4.8.5. Estimação da diferença entre médias (amostras pequenas)

Na seção anterior, sobre a estimação da *média* de um população, vimos que não há muitas diferenças entre os procedimentos para cálculo de ICs com amostras grandes ou com amostras pequenas. Na estimação da *diferença entre médias* de duas populações, porém, o procedimento para cálculo de ICs com amostras pequenas é bem mais complicado do que o com amostras grandes. Primeiro, é preciso classificar as amostras como *independentes* ou como *pareadas*. Se elas forem independentes, é preciso decidir se as variâncias das duas populações podem ser consideradas iguais ou não. Veremos abaixo como é feito o cálculo do IC nestas três situações.

4.8.5.1. Amostras pequenas independentes de populações com mesmo desvio-padrão

Se as amostras são aleatórias simples, e portanto independentes entre si, a expressão do IC para amostras pequenas é similar à do IC para amostras grandes. Para amostras grandes, vimos na seção 4.8.3.1 que o IC é baseado na diferença entre as médias das amostras, definida por:

$$D = \overline{X}_1 - \overline{X}_2$$

O IC pode ser escrito como:

$$IC^{\Delta}_{(1-\alpha)} = D \pm z_{(1-\alpha)}\sigma_D$$

onde σ_D é o desvio-padrão da diferença D (também chamado de "erro-padrão").

Se as amostras forem pequenas, o IC terá esta mesma expressão básica, com a substituição da variável z pela variável t:

$$IC^{\Delta}_{(1-\alpha)} = D \pm t_{(1-\alpha,g.l.)}\sigma_D \tag{1}$$

As fórmulas para calcular o desvio-padrão σ_D serão as mesmas que já vimos nos testes de médias com amostras pequenas (seções 4.7.1. a 4.7.3); iremos repeti-las abaixo.

Se supomos que as duas populações têm o mesmo desvio-padrão σ, podemos usar como estimador deste σ a média ponderada entre os desvios-padrões s_1 e s_2 das duas amostras. Esta média será o "desvio-padrão combinado", representado por s_{comb}, e calculado como:

$$s_{comb} = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}} \tag{2}$$

A partir de s_{comb} podemos estimar o desvio-padrão σ_D da diferença de médias por meio de:

$$\sigma_D = s_{comb}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \qquad (3)$$

Inserindo estes valores em (1), obtemos a expressão do IC:

$$IC^{\Delta}_{(1-\alpha)} = D \pm t_{(1-\alpha, g.l.)}\sigma_D$$

$$IC^{\Delta}_{(1-\alpha)} = (\overline{X}_1 - \overline{X}_2) \pm t_{(1-\alpha, g.l.)}\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \qquad (4)$$

A variável t neste caso terá o número de graus de liberdade determinado por:

$$g.l. = n_1 + n_2 - 2 \qquad (5)$$

4.8.5.2. Amostras pequenas independentes de populações com desvios-padrões diferentes

Antes de se fazer o teste do item anterior, é importante verificar se as variâncias das populações podem ser mesmo consideradas iguais, pelo menos aproximadamente. Como regra prática, alguns autores afirmam que os desvios-padrões populacionais (σ_1 e σ_2) podem ser considerados iguais se os desvios-padrões amostrais (s_1 e s_2) diferirem entre si por um fator menor do que 2 (isto é, se um deles não for mais do que o dobro do outro). É mais prudente contudo fazer antes um *teste de igualdade de vari-âncias*; ou então, fazer dois testes de diferença entre médias, um considerando que as duas populações têm mesmo desvio-padrão, como no item anterior, e outro considerando que os desvios-padrões são diferentes, como visto a seguir (alguns programas estatísticos fazem automaticamente os dois testes).

Se os dois desvios-padrões populacionais são consideradas diferentes, a expressão do IC é simples, praticamente idêntica à usada para amostras grandes (seção 4.8.3.1 (ii)); a única diferença é a substituição da variável z por t:

$$IC^{\Delta}_{0.95} = \left(\overline{X}_1 - \overline{X}_2\right) \pm t.\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \qquad (6)$$

O que se torna difícil, no entanto, é calcular o número de graus de liberdades da variável t. Este número será dado por:

$$g.l. = \frac{\left(\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}} \qquad (7)$$

4.8.5.3. Amostras pequenas pareadas

Outra forma de comparar médias de duas populações é usando amostras *pareadas*, ao invés de amostras *independentes*. Neste tipo de planejamento, as amostras têm o mesmo tamanho e cada elemento da amostra A tem seu *par* na amostra B, um elemento com características muito semelhantes. Isto é feito, por exemplo, em estudos que buscam verificar a eficácia de um tratamento (lembre-se de que a palavra "tratamento" aqui significa qualquer intervenção feita nos elementos de uma amostra; veja seção 4.7.2); por exemplo, o efeito de dois tipos de medicamento sobre pacientes portadores de uma doença. Neste caso, a cada paciente da amostra A corresponde um paciente da amostra B, geralmente um paciente do mesmo sexo, mesma idade, mesma raça, mesmo tipo de doença, etc.; ao final do experimento, se houver diferença significativa entre os resultados observados nas duas amostras, esta diferença só pode ser devida ao efeito do tratamento, já que os pacientes nas duas amostras eram praticamente iguais entre si no início do experimento.

Um tipo muito comum de estudo que usa amostras pareadas são os estudos *antes-depois*, nos quais uma variável é medida numa amostra *antes* do tratamento, e medida de novo na mesma amostra *depois* do tratamento; se houver alguma diferença significativa entre os resultados, esta diferença só pode ser devida ao tratamento, já que os pacientes são os mesmos nas duas amostras.

Os resultados podem ser organizados num quadro como este:

Quadro 1. Resultados de experimento com amostras pareadas

Par	Resposta ao Tratamento 1	Resposta ao Tratamento 2	Diferença
1	x_1	y_1	$d_1 = x_1 - y_1$
2	x_2	y_2	$d_2 = x_2 - y_2$
...	...	...	...
n	x_n	y_n	$d_n = x_n - y_n$
médias	$\overline{X}$	$\overline{Y}$	$\overline{d}$

A diferença entre as médias das populações,

$$\Delta = \mu_1 - \mu_2$$

não será agora estimada pela diferença D entre as médias amostrais

$$D = \overline{X} - \overline{Y}$$

mas sim pela média $\overline{d}$ das diferenças d calculadas *em cada par de observações*:

$$\overline{d} = \frac{1}{n}\sum d_i$$

O desvio-padrão das diferenças d será dado por:

$$s_d = \sqrt{\frac{\sum (d_i - \overline{d})^2}{n-1}}$$

A partir deste desvio-padrão s_d podemos estimar o erro-padrão da diferença média $\overline{d}$:

$$\sigma_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

O IC toma então a forma

$$IC^{\Delta}_{(1-\alpha)} = \bar{d} \pm t_{(1-\alpha, g.l.)} \sigma_{\bar{d}}$$

$$IC^{\Delta}_{(1-\alpha)} = \bar{d} \pm t_{(1-\alpha, g.l.)} \frac{s_d}{\sqrt{n}} \qquad \textbf{(8)}$$

A variável t tem neste caso o número de graus de liberdade dado por:

$$g.l. = n - 2 \qquad \textbf{(9)}$$

Um pressuposto importante, que deve que ser verificado antes do cálculo do IC, é o de que diferenças d_i tenham distribuição normal.

Exemplo 1 – Estimação com amostras pequenas independentes

Foi feita uma pesquisa para determinar o efeito no peso corporal da administração de hormônio de crescimento a ratas grávidas. Os ganhos de peso durante a gestação foram registrados para oito ratas que receberam o hormônio, e para oito que ficaram de controle. Estime por meio de um IC a diferença entre os pesos médios das ratas que receberam o hormônio e os das que não receberam.

	ganhos de peso (g)
tratamento	49 79 59 67 62 50 55 60
controle	48 49 34 43 44 34 35 45

As estatísticas amostrais são:

$n_1 = 8$ $\qquad n_2 = 8$

$\overline{X}_1 = 60{,}1$ $\qquad \overline{X}_2 = 41{,}5$

$s_1 = 9{,}72$ $\qquad s_s = 6{,}26$

O IC será baseado na diferença D encontrada entre as duas médias amostrais:

$$D = \overline{X}_1 - \overline{X}_2 = 60{,}1 - 41{,}5 = 18{,}6$$

Se consideramos que as duas populações têm o mesmo desvio-padrão σ, podemos usar a média ponderada s_{comb} como estimador deste σ:

$$s_{comb} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$s_{comb} = \sqrt{\frac{(8-1)\times 9{,}72^2 + (8-1)\times 6{,}26^2}{8+8-2}} = 8{,}18$$

O erro-padrão de D (isto é, o desvio-padrão de sua distribuição amostral) será:

$$\sigma_D = s_{comb}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\sigma_D = 8{,}18 \times \sqrt{\frac{1}{8} + \frac{1}{8}} = 4{,}09$$

O número de graus de liberdade da variável t é calculado por:

$$g.l. = n_1 + n_2 - 2$$

$$g.l. = 8 + 8 - 2 = 14$$

Para confiança de $(1 - \alpha) = 0{,}95$ e $g.l. = 14$, o valor de t dado pela tabela é:

$$t = \pm\, 2{,}14$$

Inserindo agora todos estes valores na expressão do IC dada em (**1**), obtemos o IC desejado:

$$IC^{\Delta}_{(1-\alpha)} = D \pm t_{(1-\alpha,\, g.l.)}\sigma_D$$

$$IC_{0,95}^{\Delta} = 18,6 \pm (2,14 \times 4,09)$$
$$IC_{0,95}^{\Delta} = 18,6 \pm 8,75$$
$$IC_{0,95}^{\Delta}:\ (9,9\ a\ 27,4)\, g$$

Note que esta estimativa da média da diferença de peso entre as ratas tratadas e as não tratadas não é muito precisa (a margem de erro é de 2×8,75 = 17,5 g), o que se explica pela tamanho das amostras, ambas muito pequenas. Esta estimativa, porém, mesmo que imprecisa, mostra que realmente existe uma diferença significativa entre os pesos dos dois grupos de ratas; se fizermos um teste de significância, o valor-p encontrado será de

```
p-value = 0.0004
```

o que é uma evidência muito forte de que o tratamento realmente aumenta o peso das ratas.

Os testes de normalidade das duas amostras, e também o teste de significância da diferenças, são feitos no código em R mostrado abaixo. Neste código foram feitos dois ICs, um considerando que as variâncias das duas populações sejam iguais, outro considerando que sejam diferentes; a diferença entre os ICs calculados pelos dois métodos foi pequena.

Calculo do IC usando o R

```
> trat = c(49,79,59,67,62,50,55,60)
> contr = c(48,49,34,43,44,34,35,45)
# estatisticas amostrais
> ma = mean(trat)
> mc = mean(contr)
> sda = sd(trat)
> sdc = sd(contr)
# testes de normalidade
# kolmogorov-smirnov
> mu=mean(trat)
> sig=sd(trat)
> z = (trat-mu)/sig
> ks.test(z,"pnorm")
     One-sample Kolmogorov-Smirnov test
     data:  z
     D = 0.1735, p-value = 0.9376
     alternative hypothesis: two-sided
> mu=mean(contr)
> sig=sd(contr)
> z = (contr -mu)/sig
> ks.test(z,"pnorm")
     One-sample Kolmogorov-Smirnov test
     data:  z
     D = 0.2256, p-value = 0.8103
     alternative hypothesis: two-sided
     Warning message:
     In ks.test(z, "pnorm"): não é possível calcular os
     níveis descritivos corretos com empates
# shapiro
> shapiro.test(trat)
     Shapiro-Wilk normality test
     W = 0.9315, p-value = 0.5297
```

```
> shapiro.test(contr)
      Shapiro-Wilk normality test
      W = 0.8557, p-value = 0.1086
# achando o valor de t
> n1 = n2 = length(trat)
> gl = n1+n2-2
> p = 0.025
> t = qt(p,gl)
> t
      [1] -2.144787
# Calculo do IC
> D = 60.1-41.5
> D
      [1] 18.6
> scomb = sqrt((7*9.72^2+7*6.26^2)/(8+8-2))
> scomb
      [1] 8.175145
> sigD = 8.18*sqrt(2/8)
> sigD
      [1] 4.09
> L1 = D - abs(t)*sigD
> L2 = D + abs(t)*sigD
> L1
      [1] 9.8474
> L2
      [1] 27.3526
# Fazendo o IC no R
t.test(trat,contr,var.equal = T)
      Two Sample t-test
      data:  trat and contr
      t = 4.5584, df = 14, p-value = 0.0004465
      alternative hypothesis: true difference in means is not
      equal to 0
      95 percent confidence interval:
        9.861716 27.388284
      sample estimates:
      mean of x mean of y
         60.125    41.500
t.test(trat,contr,var.equal = F)
      Welch Two Sample t-test
      data:  trat and contr
      t = 4.5584, df = 11.953, p-value = 0.0006631
      alternative hypothesis: true difference in means is not
      equal to 0
      95 percent confidence interval:
        9.718809 27.531191
      sample estimates:
      mean of x mean of y
         60.125    41.500
```

Exemplo 2. Estimação com amostras pequenas pareadas

Pesquisadores desejam determinar se um medicamento contra enxaqueca tem o efeito colateral não desejado de reduzir a pressão sanguínea diastólica dos usuários (medida em *mm/Hg*). Uma amostra de 12 pacientes foi selecionada. Primeiro, a pressão normal de cada um foi medida; em seguida, os pacientes foram submetidos ao tratamento com o medicamento e tiveram suas pressões medidas novamente. Faça um IC para a diferença entre as médias das pressões antes e depois do tratamento, e verifique se os resultados obtidos corroboram a hipótese de redução de pressão.

par	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
antes	70	80	72	76	76	76	72	78	82	64	74	92	74	68	84
depois	68	72	62	70	58	66	68	52	64	72	74	60	74	72	74
d	-2	-8	-10	-6	-18	-10	-4	-26	-18	8	0	-32	0	4	-10

Este é um exemplo do uso de amostras *pareadas*, e o IC será baseado na média das diferenças *d* observadas em cada par, calculada como:

$$d = depois - antes$$

$$\bar{d} = \frac{1}{n}\sum d_i = -8{,}8$$

O desvio-padrão de *d* é

$$s_d = \sqrt{\frac{\sum (d_i - \bar{d})^2}{n-1}} = 11{,}0$$

O número de graus de liberdade da variável *t* será:

$$g.l. = n - 1 = 15 - 1 = 14$$

Para uma confiança de $1 - \alpha = 0{,}95$, valor de $t_{(0{,}95,\ g.l.\ =\ 14)}$ fornecido pela tabela é de

$$t = \pm\ 2{,}14$$

Inserindo estes valores na expressão em (8), obtemos o IC desejado:

$$IC^{\Delta}_{(1-\alpha)} = \bar{d} \pm t_{(1-\alpha, g.l.)} \frac{s_d}{\sqrt{n}}$$

$$IC^{\Delta}_{(1-\alpha)} = -8{,}8 \pm 2{,}14 \times \frac{11{,}0}{\sqrt{15}} \quad \rightarrow \quad IC^{\Delta}_{0{,}95} : (-14{,}9 \ \ a \ -2{,}7) \quad mg\,/\,Hg$$

Como no exemplo anterior, note que esta estimativa da média da diferença entre os dois grupos também não é muito precisa (a margem de erro é de 12,9 mg/Hg). Contudo, esta estimativa já é suficiente para mostrar que realmente ocorreu uma diferença significativa entre as pressões médias dos pacientes, durante o tratamento. Se fizermos um teste de significância, o valor-p encontrado será de `p = 0.008,` o que é uma evidência forte de que o tratamento realmente faz cair a pressão dos pacientes tratados. (Este teste é feito no código em R mostrado abaixo; neste código também é feito o teste da normalidade das diferenças entre cada par).

Cálculo usando o R

```
# dados
> ant = c(70,80,72,76,76,76,72,78,82,64,74,92,74,68,84)
> dep = c(68,72,62,70,58,66,68,52,64,72,74,60,74,72,74)
> dif = dep - ant

# testes de normalidade
# kolmogorov-smirnov
> mu=mean(dif)
> sig=sd(dif)
> z = (dif-mu)/sig
> ks.test(z,"pnorm")
     One-sample Kolmogorov-Smirnov test
     data:  z
     D = 0.1898, p-value = 0.6523
     alternative hypothesis: two-sided
# shapiro
> shapiro.test(dif)
     Shapiro-Wilk normality test
     data:  dif
     W = 0.9571, p-value = 0.6428

# achando o valor de t
> n = length(dif)
> gl = n-1
> p = 0.025
> qt(p,gl)
     [1] -2.144787

# Calculo do IC à mão
> L1 = -8.8 - 2.14*11.0/sqrt(n)
> L2 = -8.8 + 2.14*11.0/sqrt(n)
> L1
     [1] -14.87800
> L2
     [1] -2.721998

# Cálculo do IC usando o R
> t.test(ant,dep,paired=T)
     data:  dep and ant
          t = -3.1054, df = 14, p-value = 0.00775
     alternative hypothesis:
          true difference in means is not equal to 0
     95 percent confidence interval:
           -14.877917  -2.722083
     sample estimates:
     mean of the differences = -8.8
```