4.8.3. Estimação de diferenças (amostras grandes)

4.8.3.1. Estimação da diferença entre médias
4.8.3.2. Estimação da diferença entre proporções

Nas seções 4.6.1 e 4.6.2, vimos os testes de hipóteses feitos para verificar se as diferenças entre médias ou proporções observadas em duas amostras grandes são ou não significativas. Veremos agora como estimar as diferenças entre as médias e as proporções nas populações, a partir do que foi observado nas amostras, por meio de intervalos de confiança.

4.8.3.1. Estimação da diferença entre médias

Para comparar as médias μ_1 e μ_2 de duas populações independentes, podemos tirar amostras (grandes) de cada um delas, e calcular a diferença D entre as médias encontradas nestas amostras:

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

Na seção 4.6.1, vimos como fazer um teste de hipótese para verificar se esta diferença D é significativa ou não. Veremos agora como usar D para estimar por meio de um intervalo de confiança a diferença entre as médias das duas populações, $\Delta = \pi_1 - \pi_2$.

(i) Distribuição amostral da diferença entre médias

Na seção 4.6.1.1, vimos um teorema que afirma que, se retirarmos amostras aleatórias grandes de duas populações independentes, a diferença D entre as médias destas amostras será uma variável aleatória cuja distribuição tende para a normal. Reproduzimos abaixo este teorema:

Teorema: *Distribuição amostral da diferença D entre médias*
Duas populações têm médias μ_1 e μ_2, cuja diferença é $\Delta = \mu_1 - \mu_2$. Se destas populações retiramos amostras aleatórias grandes de tamanhos n_1 e n_2, a diferença D entre as médias destas amostras,

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

seguirá uma distribuição que tende para a normal, com parâmetros:

$$D \rightarrow N(\mu_D, \sigma_D^2) \text{ quando } n \rightarrow \infty$$

$$\text{onde: } \mu_D = \Delta = \mu_1 - \mu_2 \quad \text{e} \quad \sigma_D = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Nas aplicações práticas, as variâncias σ das duas populações são em geral desconhecidas, e teremos que estimá-las a partir das variâncias s das amostras:

$$\sigma_D = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cong \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

(ii) Intervalo de confiança para a diferença Δ entre médias populacionais

Para estimarmos a diferença Δ por meio de um intervalo de confiança, usaremos um procedimento similar ao que já foi usado para os ICs de proporções (seção 4.8.1.2) e de médias (seção 4.8.2.2). A partir da distribuição amostral dada no teorema acima, podemos construir um intervalo em torno de Δ que tenha uma probabilidade conhecida (por exemplo, de 0,95) de conter a diferença D entre as médias das duas amostras:

$$P(\Delta - 1{,}96\sigma_D < D < \Delta + 1{,}96\sigma_D) = 0{,}95$$

Este intervalo pode ser escrito na forma:

$$\Delta \pm 1{,}96\sigma_D \;\rightarrow\; \Delta \pm 1{,}96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Da mesma forma, podemos construir um intervalo em torno de *D* que tenha uma probabilidade conhecida de conter o valor real de Δ; este intervalo é o *intervalo de confiança* para Δ, definido como:

$$IC_{\Delta}^{0,95}:\; D \pm 1{,}96\sigma_D$$

$$IC_{\Delta}^{0,95}:\; (\overline{X}_1 - \overline{X}_2) \pm 1{,}96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

(iii) Exemplos

Na seção 4.6.1.3, vimos dois exemplos de *testes de hipótese*, feitos para testar se havia diferença significativa entre as médias de duas populações. Retomaremos agora estes exemplos, e iremos estimar por meio de ICs as diferenças entre as médias destas populações.

Exemplo 1 : Peso ao nascer de crianças, de acordo com hábitos de fumo da mãe

Este exemplo usa dados de uma pesquisa feita para avaliar quais fatores podem afetar o peso ao nascer de uma criança. A Tab. **1** mostra as médias dos pesos ao nascer em duas amostras de crianças, separadas de acordo de acordo com o hábito de fumar de suas mães. Estimaremos por meio de ICs a diferença entre as médias das populações representadas por estas amostras, para avaliar quão grande é o efeito do fumo no peso do recém-nascido.

Tabela 1. Pesos ao nascer de crianças, de acordo com hábitos de fumo da mãe

	amostra	
estatística	(1) Mães não-fumantes	(2) Mães fumantes
média ($\overline{X}$)	3055 g	2773 g
desvio-padrão (s)	752 g	660 g
tamanho da amostra (n)	115	74

As estatísticas amostrais, retiradas da tabela, são:

$$\overline{X}_1 = 3055g \qquad s_1 = 752 \qquad n_1 = 115 \qquad \text{(mães não-fumantes)}$$
$$\overline{X}_2 = 2773g \qquad s_2 = 660 \qquad n_2 = 104 \qquad \text{(mães fumantes)}$$

A diferença entre as médias encontradas nas amostras foi portanto:
$$\overline{X}_1 - \overline{X}_2 = 3055 - 2773 = 282g$$

O intervalo de confiança pode ser calculado como:
$$IC_{\Delta}^{0,95}: \ (\overline{X}_1 - \overline{X}_2) \pm 1{,}96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$
$$IC_{\Delta}^{0,95}: \ (3055 - 2773) \pm 1{,}96\sqrt{\frac{752^2}{115} + \frac{660^2}{74}}$$
$$IC_{\Delta}^{0,95}: \ 282 \pm 1{,}96 \times 95{,}4$$
$$IC_{\Delta}^{0,95}: \ (95{,}0 \ \ a \ \ 469{,}0)\, g$$

Para interpretar este resultado, notamos:
- o resultado não é muito preciso (a amplitude do intervalo é grande). Para que ele fosse mais preciso, seria preciso aumentar o tamanho das amostras.
- apesar do intervalo não ser muito preciso, ele corresponde a um resultado *significativo* num teste de hipóteses. O IC abrange apenas valores positivos; isto significa que é provável que o peso médio das crianças de mães não-fumantes seja maior do que o das crianças de mães fumantes;
- o limite inferior do intervalo (95,0 g) está bem distante do zero. Isto quer dizer que um teste de hipóteses sobre estes dados conseguiria um resultado *altamente* significativo, com valor-p bem próximo de zero (da fato, no teste de hipótese feito na seção 4.6.1.3 obtivemos valor-p=0,007).

Exemplo 2 : Peso ao nascer de crianças, de acordo com seu sexo

Para verificar se o peso ao nascer da criança depende de seu sexo, estimaremos por meio de um IC a diferença entre as médias de pesos dos meninos e das meninas de uma amostra de 503 crianças. As estatísticas amostrais são mostradas na Tab. 2.

Tabela 2. Pesos ao nascer de crianças, de acordo com o sexo

	pesos ao nascer	
	(1) meninos	(2) meninas
média ($\overline{X}$)	3035.3 g	3034.6 g
desvio-padrão (s)	901.2 g	651.5 g
tamanho da amostra (n)	264	239

As estatísticas amostrais, retiradas da tabela, são:
$$\overline{X}_1 = 3035{,}3g \qquad s_1 = 901{,}2 \qquad n_1 = 264 \qquad \text{(meninos)}$$
$$\overline{X}_2 = 3034{,}6g \qquad s_2 = 651{,}5 \qquad n_2 = 239 \qquad \text{(meninas)}$$

A diferença entre as médias encontradas nas amostras foi:

$$\overline{X}_1 - \overline{X}_2 = 3035{,}3 - 3034{,}6 = 0{,}7g$$

Note que a diferença média observada entre os pesos dos meninos e das meninas foi de menos de um grama, o que não parece ser significativo. O intervalo de confiança pode ser calculado como:

$$IC_{\Delta}^{0,95} : \ (3035{,}3 - 3034{,}6) \pm 1{,}96\sqrt{\frac{901{,}2^2}{264} + \frac{651{,}5^2}{239}}$$

$$IC_{\Delta}^{0,95} : \ 0{,}7 \pm 1{,}96 \times 68{,}4$$

$$IC_{\Delta}^{0,95} : \ (-133{,}4 \ \ a \ \ 134{,}8)\, g$$

Para interpretar este resultado, notamos que o intervalo inclue o zero, e abrange tanto valores positivos quanto valores negativos; isto significa que a diferença entre as médias pesos de meninos e meninas na população talvez seja positiva (peso médio dos meninos *maior* que o das meninas), talvez negativa (peso dos meninos *menor* que o das meninas), ou talvez nula (pesos *iguais*). Não é possível dizer qual destas opções é a correta; este IC corresponde a um teste cujo resultado seja não-significativo (de fato, o teste que fizemos na seção 4.6.1.3 resultou num valor-p = 0,99).

4.8.3.2. Estimação da diferença entre proporções

Para comparar as *proporções de sucessos* (π_1 e π_2) de duas populações, podemos tirar amostras (grandes) de cada um delas, e calcular a diferença D entre as proporções P_1 e P_2 encontradas nestas amostras:

$$D = P_1 - P_2$$

Na seção 4.6.2, vimos como fazer um teste de hipótese para verificar se esta diferença D é significativa. Veremos agora como usar D para estimar a diferença Δ entre as proporções de sucessos nas duas populações,

$$\Delta = \pi_1 - \pi_2$$

por meio de um intervalo de confiança.

(i) Distribuição amostral da diferença entre proporções

Na seção 4.6.2.1 vimos um teorema que afirma que a diferença entre as proporções encontradas em duas amostras é uma variável que tende para uma distribuição normal. Reproduzimos abaixo este teorema.

> **Teorema** 1. Distribuição amostral da diferença D entre proporções amostrais
> Duas populações têm proporções de sucessos π_1 e π_2, cuja diferença é $\Delta = \pi_1 - \pi_2$. Se retiramos destas populações amostras aleatórias de tamanhos n_1 e n_2, a diferença entre as proporções de sucessos P_1 e P_2 destas amostras,

$$D = P_1 - P_2$$

terá uma distribuição que tende para a normal:

$$D \to N(\mu_D, \sigma_D^2) \qquad \text{quando } n_1 \to \infty \text{ e } n_2 \to \infty$$

com parâmetros:

$$\mu_D = \Delta = \pi_1 - \pi_2 \quad \text{e} \quad \sigma_D = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

Nas aplicações práticas, iremos precisar de estimativas iniciais das proporções de sucessos π_1 e π_2, para calcularmos σ_D; usamos como estimativas as proporções de sucessos amostrais P_1 e P_2:

$$\sigma_D = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}} \approx \sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

(ii) Intervalo de confiança para a diferença Δ entre proporções populacionais

Para construirmos um intervalo de confiança para a diferença Δ, o procedimento é o mesmo que usamos anteriormente para os ICs de proporções (seção 4.8.1.2) e de médias (seção 4.8.2.2).

A partir do teorema acima, podemos construir um intervalo em torno de Δ que tenha uma probabilidade conhecida (por exemplo, de 0,95) de conter a diferença *D* entre as proporções encontradas em duas amostras:

$$P(\Delta - 1{,}96\sigma_D < D < \Delta + 1{,}96\sigma_D) = 0{,}95$$

ou seja, o intervalo

$$\Delta \pm 1{,}96\sigma_D \;\to\; \Delta \pm 1{,}96\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

Da mesma forma, podemos construir um intervalo em torno de *D* que tenha uma probabilidade conhecida de conter o valor real de Δ; este intervalo será o intervalo de confiança para Δ, definido como:

$$IC_\Delta^{0,95}: \; D \pm 1{,}96\sigma_D$$

$$IC_\Delta^{0,95}: \; (P_1 - P_2) \pm 1{,}96\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

(iii) Exemplo

A Tab. **1** mostra os pesos na nascer de 189 crianças, classificadas em duas classes conforme o peso ao nascer, *normal* ou *baixo* (menos de 2500 g). Vemos que, nestas amostras, as proporções de crianças com baixo peso é diferente de acordo com a raça da mãe: as mães brancas tiveram uma proporção menor do que as mães não-brancas. Na seção 4.6.2.3, fizemos um teste de hipótese para verificar se a diferença entre estas proporções amostrais é significativa.

Tabela 1. Peso ao nascer da criança (normal / baixo) pela raça da mãe

	raça da mãe				Total
	(1) brancas		(2) não-brancas		
peso ao nascer	f	fr	f	fr%	
normal	73	0,760	57	0,613	130
baixo	23	0,240	36	0,387	59
Total	96	1,000	93	1,000	189

Iremos agora estimar por meio de um IC a diferença entre as proporções de crianças com baixo peso nas duas raças. Ou seja, estimaremos

$$\Delta = \pi_1 - \pi_2$$

onde π_1 : proporção, na população de mães brancas, de crianças nascidas com baixo peso
π_2 : proporção, na população de mães não-brancas, de crianças com baixo peso

Os dados são:

$n_1 = 96$ $P_1 = 0{,}24$ (mães brancas)
$n_2 = 93$ $P_2 = 0{,}387$ (mães não-brancas)

O intervalo de confiança pode ser calculado como:

$$IC_{\Delta}^{0,95}: \ (P_1 - P_2) \pm 1{,}96\sqrt{\frac{P_1(1-P_1)}{n_1} + \frac{P_2(1-P_2)}{n_2}}$$

$$IC_{\Delta}^{0,95}: \ (0{,}24 - 0{,}387) \pm 1{,}96\sqrt{\frac{0{,}24(1-0{,}24)}{96} + \frac{0{,}387(1-0{,}387)}{93}}$$

$$IC_{\Delta}^{0,95}: \ -0{,}147 \pm 0{,}131$$

$$IC_{\Delta}^{0,95}: \ -0{,}278 \ \ a \ \ -0{,}016$$

Note que o intervalo só contem valores negativos; segundo esta estimativa, portanto, a proporção π_1 de mães brancas na população que têm crianças de baixo peso é realmente menor do que a proporção π_2 de mães não-brancas. Isto corrobora o resultado do teste de hipótese feito na seção 4.6.2.3, que teve um resultado significativo.

Note porém que o limite superior deste intervalo (– 0,016) está bem próximo de zero. Quando o intervalo contem o zero, isto significa que há uma boa probabilidade de que as duas proporções populacionais sejam iguais ($\pi_1 = \pi_2$); e o intervalo corresponde a um teste com resultado não-significativo. O resultado encontrado no exemplo, portanto, é significativo mas não muito; está bem próximo de ser não-significativo (de fato, no teste de hipótese mencionado acima obtivemos um valor-p=0,0139; significativo se considerarmos um nível de significância $\alpha = 0{,}01$, mas não se considerarmos $\alpha = 0{,}05$).