

## **CAPÍTULO 3. O ESTUDO DE CARACTERES QUANTITATIVOS EM GÊMEOS**

A utilização de gêmeos em estudos para avaliar a importância do genótipo na determinação de caracteres quantitativos baseia-se, fundamentalmente, no fato de que a similaridade genotípica dos pares DZ é, em média, aquela apresentada por pares de irmãos gerados sucessivamente, pois, do mesmo modo que eles, são oriundos da união de pares de gametas distintos, produzidos pelos mesmos genitores. Essa é a razão pela qual os pares DZ têm, em média, metade de seus genes idênticos, por terem origem comum. No caso dos gêmeos MZ, cada par é oriundo de um único zigoto, de sorte que os elementos de um par MZ são geneticamente idênticos, isto é, possuem o mesmo genótipo. Aliás, na espécie humana, os pares MZ constituem a única série isogênica.

Nos estudos gemelares para avaliar a importância do genótipo na determinação do fenótipo as seguintes premissas são aceitas, desde que não haja distorções de averiguação:

1. Os gêmeos são uma amostra da população geral, o que equivale a dizer que entre os gêmeos todos os caracteres se distribuem da mesma maneira que na população geral.
2. Os gêmeos são uma amostra de todos os gêmeos.
3. Os elementos de cada par de gêmeos estão sujeitos às mesmas influências do ambiente.
4. O meio ambiente dos gêmeos é, em média, igual ao dos elementos da população geral.
5. As variáveis que atuam sobre os gêmeos MZ, provocando diferenças fenotípicas intrapar, são as mesmas que agem sobre os pares DZ.

Ao aceitar essas premissas, aceita-se, também, que, nos pares MZ, as diferenças intrapar de caracteres quantitativos são determinadas unicamente pelos fatores do ambiente, ao passo que as diferenças intrapar observadas nos pares DZ são causadas pelos fatores responsáveis pelas diferenças nos pares MZ *acrescidas da diversidade genotípica que existe em pares de irmãos*, os quais diferem, em média, em metade de seus genes.

A premissa de que os elementos de cada par de gêmeos estão sujeitos ao mesmo ambiente baseia-se no fato de que, na fase pré-natal, os gêmeos recebem as mesmas

influências da idade materna, da paridade, das condições de saúde da gestante e do efeito da ingestão de medicamentos e outras substâncias. Como se vê, esse postulado não leva em conta que, na fase intra-uterina, é possível que um dos elementos de um par MZ receba suprimento de sangue maior do que o outro ou que, em pares MZ, possa haver repartição desigual do material embrionário que lhes deu origem. Além disso, é possível que os elementos de um par de gêmeos, MZ ou DZ, tenham espaços desiguais para o seu desenvolvimento. Ainda que tais influências possam ser minimizadas, está fora de dúvida que a diversificação do ambiente dos gêmeos se acentua com a idade, sendo maior na idade adulta.

A premissa de que o ambiente dos gêmeos é, em média, igual ao dos elementos da população geral também é passível de críticas, porque os gêmeos estão sujeitos a condições mais adversas do que os conceptos únicos, não só durante a gestação mas, também, no momento do parto e no período neonatal. Isso aliás, contribui para o fato de os gêmeos apresentarem maior taxa de mortalidade perinatal do que os recém-nascidos de parto único (Beiguelman e Franchi-Pinto, 2000).

Também é passível de críticas a premissa de que as variáveis do ambiente que atuam sobre os pares MZ são as mesmas que agem sobre os pares DZ, porque, em conseqüência de sua maior similaridade, os pares MZ tendem a ser tratados de modo mais semelhante do que os DZ, pelo menos na infância. Por outro lado, visto que o nascimento de gêmeos DZ está correlacionado à idade materna e que o aumento da idade das mães está associado a condições de gestação que são, em média, piores do que as de mães jovens, e a maior risco de anomalias congênitas decorrentes ou não de aberrações cromossômicas, pode-se supor que o ambiente intra-uterino é mais adverso para uma parcela dos gêmeos DZ do que para os MZ.

Essas últimas críticas têm sido refutadas com base em dados que permitem aceitar serem os fatores de distorção apontados contrabalançados por outros, que tendem a acentuar as diferenças entre os elementos de uma parte dos pares MZ e a diminuir as existentes entre os elementos dos pares DZ, mormente se apresentarem quimerismo.

## O COEFICIENTE DE CORRELAÇÃO INTRACLASSE

Para avaliar a participação do genótipo na variação quantitativa, por intermédio do estudo de gêmeos, é importante o cálculo do *coeficiente de correlação intraclasse* ( $r_i$ ) das medidas feitas nos pares MZ e DZ. Essa é a razão pela qual teceremos algumas considerações sucintas a respeito do seu cálculo. O leitor interessado nas deduções das fórmulas aqui apresentadas poderá consultar outra obra do autor (Beiguelman, 1994).

No modelo teórico para o cálculo do coeficiente de correlação intraclasse, a respeito de uma variável  $x$  com distribuição normal, presume-se que lidamos com uma população de pares, em relação aos quais não se reconhece uma ordem *entre* nem *dentro* desses pares. Os gêmeos se adaptam bem a esse modelo quando lidamos com amostras do mesmo grupo etário e do mesmo sexo, extraídas de uma única população. Desse modo, além de não existirem critérios para classificá-los segundo a idade ou procedência, também não será possível designar qual é o gêmeo número 1 e qual o número 2 de cada par, como se faz, normalmente, em pares como pai-filho, marido-esposa, irmão mais velho-irmão mais novo. Evidentemente, se um caráter depender da idade e os gêmeos não pertencerem a um mesmo grupo etário, poderemos, por análise de regressão, fazer o ajustamento dos dados para idade, anulando, assim, seu efeito. Se, além da idade, o caráter depender do sexo, também se poderá fazer o ajustamento para essa variável, a fim de anular seu efeito.

Tendo em vista o exposto, quando estudarmos um caráter quantitativo em gêmeos e distinguirmos, em cada par, uma medida  $x$  e outra  $x'$ , isso será feito meramente para diferenciar as medidas de cada par, mas não para indicar uma ordem dentro dos pares. É por isso, também, que não se deve falar em desvio padrão de  $x$ , isto é,  $s_x$ , e desvio padrão de  $x'$ , ou seja,  $s_{x'}$ , porque, por hipótese,  $s_x$  não pode diferir de  $s_{x'}$ . Em consequência, no caso do coeficiente de correlação intraclasse, sua fórmula geral não pode ser escrita como

$$r_i = \frac{COV_{xx'}}{s_x \cdot s_{x'}} \text{ e sim como } r_i = \frac{COV_{xx'}}{s_x^2}.$$

Designando a variância entre os pares por  $s_E^2$  e a variância dentro dos pares por  $s_D^2$ , e lembrando que  $cov_{xx'} = \frac{s_E^2 - s_D^2}{2}$  enquanto  $s_x^2 = \frac{s_E^2 + s_D^2}{2}$ , fica claro que o coeficiente de correlação intraclasse pode ser calculado por intermédio de  $r_i = \frac{s_E^2 - s_D^2}{s_E^2 + s_D^2}$ .

Para obtermos as fórmulas de cálculo das variâncias entre e dentro dos pares, consideremos que retiramos de uma população  $a$  amostras de  $n = 2$  indivíduos não ordenados. Cada par fornecerá uma média  $\bar{x} = \frac{x+x'}{2}$ . A média geral ( $\bar{X}$ ) pode ser obtida de várias maneiras, pois  $\bar{X} = \frac{\sum(x+x')}{2a} = \frac{\sum x + \sum x'}{2a} = \frac{\sum \bar{x}}{a}$ . Nessas fórmulas, e nas seguintes, o símbolo de somatório ( $\Sigma$ ) significa, na realidade,  $\sum_{i=1}^a$  ou seja, a soma de todas as amostras, desde a primeira até a última  $a$  amostra, cada amostra  $a$  sendo constituída por  $n = 2$  indivíduos, isto é, um par de gêmeos.

Visto que  $n = 2$  gêmeos de um par, as fórmulas gerais das somas de quadrados entre ( $SQ_E$ ) e dentro ( $SQ_D$ ) poderão ser escritas como abaixo:

$$SQ_E = n \sum (\bar{x} - \bar{X})^2 = 2 \sum (\bar{x} - \bar{X})^2$$

$$SQ_D = \sum [(x - \bar{x})^2 + (x' - \bar{x})^2]$$

as quais, depois de desenvolvidas, fornecem as seguintes fórmulas de cálculo:

$$SQ_E = \frac{\sum (x+x')^2}{2} - \frac{(\sum x + \sum x')^2}{2a}$$

$$SQ_D = \frac{\sum (x-x')^2}{2}$$

A soma de quadrados total pode ser obtida a partir da adição  $SQ_E + SQ_D$  ou a partir de  $SQ_T = \sum (x - \bar{X})^2 + \sum (x' - \bar{X})^2$  que, depois de desenvolvida fornece a seguinte fórmula de cálculo:

$$SQ_T = \sum x^2 + \sum x'^2 - \frac{(\sum x + \sum x')^2}{2a}$$

Sabendo-se que  $SQ_E$  tem  $a-1$  graus de liberdade, que  $SQ_D$  tem  $(2-1)a$ , isto é,  $a$  graus de liberdade e que  $SQ_T$  tem  $2a-1$  graus de liberdade, as estimativas das variâncias entre, dentro e total são calculadas, respectivamente, a partir de  $s_E^2 = \frac{SQ_E}{a-1}$ ,  $s_D^2 = \frac{SQ_D}{a}$  e  $s_T^2 = \frac{SQ_T}{2a-1}$ .

Para sabermos se o coeficiente de correlação intraclasse se desvia significativamente de zero, calcula-se  $F = \frac{s_E^2}{s_D^2}$ , que é comparado ao F crítico com  $a-1$  e  $a$  graus de liberdade.

Para exemplificar, consideremos os dados da Tabela 1.3, a respeito das linhas dermatoglíficas palmares A'-d tomadas em 10 pares de gêmeos MZ do sexo masculino, todos de escolas primárias de Campinas, SP, caucasóides e do mesmo grupo etário (Beiguelman, 1971).

Tabela 1.3 Número de linhas dermatoglíficas palmares A'-d contadas em 10 pares de gêmeos monozigóticos.

| Par | $x$ | $x'$ |
|-----|-----|------|
| 1   | 106 | 111  |
| 2   | 111 | 117  |
| 3   | 104 | 104  |
| 4   | 113 | 111  |
| 5   | 102 | 104  |
| 6   | 109 | 111  |
| 7   | 155 | 143  |
| 8   | 114 | 112  |
| 9   | 103 | 114  |
| 10  | 102 | 104  |

Designando o primeiro valor de cada par de dados por  $x$  e o segundo por  $x'$ , calculamos, inicialmente os valores abaixo:

$$\Sigma x = 1.119 \quad \Sigma x' = 1.131 \quad \Sigma x + \Sigma x' = 2.250 \quad \Sigma x^2 = 127.461 \quad \Sigma x'^2 = 129.089$$

$$\frac{(\Sigma x + \Sigma x')^2}{2a} = 253.125 \quad \frac{\Sigma (x + x')^2}{2} = 256.377 \quad \frac{\Sigma (x - x')^2}{2} = 173$$

Com esses dados podemos organizar um quadro de análise da variância como o da Tabela 2.3, que serve para calcular o coeficiente de correlação intraclasse. De acordo com

os dados da Tabela 2, a contagem de linhas A'-d nos gêmeos MZ possuem coeficiente de correlação alto ( $r_i = 0,91$ ) e significativamente diferente de zero, pois  $F_{(9;10)} = 20,89$  tem  $P < 0,005$ .

Tabela 2.3 Cálculo do coeficiente de correlação intraclasse entre o número de linhas dermatoglíficas palmares A'-d em 10 pares de gêmeos monozigóticos.

| Variação | G.L. | SQ   | $s^2$  | $r_i$ | $F_{(9;10)}$ |
|----------|------|------|--------|-------|--------------|
| Entre    | 9    | 3252 | 361,33 | 0,91  | 20,89        |
| Dentro   | 10   | 173  | 17,30  |       | P<0,005      |
| Total    | 19   | 3425 | 180,26 |       |              |

As estimativas das variáveis necessárias ao cálculo do coeficiente de correlação intraclasse, inclusive a própria estimativa desse coeficiente, podem vir acompanhadas dos respectivos desvios padrão. De acordo com Kempthorn e Osborne (1961), o desvio padrão

da variância *entre* é obtido por intermédio de  $s_E^2 \sqrt{\frac{2}{a+1}}$  e o desvio padrão da variância

*dentro* é obtido por meio de  $s_D^2 \sqrt{\frac{2}{a+2}}$ . Os desvios padrão da covariância e da variância

total são idênticos e fornecidos por  $\sqrt{\frac{1}{2} \left[ \frac{(s_E^2)^2}{a+1} + \frac{(s_D^2)^2}{a+2} \right]}$ , ao passo que o desvio padrão do

coeficiente de correlação intraclasse é obtido pela resolução de  $\frac{s_E^2 \cdot s_D^2}{(s_E^2 + s_D^2)^2} \sqrt{\frac{8(2a+3)}{(a+1)(a+2)}}$ .

Em relação aos dados da Tabela 2 poderíamos, pois, escrever:

$$s_E^2 = 361,33 \pm 154,07$$

$$s_D^2 = 17,30 \pm 7,06$$

$$\text{cov}_{xx'} = 172,02 \pm 77,12$$

$$s_x^2 = 180,26 \pm 77,12$$

$$r_i = 0,91 \pm 0,05$$

## A COMPARAÇÃO DE CARACTERES QUANTITATIVOS EM GÊMEOS

Se uma amostra de gêmeos MZ for comparada a uma amostra de pares DZ, em relação a um caráter quantitativo, e se os gêmeos em estudo pertencerem ao mesmo grupo etário, tiverem sido extraídos de uma população homogênea, não estiverem sujeitos a um processo seletivo diferencial, nem forem fruto de casamentos preferenciais, a variância total desse caráter nos pares MZ não deverá diferir significativamente daquela observada nos pares DZ. Se o caráter variar segundo o sexo deveremos obter:

$$s_x^2 \text{MZMM} = s_x^2 \text{DZMM} \quad \text{e} \quad s_x^2 \text{MZFF} = s_x^2 \text{DZFF}$$

Evidentemente, se isso não ocorrer, não deverá haver diferenças significativas entre essas quatro variâncias, isto é:

$$s_x^2 \text{MZMM} = s_x^2 \text{DZMM} = s_x^2 \text{MZFF} = s_x^2 \text{DZFF}$$

Por outro lado, se o caráter em estudo nos gêmeos depender de um componente genético importante, a variância dentro dos pares MZ será significativamente menor do que a dos pares DZ, tanto entre os pares masculinos quanto entre os femininos, isto é,

$$s_D^2 \text{MZMM} < s_D^2 \text{DZMM} \quad \text{e} \quad s_D^2 \text{MZFF} < s_D^2 \text{DZFF}$$

A covariância e o coeficiente de correlação intraclasse mostrarão uma relação oposta à da variância dentro, isto é, nos pares MZ seus valores serão significativamente maiores do que nos pares DZ, o que permite escrever:

$$\begin{aligned} \text{COV}_{\text{MZMM}} > \text{COV}_{\text{DZMM}} \quad \text{e} \quad \text{COV}_{\text{MZFF}} > \text{COV}_{\text{DZFF}} \\ r_{i\text{MZMM}} > r_{i\text{DZMM}} \quad \text{e} \quad r_{i\text{MZFF}} > r_{i\text{DZFF}} \end{aligned}$$

Para exemplificar, consideremos os dados relacionados na Tabela 3.3, a respeito das linhas *a-b* contadas nas palmas das duas mãos de 38 pares MZ (20 do sexo masculino e 18 do sexo feminino) e 52 pares DZ (28 do sexo masculino e 24 do sexo feminino), todos do mesmo grupo etário, caucasóides e do mesmo nível social, averiguados em escolas primárias de Campinas, SP (Beiguelman, 1971).

Tabela 3.3 Coeficientes de correlação intraclasse do número de linhas dermatoglíficas *a-b* de 90 pares de gêmeos e elementos utilizados para o seu cálculo.

| Estimativa         | MZ              | DZ              |
|--------------------|-----------------|-----------------|
| $s_E^2$            | 207,54 ± 46,998 | 140,53 ± 27,299 |
| $s_D^2$            | 13,08 ± 2,925   | 65,88 ± 12,679  |
| cov <sub>xx'</sub> | 97,23 ± 23,545  | 37,32 ± 15,050  |
| $s_x^2$            | 109,01 ± 23,545 | 102,84 ± 15,050 |
| $r_i$              | 0,88 ± 0,036    | 0,36 ± 0,119    |
| <i>a</i>           | 38              | 52              |

Visto que o número médio de linhas *a-b* nos indivíduos do sexo masculino (81,6;  $s = 10,94$ ) não diferiu significativamente do número médio dessas linhas nos indivíduos do sexo feminino (80,7;  $s = 9,43$ ), os gêmeos MZ foram comparados aos DZ sem distinção de sexo. Na Tabela 3.3 é fácil constatar que a variância total dos pares MZ não diferiu significativamente daquela observada nos DZ ( $F_{(75;103)} = \frac{109,01}{102,84} = 1,06$ ;  $P > 0,05$ ). Já a variância dentro dos pares DZ foi significativamente maior do que a estimada para os pares MZ ( $F_{(54;38)} = \frac{65,88}{13,08} = 5,04$ ;  $P < 0,05$ ). Em oposição, tanto a covariância quanto o coeficiente de correlação intraclasse foram maiores nos pares MZ do que nos DZ em ambos os sexos.

Para averiguar se a covariância dos gêmeos MZ é significativamente maior do que a dos DZ, calcula-se um *t* obtido pela divisão da diferença entre as covariâncias pelo desvio padrão dessa diferença. Com base nos dados da Tabela 3.3 tem-se, portanto,

$$t = \frac{97,23 - 37,32}{\sqrt{\frac{(23,545)^2}{75} + \frac{(15,050)^2}{103}}} = 19,345; \text{ G.L.} = 178; P < 0,001.$$

Para investigar se o coeficiente de correlação intraclasse dos gêmeos MZ é significativamente maior do que o dos DZ, pode-se empregar um método de Snedecor

(1966), que requer o cálculo da quantidade  $z = \frac{1}{2} \ln \left( \frac{1+r_i}{1-r_i} \right)$  para os pares MZ ( $z_1$ ) e para os pares DZ ( $z_2$ ). No exemplo da tabela 3.3 tem-se que  $r_i = 0,88$  dá origem a  $z_1 = 1,3758$  para os gêmeos MZ, enquanto que  $z_2$  é igual a 0,3769 para os gêmeos DZ. A razão entre a diferença  $z_1 - z_2$  e o desvio padrão dessa diferença fornecerá um  $t$  com infinitos graus de liberdade. Sabendo-se que esse desvio padrão é obtido por intermédio de  $s_{z_1} - s_{z_2} = \sqrt{\frac{1}{a_1 - 3} + \frac{1}{a_2 - 3}}$  onde  $a_1$  é o número de pares MZ e  $a_2$  é o número de pares DZ,

podemos calcular  $t = \frac{1,3758 - 0,3769}{\sqrt{\frac{1}{35} + \frac{1}{49}}} = 4,516$ ; G.L. = 84;  $P < 0,001$ .

Em resumo, os dados da Tabela 3.3 permitem afirmar que o número de linhas dermatoglíficas palmares  $a-b$  depende, indubitavelmente, de um componente genético importante, pois os pares MZ mostram maior coeficiente de correlação intraclasse do que os DZ e todos os outros elementos que serviram para o cálculo desses coeficientes de correlação são coerentes com a hipótese da existência desse componente.

### **INVESTIGAÇÃO DA HERDABILIDADE DE CARACTERES QUANTITATIVOS EM GÊMEOS**

Os coeficientes de correlação intraclasse calculados para os pares MZ e DZ estimam qual a proporção da variância fenotípica total ( $V$ ) que é comum aos membros desses pares. Assim, designando a variância aditiva por  $V_A$ , a variância dominante por  $V_D$  e a variância devido a efeitos familiares não-genéticos sobre os gêmeos por  $V_{NG}$ , e desprezando a variância epistática devido à interação, ter-se-á, de acordo com Falconer (1960) que:

$$r_{iMZ} = \frac{V_A + V_D + V_{NG}}{V}$$

$$r_{iDZ} = \frac{0,5V_A + 0,25V_D + V_{NG}}{V}$$

Simbolizando a variância genética total por  $V_G$  que, evidentemente, inclui a variância aditiva e a dominante, pode-se escrever, também,

$$r_{iMZ} = \frac{V_G + V_{NG}}{V}$$

Se a  $V_{NG}$  fosse nula,  $r_{iMZ}$  estimaria a razão  $\frac{V_G}{V}$ , que é denominada *coeficiente de determinação genética* por Smith (1974) e simbolizada por G. Essa razão tem sido chamada por outros autores, como Mather e Jinks (1984), de *herdabilidade sensu lato*, para diferenciá-la da verdadeira herdabilidade (*herdabilidade sensu stricto*,  $h^2$ ), que mede a *porção da variância fenotípica devida à variância aditiva*, isto é,  $\frac{V_A}{V}$ . No caso de  $V_{NG}$  e  $V_D$  terem valores muito baixos, tanto  $r_{iMZ}$  quanto  $2r_{iDZ}$  poderiam estimar a herdabilidade, pois, nessa situação,  $r_{iMZ} = 2r_{iDZ} = \frac{V_A}{V} = h^2$ . Entretanto, pelo fato de não conhecermos  $V_D$  nem  $V_{NG}$ , as estimativas da herdabilidade poderão ser bastante distorcidas. Por isso, para eliminar os efeitos familiares não genéticos, Smith (1974) aceitou que eles são os mesmos para os pares MZ e DZ, e recomenda o cálculo da expressão  $2(r_{iMZ} - r_{iDZ})$ , que estima  $\frac{V_A + 1,5V_D}{V}$ , porque:

$$\frac{2(V_A + V_D + V_{NG} - 0,5V_A - 0,25V_D - V_{NG})}{V} = \frac{V_A + 1,5V_D}{V}$$

Evidentemente, essa fórmula de Smith (1974) também não é boa, pois ela dará uma super-estimativa do coeficiente de determinação genética (G) se  $V_D$  ou a variação epistática forem importantes, e se aproximará da herdabilidade se  $V_D$  for pequena.

Uma postura mais razoável, embora pouco rigorosa, é a de aceitar que a herdabilidade de um caráter quantitativo estimada pelo estudo de gêmeos deve estar por volta de  $r_{iMZ}$  ou  $2r_{iDZ}$ . Assim, os dados da Tabela 3.3 permitiriam dizer que a herdabilidade ( $h^2$ ) do número de linhas dermatoglíficas palmares *a-b* deveria estar entre 0,72 e 0,88 pois  $2r_{iDZ} = 0,72$  e  $r_{iMZ} = 0,88$ .

Em 1929 Holzinger propôs uma fórmula para estimar a herdabilidade de caracteres quantitativos pelo estudo de gêmeos, a qual tem sido amplamente utilizada, apesar das críticas que sofreu. Essa fórmula é expressa do seguinte modo:

$$\frac{r_{iMZ} - r_{iDZ}}{1 - r_{iDZ}}$$

As críticas a essa fórmula têm sentido, porque se  $r_{iMZ}$  e  $r_{iDZ}$  forem substituídos, respectivamente, por  $\frac{V_A + V_D + V_{NG}}{V}$  e  $\frac{0,5V_A + 0,25V_D + V_{NG}}{V}$ , ela passará a significar  $\frac{0,5V_A + 0,75V_D}{V - (0,5V_A + 0,25V_D + V_{NG})}$ , não estimando, pois a herdabilidade.

Por causa dessas objeções, poder-se-ia, dizer que a fórmula clássica de Holzinger (1929) estima uma herdabilidade aproximada, a qual poderia ser denominada “herdabilidade” ou “ $h^2$ ”, devendo ela ser aplicada separadamente para gêmeos do sexo masculino e do sexo feminino, no caso de os caracteres em estudo sofrerem influência do sexo. Se aplicarmos essa fórmula aos dados da Tabela 3.3, poderemos dizer que a “herdabilidade” do número de linhas  $a-b$  é 0,81, porque “ $h^2$ ” =  $\frac{0,88 - 0,36}{1 - 0,36} = 0,81$ .

Se a variância total dos pares MZ puder ser considerada como idêntica à dos pares DZ, por não haver diferença significativa entre as suas estimativas, como é o caso dos dados da Tabela 3.3, a fórmula da “herdabilidade” de Holzinger pode ser transformada numa outra, que leva em conta apenas as variâncias dentro dos pares MZ e DZ de cada sexo, e que pode ser apresentada como abaixo:

$$“h^2” = \frac{s_{DDZ}^2 - s_{DMZ}^2}{s_{DDZ}^2}$$

De fato, partindo de  $r_i = \frac{s_E^2 - s_D^2}{s_E^2 + s_D^2}$ , podemos escrever:

$$1 - r_i = 1 - \frac{s_E^2 - s_D^2}{s_E^2 + s_D^2} = \frac{2s_D^2}{s_E^2 + s_D^2}$$

Essa expressão não se alterará se dividirmos por 2 tanto o numerador quanto o denominador, o que conduz a:

$$1 - r_i = \frac{s_D^2}{\frac{s_E^2 + s_D^2}{2}} = \frac{s_D^2}{s_T^2}$$

que permite escrever:

$$r_i = 1 - \frac{s_D^2}{s_T^2}$$

Se a variância total dos gêmeos MZ for igual à dos gêmeos DZ, poder-se-á escrever:

$$"h^2" = \frac{r_{iMZ} - r_{iDZ}}{1 - r_{iDZ}} = \frac{\left(1 - \frac{s_{DMZ}^2}{s_T^2}\right) - \left(1 - \frac{s_{DDZ}^2}{s_T^2}\right)}{1 - \left(1 - \frac{s_{DDZ}^2}{s_T^2}\right)} = \frac{s_{DDZ}^2 - s_{DMZ}^2}{s_{DDZ}^2}$$

Essa última fórmula, preferida por muitos autores, como Clark (1956), deixa claro que a “herdabilidade” estimada pela fórmula de Holzinger poderia ser definida como *a proporção da variância dentro dos gêmeos DZ que deve ser atribuída a fatores genéticos*. Aplicando essa última fórmula aos dados sobre a contagem das linhas *a-b* (Tabela 3.3), podemos constatar que o valor estimado é da mesma ordem que o obtido pela fórmula que leva em conta os coeficientes de correlação intraclasse dos pares MZ e DZ, pois “ $h^2$ ” =  $\frac{65,88 - 13,08}{65,88} = 0,80$ .

Em qualquer caso, a demonstração de que a “herdabilidade” difere significativamente de zero será feita por um F obtido pela divisão da variância dentro dos pares DZ pela variância dentro dos pares MZ. Nos dados da Tabela 3.3 tem-se, pois,  $F_{(52;38)} = 5,04$ ;  $P < 0,005$  o qual indica que a “herdabilidade” das linhas *a-b* difere significativamente de zero nos indivíduos de ambos os sexos.

Uma estimativa aproximada da herdabilidade de caracteres quantitativos em amostras de gêmeos sem diagnóstico da zigossidade pode ser obtida a partir de uma fórmula sugerida pelos Professores Newton E. Morton e Henrique Krieger (Franchi-Pinto, Dal Colletto, Krieger e Beiguelman, 1999). Para isso, leva-se em conta que, em uma amostra de gêmeos, os pares discordantes quanto ao sexo (MF) são, seguramente, DZ. Se o número desses pares for subtraído do número de pares de gêmeos do mesmo sexo (MM + FF), ter-se-á, de acordo com a regra clássica de Weinberg (1901) que, o número restante estimará o número de pares MZ. Assim, por exemplo, em uma amostra de gêmeos com 250 MM, 192 MF e 256 FF, a estimativa do número de pares MZ seria 314, porque  $(250+256) - 192 = 314$ . Esses dados, por sua vez, permitiriam estimar a proporção de pares do mesmo sexo que são MZ em 62% porque  $\frac{314}{506} = 0,62$ .

Se aceitarmos que o coeficiente de correlação intraclasse de um caráter quantitativo em pares MZ estima, aproximadamente, a herdabilidade, isto é, que  $r_{iMZ} = h^2$ , enquanto

entre os pares DZ esse coeficiente estima, aproximadamente, a metade da herdabilidade, isto é,  $r_{iDZ} = \frac{h^2}{2}$ , os coeficientes de correlação intraclasse de pares concordantes ( $r_c$ ) e de discordantes quanto ao sexo ( $r_d$ ) poderiam ser expressos como abaixo, onde  $c$  é a proporção de pares MZ entre os gêmeos do mesmo sexo:

$$r_c = ch^2 + \frac{(1-c)h^2}{2}$$

$$r_d = \frac{h^2}{2}$$

Em vista do exposto, pode-se escrever a seguinte equação:

$$r_c + r_d = ch^2 + \frac{(1-c)h^2}{2} + \frac{h^2}{2}$$

que pode ser convertida na seguinte expressão:

$$h^2 = \frac{r_c + r_d}{1 + \frac{c}{2}}$$

Num estudo de gêmeos realizado por Franchi-Pinto, Dal Colletto, Krieger e Beiguelman (1999) nos quais o índice de Apgar foi ajustado para a idade gestacional e sexo, constatou-se que esse índice obtido no primeiro minuto após o nascimento mostrou  $r_c = 0,383$  e  $r_d = 0,169$ . Visto que, nesse estudo,  $c$  foi estimado em 0,62, isso permitiu estimar  $h^2$  como sendo igual a 0,42

Do exposto neste capítulo pode-se concluir que os estudos de gêmeos são ótimos para pôr à prova a hipótese de que um caráter quantitativo depende de um componente genético importante. Contudo, quando se trata de calcular a herdabilidade de um desses caracteres, eles são menos precisos do que os estudos familiares. Realmente, em estudos familiares podemos estimar o valor da herdabilidade por intermédio da regressão das médias das medidas do caráter quantitativo feitas nos filhos e nos genitores. Quando o número de famílias não é grande e existe muita variabilidade em relação ao tamanho das irmandades, pode ser desejável que a obtenção das médias das medidas tomadas nas irmandades seja ponderada com o número de elementos que ela contém (Kempthorne e Tandon, 1953; Reeve, 1955).

Se a regressão for feita entre a média das medidas obtidas em cada irmandade ( $\bar{F}$ )

sobre um dos genitores, isto é, sobre o pai ( $P$ ) ou sobre a mãe ( $M$ ), os coeficientes de regressão  $b_{\bar{F}P}$  ou  $b_{\bar{F}M}$  estimarão, cada qual a metade da herdabilidade, isto é,  $b_{\bar{F}P} = \frac{h^2}{2}$  e  $b_{\bar{F}M} = \frac{h^2}{2}$ , pois cada um dos genitores transmite a metade de seus genes à sua prole. Por outro lado, se a regressão for feita entre a média das medidas tomadas em cada irmandade e a média das medidas verificadas nos genitores ( $\bar{G}$ ), o coeficiente de regressão  $b_{\bar{F}\bar{G}}$  estimará a herdabilidade, isto é,  $b_{\bar{F}\bar{G}} = h^2$ . Isso permite escrever:

$$h^2 = 2 b_{\bar{F}P}$$

$$h^2 = 2 b_{\bar{F}M}$$

$$h^2 = b_{\bar{F}\bar{G}}$$

Apesar de qualquer um desses coeficientes de regressão servir para estimar a herdabilidade, considera-se que a melhor estimativa desse valor é obtida por intermédio de uma combinação dos três coeficientes. Para isso, pondera-se cada coeficiente com a *quantidade de informação da estimativa*, a qual é a invariância ( $I$ , isto é, o inverso do quadrado do desvio padrão do coeficiente de regressão). Desse modo, pode-se escrever

$I = \frac{1}{s_b^2}$  e expressar a fórmula da estimativa combinada da herdabilidade como abaixo:

$$h^2 = \frac{2b_{\bar{F}P} \cdot I_{\bar{F}P} + 2b_{\bar{F}M} \cdot I_{\bar{F}M} + b_{\bar{F}\bar{G}} \cdot I_{\bar{F}\bar{G}}}{I_{\bar{F}P} + I_{\bar{F}M} + I_{\bar{F}\bar{G}}}$$

Nos estudos familiares, o coeficiente de regressão das medidas tomadas nas mulheres sobre as obtidas nos maridos ( $b_{PM}$ ) permite, por sua vez, verificar se o meio comum exerce influência importante sobre o caráter quantitativo em estudo ou se existem casamentos preferenciais em relação a esse caráter. Assim, se esse coeficiente não diferir significativamente de zero, aceitaremos a hipótese de que não existe influência importante do meio sobre o caráter em estudo, nem que há casamentos preferenciais em relação a esse caráter.

Para exemplificar, consideremos os resultados da análise de regressão para investigar a herdabilidade da distância intercantal interna feita com base em medidas feitas na região orbitária em pessoas pertencentes a 62 famílias brasileiras, após ajustamento dos valores para idade e sexo (Tabela 4.3). Os valores de  $t$  expressos na Tabela 4.3 permitem concluir que:

- a) o coeficiente de regressão  $b_{PM}$  não difere significativamente de zero, o que permite descartar a hipótese de que houve casamentos preferenciais em relação à distância intercantal interna ou a de que o meio intrafamiliar comum exerceria influência importante sobre esse caráter quantitativo;
- b) todos os outros coeficientes de regressão são significativamente diferentes de zero e permitem estimar a herdabilidade da distância intercantal interna em 31 % porque:

$$h^2 = \frac{(2 \times 0,1574 \times 285,7143) + (2 \times 0,1551 \times 243,9024) + (0,3045 \times 151,5152)}{285,7143 + 243,9024 + 151,5152} = 0,3109$$

Os da Tabela 4.3 indicam, ainda, que não teríamos cometido enganos graves caso tivéssemos estimado a herdabilidade com base em apenas um dos três coeficientes de regressão utilizados. Realmente, tanto  $2b_{\bar{F}P} = 0,3148$  quanto  $2b_{\bar{F}M} = 0,3102$  ou  $b_{\bar{F}G} = 0,3054$  estão próximos do valor estimado ( $h^2 = 0,3109$ ).

Tabela 4.3 Dados para o cálculo da herdabilidade da distância intercantal interna em 62 famílias, após ajustamento dos dados para idade e sexo (Raposo-do-Amaral, 1973).

| $b$                     | $s_b$  | I        | $t$   | G.L. | P      |
|-------------------------|--------|----------|-------|------|--------|
| $b_{PM} = 0,0285$       | 0,1388 | -        | 0,205 | 60   | >0,80  |
| $b_{\bar{F}P} = 0,1574$ | 0,0594 | 285,7143 | 2,650 | 60   | <0,02  |
| $b_{\bar{F}M} = 0,1551$ | 0,0639 | 243,9024 | 2,427 | 60   | <0,02  |
| $b_{\bar{F}G} = 0,3045$ | 0,0813 | 151,5152 | 3,745 | 60   | <0,001 |

## REFERÊNCIAS

- Beiguelman, B. Genetics of *a-b* and *A'-d* ridge counts *Rev. Bras. Pesq. Med. Biol.* 4: 337-342, 1971.
- Beiguelman, B. *Curso prático de Bioestatística*. Edições da Revista Brasileira de Genética, Ribeirão Preto, 1994.
- Beiguelman, B. & Franchi-Pinto, C. Perinatal mortality among twins and singletons in a city in southeastern Brazil, 1984-1996. *Genet. Molec. Biol.* 23: 15-23, 2000.
- Clark, P.J. The heritability of certain anthropometric characters as ascertained from measurements of twins. *Am. J. Hum. Genet.* 8: 49-54, 1956.
- Falconer, D.S. *Introduction to Quantitative Genetics*. Oliver and Boyd, Edimburgh, 1960.
- Franchi-Pinto, C., Dal Colletto, G.M.D., Krieger, H. & Beiguelman, B. Genetic effect on Apgar score. *Genet. Molec. Biol.* 22: 13-16, 1999.
- Holzinger, K.J. The relative effect of nature and nurture influences on twins differences. *J. Educ. Psychol.* 20: 241-248, 1929.
- Kempthorne, O. & Tandon, O.B. The estimation of heritability by regression of offspring on parent. *Biometrics*, 9: 90-100, 1953.
- Kempthorne, O. & Osborne, R.H. The interpretation of twin data. *Am. J. Hum. Genet.* 13: 320-339, 1961.
- Mather, K. & Jinkjs, J.L. *Introdução à Genética Biométrica*. Traduzido por F.A. Moura Duarte, F.M. Sene, H.A. Rothschild, R.B. Lôbo, N. Mortari & A.P. Schlindwein. Sociedade Brasileira de Genética, Ribeirão Preto, SP, Brasil, 1984.
- Raposo-do-Amaral, C.M. *A herdabilidade das medidas da região orbitária. Contribuição ao estudo do hipertelorismo e telecanto*. Tese de Doutorado, Faculdade de Medicina, Universidade de São Paulo, 1973.
- Reeve, E.C.R. The variance of the genetic correlation coefficient. *Biometrics*, 11: 357-374, 1955.
- Smith, C. Concordance in twins: methods and interpretations. *Am. J. Hum. Genet.* 26: 454-466, 1974.
- Snedecor, W.G. *Statistical methods*. Iowa State Univ. Press, Ames, Iowa, 5a.ed, 9a. reimpressão, 1966.
- Weinberg, W. Beiträge zur Physiologie und Pathologie der Mehrlingsgeburten beim Menschen. *Pflügers Arch. Ges. Physiol.* 88: 346-430, 1901.