Parte 1.2: Regressão Linear Múltipla (Estimação e Inferência)

Econometria I - IBMEC

Marcus L. Nascimento

24 de setembro de 2025

1. Regressão Linear Múltipla

2. Estimação

3. Inferência

Regressão Linear Múltipla

Introdução

- Anteriormente, estudamos a relação entre uma **variável dependente** e uma **variável independente** através de um modelo de **regressão linear simples**.

- Em casos onde há **duas ou mais** variáveis independentes, estabelecemos uma relação com uma variável dependente através de um modelo de **regressão linear múltipla**.

- O modelo de regressão linear múltipla é descrito da seguinte forma:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon,$$

  onde:

  - β_0: coeficiente da regressão que representa o intercepto;
  - β_j: coeficiente da regressão associado com a variável indepente x_j, $j = 1, 2, \ldots, p$;
  - ε: erro aleatório.

Introdução

- De forma análoga ao caso da regressão linear simples, assume-se que $E[\varepsilon|x_1, x_2, \ldots, x_p] = 0$.
    - Variáveis independentes devem ser não correlacionadas com os fatores no termo de erro;
    - Forma funcional entre variável dependente e variáveis independentes corretamente especificada.
- Dada uma amostra de tamanho n, $(y_i, x_{i1}, x_{i2}, \ldots, x_{ip})$, $i = 1, 2, \ldots, n$, o modelo de regressão linear múltipla é dado por:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} + \varepsilon_i, \;\; i = 1, 2, \ldots, n,$$

ou

$$E[y_i|x_{i1}, x_{i2}, \ldots, x_{ip}] = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}.$$

Estimação

Introdução

- No caso do modelo de **regressão linear simples**, vimos que os estimadores de **Mínimos Quadrados Ordinários** de β_0 e β_1 eram encontrados através da minimização da função a seguir:

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2.$$

- No caso do modelo de **regressão linear múltipla**, os estimadores de **Mínimos Quadrados Ordinários** de $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ podem ser encontrados de forma análoga, ou seja, através da minimização da soma dos erros quadráticos:

$$S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p) = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \ldots - \beta_p x_{ip})^2.$$

Estimadores de Mínimos Quadrados Ordinários

O mínimo de $S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)$ é obtido através do cálculo de sua derivada com respeito a $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$, igualando o resultado a zero.

$$\begin{aligned}
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_0} &= 0; \\
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_1} &= 0; \\
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_2} &= 0; \\
&\vdots \\
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_p} &= 0.
\end{aligned}$$

Estimadores de Mínimos Quadrados Ordinários

O problema anterior de minimização recai na solução de um sistema de $p + 1$ equações lineares em $p + 1$ quantidades desconhecidas $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_p$.

$$
\begin{aligned}
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_0} &= -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \ldots - \hat{\beta}_p x_{ip}) = 0; \\
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_1} &= -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \ldots - \hat{\beta}_p x_{ip})x_{i1} = 0; \\
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_2} &= -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \ldots - \hat{\beta}_p x_{ip})x_{i2} = 0; \\
&\vdots \\
\frac{\partial S(\beta_0, \beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta_p} &= -2\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \ldots - \hat{\beta}_p x_{ip})x_{ip} = 0.
\end{aligned}
$$

Interpretação

Seja $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_p x_p$ a **modelo de regressão estimado**, temos:

- $x_1 = x_2 = \ldots = x_p = 0$: $\hat{y} = \hat{\beta}_0$.
  $\hat{\beta}_0$ é o valor médio estimado para a variável resposta, condicionado a $x_1 = x_2 = \ldots = x_p = 0$ e pode ser interpretável ou não.
- $x_1 \to x_1 + 1$: $\Delta\hat{y} = [\hat{\beta}_0 + \hat{\beta}_1(x_1+1) + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_p x_p] - [\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_p x_p] = \hat{\beta}_1$.
  $\hat{\beta}_1$ representa o quanto varia a média de y para um aumento de uma unidade da variável x_1 quando mantidas as demais variáveis constantes.
- $x_j \to x_j + 1$: $\Delta\hat{y} = [\hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_j(x_j + 1) + \ldots + \hat{\beta}_p x_p] - [\hat{\beta}_0 + \hat{\beta}_1 x_1 + \ldots + \hat{\beta}_j x_j + \ldots + \hat{\beta}_p x_p] = \hat{\beta}_j$.
  $\hat{\beta}_j$ representa o quanto varia a média de y para um aumento de uma unidade da variável x_j quando mantidas as demais variáveis constantes.

Exercício (Wooldridge, J. M. (2003))

Através de um modelo de regressão linear múltipla deseja-se encontrar a relação entre anos de educação (educ), anos de experiência (exper), anos na posição atual (tenure) e salário (em escala logarítmica). Utilizando dados de 526 trabalhadores, a equação a seguir é estimada:

$$\log(\text{salário}) = 0,284 + 0,092\ \mathit{educ} + 0,0041\ \text{exper} + 0,022\ \text{tenure}.$$

(i) Qual interpretação é dada para o coeficiente 0,092?

(ii) Qual o efeito estimado sobre o salário considerando a permanência na mesma posição por mais um ano?

Suposições

A1. (Linearidade) No modelo populacional, a variável resposta y está relacionada as variáveis independentes $x_1, \ldots, x_p$ e ao erro ε da seguinte forma:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon,$$

onde $\beta_0, \beta_1, \ldots, \beta_p$ são os parâmetros desconhecidos de interesse.

A2. (Amostragem aleatória) Pode-se utilizar uma amostra aleatória de tamanho n do modelo populacional, $\{(x_{i1}, x_{i2}, \ldots, x_{ip}, y_i) : i = 1, 2, \ldots, n\}$.

A3. (Média Condicional Zero) $E(\varepsilon | x_1, x_2, \ldots, x_p) = 0$.

Suposições

A4. (Ausência de Colinearidade Perfeita) Na amostra (e, portanto, na população) nenhum regressor é constante e não há relação linear **perfeita** entre os regressores.

A5. (Homocedasticidade) $\text{Var}(\varepsilon | x_1, x_2, \ldots, x_p) = \sigma^2$.

OBSERVAÇÃO:

- $E(\varepsilon | x_1, x_2, \ldots, x_p) = 0$ implica que todos os fatores contidos no erro devem ser não correlacionados com as variáveis explicativas, e deve ter sido usada a forma funcional correta.

Suposições

Possíveis causas para violação da suposição A3 ($E(\varepsilon|x_1, x_2, \ldots, x_p) = 0$):

- Omissão de variável explicativa importante, correlacionada com x_1, x_2, ... ou x_p;
- Forma funcional especificada incorretamente;
- Erro de medida em x_1, x_2, ... ou x_p;
- Simultaneidade entre y e x_1, x_2, ... ou x_p.

Propriedades

Sob as suposições A1-A4, os estimadores de Mínimos Quadrados Ordinários $\hat{\beta}_j$, $j = 1, 2, \ldots, p$, são não viesados:

$$E(\hat{\beta}_j) = \beta_j, \; j = 0, 1, 2, \ldots, p.$$

Como a exclusão ou inclusão de variáveis independentes afeta tal propriedade?

- Inclusão de variável irrelevante: não altera a propriedade de ausência de viés;
- Omissão de variável relevante (caso em que $p = 2$ e $p = 1$ é utilizado):

$$E(\tilde{\beta}_1) = \beta_1 + \beta_2 \frac{\sum_{i=1}^{2} (x_{i1} - \bar{x}_1) x_{i2}}{\sum_{i=1}^{2} (x_{i1} - \bar{x}_1)^2}$$

Propriedades

	Corr(x_1,x_2) > 0	**Corr(x_1,x_2) < 0**
$\beta_2 > 0$	positive bias	negative bias
$\beta_2 < 0$	negative bias	positive bias

Propriedades

Observações:

- Viés depende tanto dos sinais quanto das magnitudes;
- Em geral, se $p > 1$, a omissão de qualquer variável relevante faz com que todos os estimadores de mínimos quadrados sejam viesados;
- A menos que a variável omitida seja irrelevante ou não correlacionada com as demais variáveis explicativas presentes no modelo, os estimadores de mínimos quadrados serão viesados.

Propriedades

- As suposições A1-A5 conjuntamente são conhecidas como **suposições de Gauss-Markov**.
- Sob as suposições de Gauss-Markov, as variâncias dos estimadores de Mínimos Quadrados Ordinários são descritas da seguinte forma:

$$\mathrm{Var}(\hat{\beta}_j) = \frac{\sigma^2}{\mathrm{SST}_j(1 - R_j^2)}, \; j = 1, 2, \ldots, p,$$

onde

  - $\mathrm{SST}_j = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2$ é a variação amostral em x_j;
  - R_j^2 é o coeficiente de determinação calculado a partir da regressão de x_j nas demais variáveis independentes (incluindo intercepto).

Propriedades

- Componentes das variâncias dos estimadores: σ^2, SST_j e R_j^2.
- Variância do erro (σ^2): Quanto maior o σ^2, maior a variância dos estimadores.
    - Mais "Ruído" no modelo de regressão múltipla $\rightarrow$ Maior dificuldade em estimar os efeitos das variáveis independentes em y $\rightarrow$ Maior variância dos estimadores.
- Variação amostral total em x_j (SST_j): Quanto maior o SST_j, menor $\text{Var}(\hat{\beta}_j)$.
    - Para estimar β_j, é desejável que haja o máximo de variação amostral possível em x_j (aumento do tamanho da amostra n).

Propriedades

- Relação linear entre as variáveis independentes (R_j^2): Quanto maior o R_j^2, maior a variância dos estimadores.
    - Um R_j^2 próximo a 1 indica que as demais variáveis independentes (parte ou alguma das variáveis independentes) explicam muito da variação de x_j;
    - Forte relação linear entre x_j e as demais variáveis independentes (parte ou alguma das variáveis independentes) pode implicar em um aumento na variância dos estimadores;
    - Dado σ^2 e SST_j, a menor variância para $\hat{\beta}_j$ é obtida quando $R_J^2 = 0$, ou seja, quando não há correlação entre x_j e as demais variáveis independentes.
- Inclusão de variável irrelevante: Geralmente aumenta as variâncias dos demais estimadores de MQO.

Propriedades

(Teorema de Gauss-Markov) Melhores estimadores lineares não-viesados (BLUE):

- Sob as suposições A1-A5, os estimadores de MQO para os parâmetros $\beta_0, \beta_1, \beta_2, \ldots, \beta_p$ são os melhores dentre todos os estimadores da classe dos lineares não-viesados;

- Além de serem não-viesados, apresentam a menor variância dentre os demais estimadores não-viesados.

Estimador para a variância do erro (σ^2)

Seja $e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2} - \ldots - \hat{\beta}_p x_{ip}$ o **resíduo** da regressão linear. Para obtermos um estimador não enviesado de σ^2, analisamos a dispersão em torno da reta de regressão estimada:

$$\text{SSR} = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \text{ (Soma de quadrados dos resíduos) } .$$

Sob as suposições A1-A5, $E(\sum_{i=1}^{n} e_i^2) = (n-p-1)\sigma^2$, logo um estimador não viesado de σ^2 é

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-p-1} = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-p-1}.$$

Cálculo das variâncias dos estimadores de MQO

Com base em uma amostra, é possível encontrar estimativas para as variâncias de $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$ substituindo σ^2 por $\hat{\sigma}^2$ (estimador não viesado para σ^2) na expressão para $\mathrm{Var}(\hat{\beta}_j)$:

$$\widehat{\mathrm{Var}}(\hat{\beta}_j) = \frac{\hat{\sigma}^2}{\mathrm{SST}_j(1 - R_j^2)}, \ j = 1, 2, \dots, p.$$

Exercício (Wooldridge, J. M. (2003))

Utilizando dados de 4.137 estudantes universitários, a seguinte equação foi estimada via Mínimos Quadrados Ordinários:

$$GPA = 1,392 - 0,0135\ hsperc + 0,00148\ sat,$$

onde *hsperc* é o percentil do aluno no ensino médio, por exemplo, $hsperc = 5$ indica que o aluno terminou o ensino médio entre os 5% melhores classificados, e sat é a nota combinada dos exames oral e de matemática.

(i) Por que faz sentido que o coeficiente de *hsperc* seja negativo?

(ii) Qual o *GPA* predito quando $hsperc = 20$ e $sat = 1050$?

(iii) Considere dois estudantes (A e B) que tenham se formado no mesmo percentil no ensino médio, porém o *sat* do estudante A foi 140 pontos maior. Qual a diferença predita para o *GPA* dos dois estudantes?

(iv) Fixado *hsperc*, calcule a diferença no *sat* que implica uma diferença de 0,50 no *GPA*.

Inferência

Introdução

- Até este momento, obtemos os estimadores de mínimos quadrados ordinários considerando um modelo de regressão linear múltipla e estudamos suas propriedades.

- A seguir, voltaremos nossa atenção para o problema de testes de hipóteses acerca dos parâmetros de um modelo de regressão.

- Para tal, além das suposições de Gauss-Markov (A1-A5), adicionamos a seguinte suposição:

  A6. O erro estocástico ε é independente das variáveis explicativas e segue distribuição normal com média zero e variância σ^2, $\varepsilon \sim N(0, \sigma^2)$.

Observações:

- Para aplicações com dados do tipo *cross-sectional*, as suposições A1-A6 são conhecidas como suposições do modelo linear clássico (suposições CLM).

- Podemos resumir as suposições CLM na população da seguinte forma:

$$y|x_1, x_2, \ldots, x_p \sim N(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p, \sigma^2).$$

- Sob as suposições CLM, os estimadores de mínimos quadrados são **estimadores não-viesados de variância mínima**.

Distribuição amostral de $\hat{\beta}_j$

Sob as suposições CLM (MLR.1 a MLR.6), condicionado nos valores amostrais das variáveis explicativas, temos que

$$\hat{\beta}_j \sim N\left(\beta_j, \frac{\sigma^2}{\text{SST}_j(1 - R_j^2)}\right).$$

Da expressão anterior, temos que

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\frac{\sigma^2}{\text{SST}_j(1-R_j^2)}}} \sim N(0, 1).$$

Teste t

- Na distribuição amostral de $\hat{\beta}_j$, σ^2 é um parâmetro desconhecido e, portanto, deverá ser estimado.

- "Substituiremos" σ^2 por $\hat{\sigma}^2$.

- Dessa forma, precisamos estudar a distribuição de probabilidades da nova variável aleatória resultante, ou seja,

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\frac{\hat{\sigma}^2}{\mathrm{SST}_j(1-R_j^2)}}}.$$

Teste t

Sob as suposições CLM (A1-A6),

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\frac{\hat{\sigma}^2}{\mathrm{SST}_j(1-R_j^2)}}} \sim t_{n-(p+1)},$$

onde $p+1$ é o número de parâmetros desconhecidos no modelo populacional $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p$.

Com base neste resultado, podemos testar hipóteses envolvendo $\hat{\beta}_j$. Em particular, estamos interessados no seguinte teste:

$$\begin{aligned} \mathrm{H}_0 : \beta_j &= 0 \\ \mathrm{H}_1 : \beta_j &\neq 0 \ (\beta_j < 0 \text{ ou } \beta_j > 0). \end{aligned}$$

Teste t

Neste caso, a estatística de teste é denominada **estatística** t de $\hat{\beta}_j$ e é definida por

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j}{\sqrt{\frac{\hat{\sigma}^2}{\mathrm{SST}_j(1-R_j^2)}}}.$$

Logo, para um nível de significância α, rejeitamos H_0 se

- $|t_{\hat{\beta}_j}| > t_{n-p-1;1-\alpha/2}$ (teste bilateral);
- $t_{\hat{\beta}_j} > t_{n-p-1;1-\alpha}$ quando $\mathrm{H}_1 : \beta_j > 0$;
- $t_{\hat{\beta}_j} < t_{n-p-1;\alpha}$ quando $\mathrm{H}_1 : \beta_j < 0$.

Teste t

- Uma forma alternativa de decidirmos se rejeitamos ou não H_0 consiste em encontrar o valor-p.

- Dado o valor observado da estatística t, o valor-p é o menor nível de significância para o qual a hipótese nula é rejeitada.

- Para um teste onde $\mathrm{H}_0 : \beta_j = 0$, valor-$p$ é dado por

$$P(|T| > |t|),$$

  onde T denota uma variável aleatória com distribuição t-Student com $n - p - 1$ graus de liberdade e t denota o valor numérico da estatística de teste.

- Valores-p pequenos são evidência contra a hipótese nula; valores-p altos fornecem pouca evidência contra H_0.

Teste t

- Uma terceira forma de decidirmos se rejeitamos ou não $\mathrm{H}_0 : \beta_j = 0$ é através do intervalo de confiança.

- Considerando a distribuição da estatística de teste, é possível construir um intervalo de confiança tal que

$$\mathrm{IC}(\beta_j;\gamma) = \left(\hat{\beta}_j - t_{n-p-1;\frac{\alpha}{2}}\sqrt{\frac{\hat{\sigma}^2}{\mathrm{SST}_j(1-R_j^2)}}, \hat{\beta}_j + t_{n-p-1;1-\frac{\alpha}{2}}\sqrt{\frac{\hat{\sigma}^2}{\mathrm{SST}_j(1-R_j^2)}}\right),$$

onde $\gamma = 1 - \alpha$ e $t_{n-p-1;\frac{\alpha}{2}} = t_{n-p-1;1-\frac{\alpha}{2}}$.

- Sob um teste bilateral, rejeitamos H_0 se $0 \notin \mathrm{IC}(\beta_j;\gamma)$.

Teste F

- Através do Teste t é possível testar se uma particular variável independente x_j tem efeito sobre a variável resposta y.

- Em algumas situações, pode ser de interesse testar se um grupo de variáveis independentes tem efeito sobre a variável resposta, ou seja,

$$\begin{aligned} H_0 &: \beta_{j_1} = \beta_{j_2} = \ldots = \beta_{j_q} = 0 \\ H_1 &: H_0 \text{ não é verdadeira,} \end{aligned}$$

  onde $q < p$ é o tamanho do grupo a ser testado.

- Para tal, é necessário construir uma estatística de teste e uma possível maneira seria olhar no aumento relativo no SSR entre o modelo irrestrito $(x_1, x_2, \ldots, x_p)$ e o modelo restrito (ou seja, sem as variáveis $x_{j_1}, x_{j_2}, \ldots, x_{j_q}$).

Teste F

A estatística F é definida da seguinte forma:

$$F = \frac{(\mathrm{SSR}_r - \mathrm{SSR}_{ur})/q}{\mathrm{SSR}_{ur}/(n-p-q)},$$

onde

- SSR_r é a soma dos quadrados dos resíduos do modelo restrito;
- SSR_{ur} é a soma dos quadrados dos resíduos do modelo irrestrito.

Sob H_0 e assumindo que as suposições A1-A6 sejam válidas, $F \sim F_{q,n-p-1}$. Neste caso, refeitamos H_0 se $F > c$, onde c é obtido através da distribuição $F_{q,n-p-1}$.

Exercício (Wooldridge, J. M. (2003))

Considere que as taxas de aprovação de empréstimos em uma comunidade sejam determinadas por

$$apprate = \beta_0 + \beta_1 percmin + \beta_2 avginc + \beta_3 avgwlth + \beta_4 vgdebt + \varepsilon,$$

onde *percmin* é a porcentagem de minorias na comunidade, *avginc* é a renda média, *avgwlth* é a riqueza média, e *avgdebt* é alguma medida das obrigações médias de dívida.

(i) Como enunciar a hipótese nula de que não há diferença nas taxas de aprovação de empréstimos entre bairros devido à composição racial e étnica, uma vez que a renda média, a riqueza média e a dívida média tenham sido controladas?

(ii) Como enunciar a hipótese alternativa de que há discriminação contra minorias nas taxas de aprovação de empréstimos?

Exercício (Wooldridge, J. M. (2003))

Considere a relação entre o desempenho individual em um teste padronizado (*score*) e fatores relacionados à escola e fatores específicos do aluno. Os fatores escolares incluem tamanho médio da turma (*classize*), despesas por aluno (*expend*), remuneração média dos professores (*tchcomp*) e total de matrículas da escola (*enroll*). As variáveis específicas do aluno são renda familiar (*faminc*), escolaridade da mãe (*motheduc*), escolaridade do pai (*fatheduc*) e número de irmãos (*siblings*).
O modelo especificado é dado pela seguinte equação:

$$\begin{aligned} score &= \beta_0 + \beta_1 classize + \beta_2 expend + \beta_3 tchcomp + \beta_4 enroll + \beta_5 faminc \\ &+ \beta_6 motheduc + \beta_7 fatheduc + \beta_8 siblings + \varepsilon. \end{aligned}$$

(i) Enuncie a hipótese nula de que as variáveis específicas do estudante não têm efeito sobre o desempenho em testes padronizados, uma vez que os fatores relacionados à escola tenham sido controlados.
(ii) Qual o valor de p e q neste exemplo?
(iii) Escreva a versão restrita do modelo.