MÉTODO DOS MQ's GENERALIZADOS E PONDERADOS

Frederico Machado Almeida
frederico.almeida@unb.br

Departamento de Estatística
Instituto de Exatas
Universidade de Brasília

- Então o momento vimos que, a suposição principal do modelo de regressão linear,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{1}$$

é a de que o vetor dos erros $\boldsymbol{\epsilon}$ é normalmente distribuídos com média $\mathbf{0}$, e variância $\sigma^2\mathbf{I}$. Ou seja, que os erros são homoscedásticos.

- Entretanto, em muitas situações reais, a variância dos erros pode ser diferente para cada observação (problema de heterocedasticidade). Isto é, $\epsilon_i \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_i^2\right)$.
  - Até então vimos que, uma solução viável para estabilizar a variância dos erros seria transformar a variável de resposta y_i.
  - A transformação da variável resposta tem as seguintes desvantagens: (i) ela nem sempre é eficaz; (ii) pode causar sérios problema de correlação serial; (iii) pode resultar em uma relação não muito adequada para modelar o banco de dados.

- Quando a relação linear for adequada para modelar (1), porém, com variância não-constante, uma alternativa à transformação da variável resposta, seria usar alguns métodos de estimação considerados serem os mais robustos que os mínimos quadrados ordinários (MQO).
- Para contornar o problema da heterocedasticidade, iremos considerar como uma alternativa ao método dos MQO, os seguintes métodos de estimação:
  - Mínimos Quadrados Generalizados (MQG)
  - Mínimos Quadrados Ponderados (MQP): é considerado um caso particular dos MQG.

Método dos Mínimos Quadrados Generalizados

- Considere o modelo de regressão linear múltipla

$$y_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \epsilon_i, \text{ com } \mathrm{Var}(\epsilon_i) = \sigma_i^2,$$

como $i = 1, \cdots, n$. Aqui, x_{ij} são elementos de $\mathbf{X}$. E, β_j com $j = 0, \cdots, p$ são os parâmetros do modelo.

- Como $\mathrm{Var}(\epsilon_i) = \sigma_i^2$, então, $\mathrm{Var}(\epsilon) = \sigma_i^2 \mathbf{I}$ e portanto,

$$\mathrm{Var}(\boldsymbol{\epsilon}) = \begin{pmatrix} \sigma_1^2 & 0 & \cdots 0 \\ 0 & \sigma_2^2 & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ 0 & 0 & \cdots \sigma_n^2 \end{pmatrix} = \sigma^2 \mathbf{V}$$

Método dos Mínimos Quadrados Generalizados

- Sendo $\mathbf{V}$ uma matriz $n \times n$ positiva definida. Assim, usando a notação matricial obtemos:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \text{ com } \operatorname{Var}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{V}. \tag{2}$$

- Tal situação em que $\operatorname{Var}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{V}$ pode ser interpretada da seguinte forma:
  - Sendo $\mathbf{V}$ uma matriz diagonal, mas com elementos diferentes, então, os y_i's e os ϵ_i's `são não correlacionados`.
  - Se os elementos em $\mathbf{V}$ ($v_{ij} \neq 0$, para todo $i \neq j$), então, podemos afirmar que há correlação serial entre os erros.
  - Podemos então assumir que, para todo $i = 1, \cdots, n$ temos que $\sigma_i^2 = \sigma^2 v_{ii}$.
- Como a variância dos erros é viesada, a variância dos $\hat{\beta}$ também será viesada. Como consequência, os testes estatísticos são fortemente afetados

Método dos Mínimos Quadrados Generalizados

- Sendo $\mathbf{V}$ uma matriz não-singular e positiva definida, então, existe uma matriz $\mathbf{K}_{n\times n}$ não-singular e simétrica, tal que

$$\mathbf{K}^\top\mathbf{K} = \mathbf{K}\mathbf{K}^\top = \mathbf{K}\mathbf{K} = \mathbf{V},$$

onde $\mathbf{K} = \left(\sqrt{v_{ij}}\right)_{i,j=1,\cdots,n}$.

- É possível transformar o modelo em (2) da seguinte forma:

$$\underbrace{\mathbf{K}^{-1}\mathbf{y}}_{\mathbf{y}_g} = \underbrace{\mathbf{K}^{-1}\mathbf{X}}_{\mathbf{X}_g}\beta + \underbrace{\mathbf{K}^{-1}\boldsymbol{\epsilon}}_{\epsilon_g}, \text{ com } \operatorname{Var}(\boldsymbol{\epsilon}) = \sigma^2\mathbf{V}. \tag{3}$$

Aqui, $\mathbf{y}_g = \mathbf{K}^{-1}\mathbf{y}$, $\mathbf{X}_g = \mathbf{K}^{-1}\mathbf{X}$ e $\epsilon_g = \mathbf{K}^{-1}\boldsymbol{\epsilon}$ denotam as transformações do vetor de respostas, da matriz desenho e do vetor de erros aleatórios, respectivamente.

Método dos Mínimos Quadrados Generalizados

- Com base nas variáveis transformadas, o modelo de regressão pode ser dada por:

$$\mathbf{y}_g = \mathbf{X}_g\beta + \epsilon_g, \text{ com } \mathbb{E}\left(\boldsymbol{\epsilon}_g\right) = \mathbf{0} \text{ e } \mathrm{Var}\left(\boldsymbol{\epsilon}_g\right) = \sigma^2\mathbf{I}. \tag{4}$$

Prova:

$$\begin{aligned}
(i)\ \mathbb{E}\left(\boldsymbol{\epsilon}_g\right) &= \mathbb{E}\left(\mathbf{K}^{-1}\boldsymbol{\epsilon}\right) = \mathbf{K}^{-1}\mathbb{E}\left(\boldsymbol{\epsilon}\right) = \mathbf{0} \\
(ii)\ \mathrm{Var}\left(\boldsymbol{\epsilon}_g\right) &= \mathbb{E}\left[\left(\boldsymbol{\epsilon}_g - \mathbb{E}\left(\boldsymbol{\epsilon}_g\right)\right)\left(\boldsymbol{\epsilon}_g - \mathbb{E}\left(\boldsymbol{\epsilon}_g\right)\right)^\top\right] = \mathbb{E}\left(\boldsymbol{\epsilon}_g\boldsymbol{\epsilon}_g^\top\right) \\
&= \mathbb{E}\left[\mathbf{K}^{-1}\boldsymbol{\epsilon}\left(\mathbf{K}^{-1}\boldsymbol{\epsilon}\right)^\top\right] = \mathbb{E}\left[\mathbf{K}^{-1}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top\left(\mathbf{K}^\top\right)^{-1}\right] \\
&= \mathbf{K}^{-1}\mathbb{E}\left[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top\right]\left(\mathbf{K}^\top\right)^{-1} \\
&= \mathbf{K}^{-1}\sigma^2\mathbf{V}\left(\mathbf{K}^\top\right)^{-1} \\
&= \sigma^2\mathbf{I}.
\end{aligned}$$

Método dos Mínimos Quadrados Generalizados

- Assim, sob as suposições usuais segue que, $\boldsymbol{\epsilon}_g \sim \mathcal{N}\left(\mathbf{0}, \sigma^2\mathbf{I}\right)$.
- Os parâmetros do modelo em (4) podem ser estimados minimizando a soma do quadrado dos erros. Tal que,

$$\begin{aligned} SQE\left(\beta\right) &= \left(\mathbf{y}_g - \mathbf{X}_g\beta\right)^\top\left(\mathbf{y}_g - \mathbf{X}_g\beta\right) \\ &= \mathbf{y}_g^\top\mathbf{y}_g - 2\beta^\top\mathbf{X}_g^\top\mathbf{y}_g + \beta^\top\mathbf{X}_g^\top\mathbf{X}_g\beta. \end{aligned}$$

- O estimador dos MQG $\hat{\beta}_g$ de β é dado por:

$$\begin{aligned} \hat{\beta}_g &= \left(\mathbf{X}_g^\top\mathbf{X}_g\right)^{-1}\mathbf{X}_g^\top\mathbf{y}_g \\ &= \left(\mathbf{X}^\top\left(\mathbf{K}^{-1}\right)^\top\mathbf{K}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\top\left(\mathbf{K}^{-1}\right)^\top\mathbf{K}^{-1}\mathbf{y} \\ &= \left(\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{y}. \end{aligned}$$

Método dos Mínimos Quadrados Generalizados

- Pode-se mostrar facilmente que, $\mathbb{E}\left(\hat{\boldsymbol{\beta}}_g\right) = \boldsymbol{\beta}$ e $\text{Var}\left(\hat{\boldsymbol{\beta}}_g\right) = \sigma^2\left(\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X}\right)^{-1}$. Sendo $\hat{\boldsymbol{\beta}}_g$ um estimador não viesado com variância uniformemente menor (`ENVUM`).
- Sob as suposições usuais nos modelos de regressão linear, segue que, $\hat{\boldsymbol{\beta}}_g \sim \mathcal{N}\left(\boldsymbol{\beta}, \sigma^2\left(\mathbf{X}^\top\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\right)$.
- Inferências sobre $\hat{\boldsymbol{\beta}}_g$ (testes de hipóteses e intervalos de confiança) segue a mesma ideia apresentada nas aulas anteriores.

Método dos Mínimos Quadrados Generalizados

Exemplo Prático!!!

Método dos Mínimos Quadrados Ponderados

- O método MQP pode ser visto como um caso particular do método dos MQG.
- Geralmente é recomendado quando os erros são não correlacionados, $\text{Cov}(\epsilon_i, \epsilon_j) = 0$, $\forall_{i \neq j}$, mas com variâncias diferentes. Isto é, $\text{Var}(\epsilon_i) = \sigma_i^2$.
- Assim, vamos denotar por $\mathbf{W}$ uma matriz de pesos, tal que, $\mathbf{W} = \mathbf{V}^{-1}$. Ou seja,

$$\text{Var}(\boldsymbol{\epsilon}) = \sigma^2 \begin{pmatrix} \frac{1}{w_1} & 0 & \cdots 0 \\ 0 & \frac{1}{w_2} & \cdots 0 \\ \vdots & \vdots & \ddots \vdots \\ 0 & 0 & \cdots \frac{1}{w_n} \end{pmatrix} = \sigma^2 \mathbf{W}^{-1}.$$

Segue que, $\text{Var}(\epsilon_i) = \frac{\sigma^2}{w_i}$.

Método dos Mínimos Quadrados Ponderados

- Sendo w_i, com $i = 1, \cdots, n$ quantidades conhecidas. De igual forma que $\mathbf{V}$ é uma matriz diagonal, $\mathbf{W} = \text{diag}\left(w_{ii}\right)$.
- Como $\mathbf{K} = \left(\sqrt{v_{ij}}\right)_{i,j=1,\cdots,n}$, segue imediatamente que $\mathbf{K} = \mathbf{V}^{1/2}$. Logo, obtemos o seguinte resultado:

$$\mathbf{W}^{1/2} = \mathbf{V}^{-1/2} = \mathbf{K}^{-1}.$$

- Com base no resultado do item anterior, é possível modificar a equação apresentada em (3) da seguinte forma:

$$\underbrace{\mathbf{W}^{1/2}\mathbf{y}}_{\mathbf{y}_w} = \underbrace{\mathbf{W}^{1/2}\mathbf{X}}_{\mathbf{X}_w}\beta + \underbrace{\mathbf{W}^{1/2}\epsilon}_{\epsilon_w}, \text{ com } \text{Var}\left(\epsilon\right) = \sigma^2\mathbf{W}^{-1}. \tag{5}$$

Método dos Mínimos Quadrados Ponderados

- Sendo, $\mathbf{y}_w = \mathbf{W}^{1/2}\mathbf{y}$, $\mathbf{X}_w = \mathbf{W}^{1/2}\mathbf{X}$ e $\epsilon_w = \mathbf{W}^{1/2}\boldsymbol{\epsilon}$ as variáveis ponderadas.

- Segue então que,

$$\mathbf{y}_w = \mathbf{X}_w\beta + \epsilon_w, \text{ com } \mathbb{E}(\boldsymbol{\epsilon}_w) = \mathbf{0} \text{ e } \mathrm{Var}(\boldsymbol{\epsilon}_w) = \sigma^2\mathbf{I}. \tag{6}$$

- Aqui, $\mathbf{y}_w = \left(w_1^{1/2}y_1, \cdots, w_n^{1/2}y_n\right)^\top$; $\mathbf{X}_w = \left(w_{ij}^{1/2}x_{ij}\right)$ e, $\epsilon_w = \left(w_1^{1/2}\epsilon_1, \cdots, w_n^{1/2}\epsilon_n\right)^\top$. Como consequência,

$$\epsilon_{wi} = w^{1/2}\epsilon_i = w^{1/2}\left(y_i - \beta_0 - \sum_{i=1}^{n}\beta_j x_{ij}\right).$$

Método dos Mínimos Quadrados Ponderados

- A soma do quadrado dos erros, no contexto ponderado, é dada por:

$$
\begin{aligned}
SQE(\beta) &= \sum_{i=1}^{n} w_i \left(y_i - \beta_0 - \sum_{i=1}^{n} \beta_j x_{ij} \right)^2 \\
&= (\mathbf{y}_w - \mathbf{X}_w \beta)^\top (\mathbf{y}_w - \mathbf{X}_w \beta) \\
&= \mathbf{y}_w^\top \mathbf{y}_w - 2\beta^\top \mathbf{X}_w^\top \mathbf{y}_w + \beta^\top \mathbf{X}_w^\top \mathbf{X}_w \beta.
\end{aligned}
$$

- Derivando a $SQE(\beta)$ em em função de β, e igualando a zero obtemos o estimador dos MQP $\hat{\beta}_w$ de β. Cuja expressão é dada por:

$$
\begin{aligned}
\hat{\beta}_w &= \left(\mathbf{X}_w^\top \mathbf{X}_w \right)^{-1} \mathbf{X}_w^\top \mathbf{y}_w \\
&= \left(\mathbf{X}^\top \left(\mathbf{W}^{1/2} \right)^\top \mathbf{W}^{1/2} \mathbf{X} \right)^{-1} \mathbf{X}^\top \left(\mathbf{W}^{1/2} \right)^\top \mathbf{W}^{1/2} \mathbf{y} \\
&= \left(\mathbf{X}^\top {\color{blue}\mathbf{W}} \mathbf{X} \right)^{-1} \mathbf{X}^\top {\color{blue}\mathbf{W}} \mathbf{y}.
\end{aligned}
$$

Método dos Mínimos Quadrados Ponderados

- Pode-se mostrar facilmente que, $\hat{\boldsymbol{\beta}}_w$ é um estimador não-tendencioso para $\boldsymbol{\beta}$, com variância $\text{Var}\left(\hat{\boldsymbol{\beta}}_w\right) = \sigma^2 \left(\mathbf{X}^\top \mathbf{W} \mathbf{X}\right)^{-1}$.
- Como $\mathbf{W}$ é uma matriz simétrica, segue imediatamente que, $\mathbf{W}^{1/2}$ é igualmente simétrica, e portanto, $\mathbf{W}^{1/2}\left(\mathbf{W}^{1/2}\right)^\top = \mathbf{W}$. Como consequência, $\mathbf{W}^{-1/2}\left(\mathbf{W}^{-1/2}\right)^\top = \mathbf{W}^{-1}$.

Método dos Mínimos Quadrados Ponderados

Como escolher os pesos, w_i?

- A literatura dispõem de vários métodos de como os pesos, w_i devem ser obtidos. De forma geral, a escolha dos pesos depende dos seguintes fatores:
  - (i) Se $\text{Var}(Y_i) = X_i\sigma^2$ (com $X_i > 0$) então, $w_i \propto 1/X_i$, e
  - (ii) Se $\text{Var}(Y_i) = X_i^2\sigma^2$ (para todo X_i) então, $w_i \propto 1/X_i^2$
- Outras formas de obter os pesos estão resumidas abaixo:

Método dos Mínimos Quadrados Ponderados

Resposta	$\bar{y}_i$
Tamanho	n_i
Variância amostral	S_i^2

Escolha de w_i:

1. $w_i \propto n_i$ se $\text{Var}\left(\bar{Y}_i\right) = \sigma^2/n_i$
2. $w_i \propto \frac{n_i}{S_i^2}$ se $\text{Var}\left(\bar{Y}_i\right) = \sigma^2/n_i$
3. $w_i \propto \frac{1}{S_i^2}$ se $\text{Var}\left(\bar{Y}_i\right) = \sigma^2/m$.

Sendo m o número de grupos em que os dados são divididos.

ANÁLISE DE COVARIÂNCIA (*ANCOVA*)

Análise de Covariância

- Os modelos da análise de covariância (ANCOVA) são, na verdade, uma mistura de modelos de regressão com modelos de análise de variância.
- A situação em que esses modelos são utilizados é, em essência, aquela em que as observações da variância dependente, Y são obtidas sob k condições diferentes, de forma que, para cada uma dessas condições, temos um grupo de observações.
- Tais grupos de observações de Y são representadas no modelo de regressão, por variáveis fictícias (variáveis *dummies*).
- Por exemplo, o modelo de regressão pode ser dado por:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \beta_{p+1} z_1 + \cdots + \beta_{p+k-1} z_{k-1} + \epsilon, \qquad (7)$$

com $z_j = 1$, se a observação pertence ao grupo j, e 0 caso contrário.

Análise de Covariância

- Como o objetivo principal do modelo de covariância é comparar a média μ_y nos diferentes grupos, o modelo em (7) pode ser resumido em,

$$
\begin{aligned}
y &= \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p + \epsilon, \text{ se } \ z_j = 0 \ \ \forall j \\
y &= (\beta_0 + \beta_{p+1} + \cdots + \beta_{p+k-1}) + \beta_1 x_1 + \cdots + \beta_p x_p + \epsilon, \text{ se } \ z_j = 1 \, \forall j
\end{aligned}
$$

- Porém, vários outros submodelos intermediários podem ser listados.
- Além de listar os possíveis submodelos, o pesquisador pode estar interessado em incluir um efeito multiplicativo no modelo de regressão linear (*efeito de interação*).

Análise de Covariância (Interação de Variáveis)

- Existem diferentes formas de incluir variáveis qualitativas no modelo. No entanto, a forma que tem frequentemente usada consiste em criar variáveis dummies.
- Em determinadas situações, variáveis quantitativas podem ser codificadas em níveis, podendo originar uma ou mais variáveis qualitativas.
- Por *interação* entende-se como sendo a situação onde o efeito combinado de duas ou mais variáveis independentes na média da variável resposta pode variar em função de acordo com os níveis da variáveis independentes.

Análise de Covariância

Exemplo 1: Suponha que um médico está conduzindo um experimento no qual um grupo de pacientes foram subdivididos aleatoriamente em dois subgrupos: o grupo que recebeu o medicamento, e o que não recebeu. O objetivo do estudo é estudar o nível de redução do colesterol no organismo, y. Suponha que tenhamos também as variáveis, peso (x_1), idade (x_2). A variável tratamento foi categorizado tal que,

$$x_3 = \begin{cases} tto\ A, & \text{se recebeu o tratamento A} \\ tto\ B, & \text{se recebeu o tratamento B.} \end{cases}$$

Assim, o modelo de regressão pode ser formulado da seguinte forma:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i. \tag{8}$$

Análise de Covariância

- Vamos definir a variável x_{3i} tal que,

$$x_3 = \begin{cases} 1, & \text{tto A} \\ 0, & \text{tto B.} \end{cases}$$

- Para cada nível da variável x_{3i}, os seguintes submodelos podem ser definidos:

$$\begin{aligned} y_{Ai} &= (\beta_0 + \beta_3) + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \text{ tto A} \\ y_{Bi} &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \text{ tto B.} \end{aligned}$$

- Ou seja, a ideia da ANCOVA consiste na possibilidade de incluir não só variáveis quantitativas no modelo, mas também, de variáveis qualitativas. Como por exemplo, sexo, escolaridade, classe social, etc.

Análise de Covariância (Interação de Variáveis)

Definition (Interação de Variáveis)

A interação de variáveis se refere a condição em que o efeito de certas variáveis preditoras no modelo não é aditivo, mas sim, multiplicativo.

- Incluindo termo(s) de interação no modelo apresentado em (8) obtemos,

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_{13} x_{1i} x_{3i} + \beta_{23} x_{2i} x_{3i} + \epsilon_i. \quad (9)$$

- Neste caso obtemos os seguintes submodelos,

$$\begin{aligned} y_{Ai} &= (\beta_0 + \beta_3) + (\beta_1 + \beta_{13})\, x_{1i} + (\beta_2 + \beta_{23})\, x_{2i} + \epsilon_i, \text{ tto A} \\ y_{Bi} &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \text{ tto B.} \end{aligned}$$

Análise de Covariância (Interação de Variáveis)

Observe que a interpretação dos coeficientes de regressão é tal que,

- $(\beta_1 + \beta_{13})$: representa a variação (aumento ou redução) na média de y em pacientes que receberam o tto A, para cada variação unitária do peso (x_1), mantendo constante a idade (x_2).
- $(\beta_2 + \beta_{23})$: representa a variação na média de y em pacientes que receberam o tto A, para cada variação unitária da idade (x_2), mantendo constante a variável peso (x_1).
- Note que nesse caso o intercepto $\beta_0 + \beta_3$ não tem uma interpretação prática, já que, tanto x_{1i} como x_{2i} não podem assumir o valor 0.
- Para pacientes que receberam o tratamento do tipo B, a interpretação segue a mesma forma de um modelo usual.

Análise de Covariância (Interação de Variáveis)

Exemplo 2: Considere o seguinte exemplo hipotético no qual o diretor de uma empresa pretende avaliar o impacto de certas variáveis na modelagem do escore de satisfação dos mesmos (assumindo valores de 0 a 100), e podendo assumir valores decimais. As variáveis de interesse são, x_1 : sexo (0-M e 1-F); x_2 : idade (em anos); x_3 : nível de escolaridade (1-Superior, 0-caso contrário); x_4 : salário (em reais), e x_5 : experiência profissional (1-Com experiência, e 0-sem experiência).

- O modelo de interesse é dado por:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \epsilon_i. \qquad (10)$$

- Note que por alguma eventualidade, o diretor da empresa pode estar interessado em avaliar, não só o efeito aditivo das covariáveis de regressão, mas também, o efeito aditivo (baseado no conhecimento à priori).

Análise de Covariância (Interação de Variáveis)

- x_1x_4 : interação entre as variáveis sexo e salário.
- x_2x_5 : interação entre as variáveis idade e experiência profissional.
- x_4x_3 : interação entre as variáveis salário e escolaridade.
- $x_1x_3x_4$: interação entre as variáveis sexo, escolaridade e salário, etc.

Ou seja, podemos escrever o novo modelo da seguinte forma:

$$
\begin{aligned}
y_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_{14} x_{1i} x_{4i} \\
&\quad + \beta_{25} x_{2i} x_{5i} + \beta_{34} x_{3i} x_{4i} + \beta_{134} x_{1i} x_{3i} x_{4i} + \epsilon_i.
\end{aligned} \tag{11}
$$

Análise de Covariância (Interação de Variáveis)

Com base no modelo apresentado em 12, vários submodelos podem ser obtidos, alguns dos quais incluem:

$$
\begin{aligned}
y_{Fi} &= (\beta_0 + \beta_1) + \beta_2 x_{2i} + \beta_3 x_{3i} + (\beta_4 + \beta_{14})\, x_{4i} + \beta_5 x_{5i} + \beta_{25} x_{2i} x_{5i} \\
&\quad + (\beta_{34} + \beta_{134})\, x_{3i} x_{4i} + \epsilon_i. \\
y_{Mi} &= \beta_0 + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_{25} x_{2i} x_{5i} + \beta_{34} x_{3i} x_{4i} + \epsilon_i. \\
&\dots \\
y_{FxSup} &= (\beta_0 + \beta_1 + \beta_3) + \beta_2 x_{2i} + (\beta_4 + \beta_{14} + \beta 34 + \beta_{134})\, x_{4i} + \beta_5 x_{5i} \\
&\quad + \beta_{25} x_{2i} x_{5i} + \epsilon_i.
\end{aligned}
$$

$(\beta_4 + \beta_{14} + \beta 34 + \beta_{134})$: representa a mudança na média da variável resposta, nos funcionários do sexo feminino, com escolaridade superior, para cada variação unitária do salário dos funcionários, mantendo constante os demais fatores do modelo.