

Lista 2 - Econometria II (2022)

Professora: Fabiana Rocha
Monitor: Thiago Pastorelli Rodrigues

Exercício 1. Assinale V ou F nas seguintes assertivas e justifique:

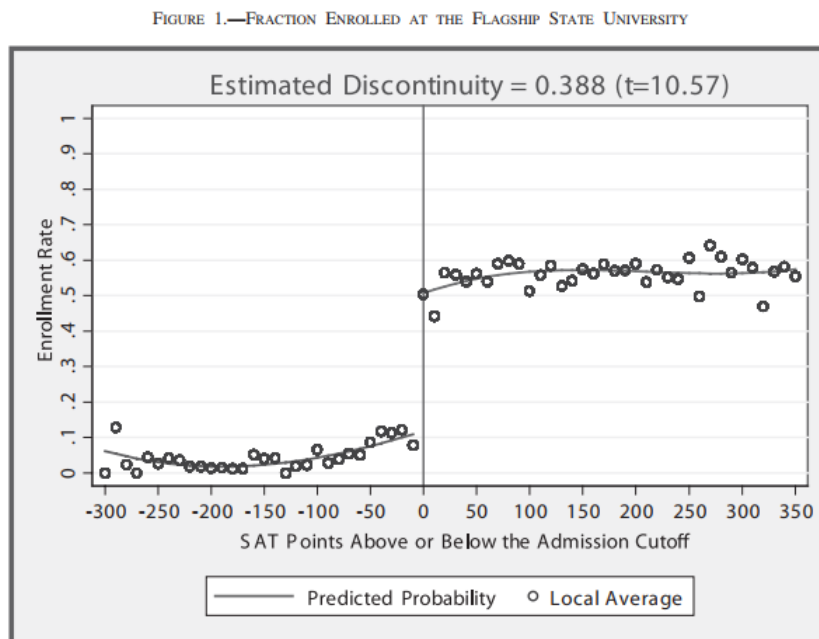
- Uma hipótese necessária à identificação dos parâmetros causais num modelo de regressão descontínua é a continuidade da variável dependente em relação ao cut-off na running variável.
- O RDD fuzzy consiste no fato que a participação no tratamento salta de 0 para 1 no cut-off da running variável.
- O efeito causal estimado com o RDD é o ATE.
- A manipulação de participação no tratamento por parte dos participantes pode enviesar a estimação do efeito causal no RDD.
- A existência de massa de probabilidade não-nula para certos pontos na running variável compromete a identificação, já que há fatores externos impactando essa variável, além do fato de o número de observações na proximidade do cut-off poder ser reduzido.

Exercício 2. Assumindo efeitos de tratamento constantes e um sharp RD design, o efeito causal de interesse no RDD é dado por:

$$\begin{aligned}\delta &= \lim_{X_i \rightarrow X_0^+} E[Y_i^1 | X_i = X_0] - \lim_{X_i \rightarrow X_0^-} E[Y_i^0 | X_i = X_0] \\ &= \lim_{X_i \rightarrow X_0^+} E[Y_i | X_i = X_0] - \lim_{X_i \rightarrow X_0^-} E[Y_i | X_i = X_0]\end{aligned}$$

Explique o que isto significa.

Exercício 3. Neste item, vamos analisar algumas escolhas realizadas num trabalho que utiliza RDD. Considere a seguinte figura:



Neste trabalho, Hoekstra (2009) busca avaliar o impacto de o aluno ter uma nota igual ou superior ao exigido por universidades prestigiosas dos estados americanos sobre seus salários futuros. Responda com base neste artigo aos seguintes itens:

- i. Pela última figura, o design utilizado para a regressão descontínua neste estudo consistiu de sharp ou fuzzy design? Justifique.
- ii. Dentre as opções na modelagem do RDD, é usual fazer-se regressões locais abaixo e acima do cut-off da running variável. O que se pretende exatamente captar com esse tipo de regressão?
- iii. Explique o porquê de Hoekstra (2009) centralizar as notas em torno de 0. Como você faria isso na prática?
- iv. A hipótese de continuidade é crucial na identificação do LATE em qualquer RDD. Que argumentos favoráveis e desfavoráveis podem ser dados no caso do estudo de Hoekstra (2009) para a validade desta hipótese? Qual lado dessa argumentação você considera mais razoável?

Exercício 4. Descreva quais as consequências de se ignorar os seguintes fatos sobre o efeito causal do RDD.

- i. A regra de tratamento é conhecida antecipadamente aos participantes.
- ii. Os agentes estão interessados em ajustes na sua situação de tratamento.
- iii. Agentes têm tempo de ajuste.
- iv. O cut-off é endógeno em relação a fatores que causam o resultado potencial a se deslocar.
- v. Há massa de probabilidade não-nula concentrada em alguns pontos da running variável.

Exercício 5. (Adaptado de Avaliação Econômica de projetos sociais - Itaú Social) O governo gostaria de saber se uma política que expandisse o acesso a planos de saúde teria um efeito positivo na saúde dos indivíduos, pois os levaria a ir mais vezes ao médico. Para responder a tal pergunta, usamos como medida de acesso ao serviço de saúde o número de vezes que o indivíduo i foi ao médico nos últimos 12 meses. O modelo que gostaríamos de estimar é: $Y_i = \beta_0 + \beta_1 C_i + \gamma X_i + \varepsilon_i$, em que Y_i é o número de vezes que o indivíduo i foi ao médico, C_i é uma variável binária que é igual a 1 se o indivíduo i possui plano de saúde, X_i é um vetor de características socioeconômicas do indivíduo i e ε_i é o termo não-observável da equação. Nos Estados Unidos, ao completar 65 anos, todos os indivíduos têm acesso a um plano de saúde completo que é pago pelo governo. Card, Dobkin e Maesta (2008) usam a descontinuidade quando o indivíduo completa 65 anos para estimar o efeito do plano de saúde no número de consultas médicas. A ideia é que os indivíduos com 64 anos são semelhantes aos indivíduos com 65 anos, exceto que o segundo grupo tem plano de saúde completo provido pelo governo.

- i. Alguns indivíduos se aposentam aos 65 anos e têm queda da renda anual. Como isso poderia causar um viés no estimador de regressão descontínua?

Considere o seguinte modelo que relaciona idade com acesso a plano de saúde: $C_i = \alpha_0 + \delta X_i + \pi m_i + \omega_i$, em que m_i é uma variável binária que assume valor igual a 1 se o indivíduo tem 65 anos ou mais e 0, caso contrário.

- ii. Qual a interpretação de π ?

Exercício 6. Sobre os modelos de primeiras diferenças e efeitos fixos, responda:

- Quais as hipóteses necessárias para consistência nos dois modelos? Compare-as.
- Em qual modelo perdemos mais graus de liberdade na estimação? Explique.
- Explique em quais casos um estimador é mais eficiente que o outro.
- Para $T = 2$ (2 períodos), mostre que o estimador de primeiras diferenças e de efeitos fixos são idênticos.

Exercício 7. Considere um modelo com uma única variável explicativa. Uma das equações utilizadas para obter o estimador de efeitos fixos é a seguinte:

$$\bar{y}_i = \beta_0 + \beta_1 \bar{x}_i + a_i + \bar{u}_i$$

onde a barra sobre a variável representa sua média no tempo. Podemos assumir que $E(a_i) = 0$, pois incluímos o intercepto no modelo. Suponha que $\bar{u}_i$ é não-correlacionado com $\bar{x}_i$, mas $Cov(x_{it}, a_i) = \sigma_{xa}$ para todo t (e i , pois a amostra é aleatória na cross-section).

- Seja $\tilde{\beta}_1$ o *between estimator*. Mostre que:

$$plim \tilde{\beta}_1 = \beta_1 + \frac{\sigma_{xa}}{Var(\bar{x}_i)}$$

- Assuma que, para todo $t = 1, 2, \dots, T$, x_{it} seja não correlacionado no tempo e tenha variância constante igual a σ_x^2 . Mostre que:

$$plim \tilde{\beta}_1 = \beta_1 + T \left(\frac{\sigma_{xa}}{\sigma_x^2} \right)$$

- Se as variáveis explicativas não são muito correlacionadas no tempo, o que o teim (b) sugere sobre a inconsistência no *between estimator* quando existem mais períodos de tempo?

Exercício 8. Suponha que você tem $T = 2$ anos de dados para o mesmo grupo de N indivíduos. Considere o seguinte modelo de determinação dos salários:

$$\log(sal) = \theta_1 + \theta_2 d2_t + z_{it}\gamma + \delta_1 feminino_i + \delta_2 d2_t \cdot feminino_i + c_i + u_{it}$$

Em que os efeitos não observados podem ser correlacionados com z_{it} e $feminino$. A variável $d2_t$ é indicadora do período de tempo, em que $d2_t = 1$ se $t = 2$ e $d2_t = 0$ se $t = 1$. Assuma que:

$$E(u_{it} | feminino_i, z_{i1}, z_{i2}, c_i) = 0, \quad t = 1, 2$$

- Sem mais suposições, quais parâmetros podem ser consistentemente estimados?
- Interprete θ_2 e δ_2 .
- Mostre que a equação diferenciada no tempo pode ser escrita como:

$$\Delta \log(sal) = \theta_2 + \Delta z_i \gamma + \delta_2 feminino_i + \Delta u_i$$

Exercício 9. Suponha que você queira estimar o seguinte modelo:

$$y_{it} = \beta_0 + \beta_1 y_{it-1} + \beta_2 x_{it} + u_{it}$$

$$u_{it} = v_{it} + c_i$$

Explique por que não podemos estimá-lo por nenhum dos métodos de painel aprendidos em sala. (Dica: olhe para a hipótese de exogeneidade de cada modelo).

Exercício 10. Queremos analisar como o salário dos trabalhadores se comporta quando as firmas fazem investimento interno, usando um painel de indivíduos. Para isso, vamos considerar o salário do trabalhador i , na firma j e no ano t , além disso seja o investimento que a firma faz no ano t e demais covariadas denotadas por x_{ijt} :

$$\log(\text{sal})_{ijt} = \beta_0 + \beta_1 \text{invest}_{jt} + \beta_2 x_{ijt} + \text{efeitosfixos} + \mu_t + u_{ijt}$$

Suponha, ainda, que os trabalhadores estão alocados em 50 firmas.

- O que são os efeitos fixos? O que eles captam nesse caso?
- Se considerarmos que alguns trabalhadores trocaram de emprego ao longo do período analisado, quais efeitos fixos você pode incluir? Justifique.
- Se os trabalhadores não puderem trocar de emprego, quais seriam os efeitos fixos?

Exercício 11. Suponha que estamos interessados em estimar o efeito da oferta de microcrédito o sobre o consumo de famílias em pobreza extrema. Para isto são realizadas entrevistas anuais com uma amostra aleatória de famílias que têm renda per capita menor do que um valor predeterminado $\bar{y}$ qualquer. O modelo estimado é:

$$l\text{consumo}_{imt} = \beta_0 + \beta_1 l\text{credito}_{imt} + \beta_2 \text{tamfam}_{imt} + v_{imt}$$

onde $l\text{consumo}_{imt}$ é o consumo da família i , no município m , no ano t (em log); $l\text{credito}_{imt}$ é o volume de microcrédito tomado pela família (em log); tamfam_{imt} é o tamanho da família e v_{imt} é o termo de erro.

- Se você tivesse que incluir efeitos fixos neste modelo, quais seriam escolhidos? Como você interpreta estes efeitos?
- Suponha que o termo de erro u_{imt} é não-correlacionado com as variáveis do seu modelo, inclusive com os efeitos fixos que você adicionou. Existe alguma situação em que você obterá estimadores viesados? Explique.

Exercício 12. Assinale se as afirmativas são verdadeiras (V) ou falsas (F) e justifique sua resposta.

- Um banco de dados em painel é tal que acompanhamos a mesma observação da cross-section ao longo do tempo.
- Quando temos um banco de dados de painel podemos estimar nosso modelo por MQO agrupado (sem nenhuma transformação das variáveis) caso o erro idiossincrático tenha média zero.
- O estimador de primeiras diferenças exige que as variáveis de controle utilizadas na análise variem ao longo do tempo.
- Para dois períodos de tempo, os estimadores de efeitos fixos, primeiras diferenças e efeitos aleatórios são idênticos.
- O estimador de efeitos aleatórios permite a inclusão de variáveis de controle constantes ao longo do tempo.
- O teste de Hausman serve para testar se o estimador de efeitos fixos é consistente.

Exercício 13. Considere o seguinte modelo:

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_3 x_{3it} + u_{it}$$

onde $u_{it} = e_i + v_{it}$. Assuma que $v_{it} \perp v_{js}, u_i, X_{it}$ para todo i, t, j, s com $t \neq s$ e $i \neq j$, $\text{var}(u|X) = \sigma_u^2$, $\text{var}(v|X, u) = \sigma_v^2$ e $E[v|u, X] = 0$.

Suponha que $E[u|X] = 0$. Derive a distribuição do estimador de Mínimos Quadrados Agrupados de $\beta = [\beta_0 \ \beta_1 \ \beta_2 \ \beta_3]'$.

Exercício 14. (Equivalência entre efeitos fixos *within estimator* e mínimos quadrados de variáveis dummies) Considere a seguinte equação:

$$y_{it} = \alpha_i + \beta X_{i,t} + e_{i,t}$$

- Mostre que o estimador de efeitos fixos para β da equação acima equivale à estimação via MQO numa regressão de y_{it} contra x_t , com a inserção de $N-1$ dummies d_i , com $i = 2, \dots, N$, com $d_j = 1\{j = i\}$.
- Adicionalmente, argumente o porquê de duas abordagens (N e $N-1$, sem e com intercepto, respectivamente) para esta regressão de dummies serem idênticas e por que não podemos ter uma regressão com N dummies para cada indivíduo e um intercepto?

Exercício 15. Um modelo do tipo:

$$y_{it} = \alpha_i + \beta X_{i,t} + e_{i,t}$$

é conhecido como do tipo *one-way fixed effects* já que ele explicitamente modela características individuais constantes ao longo do tempo, tanto observáveis como não-observáveis. Responda aos seguintes itens:

- O parâmetro α_i também é conhecido como efeitos fixos da observação i . Nele temos fatores observáveis e não-observáveis constantes ao longo do tempo. Suponha agora que tenhamos uma adaptação do último modelo no seguinte formato.

$$y_{it} = \alpha_i + \beta X_{i,t} + \gamma \text{race}_i + e_{i,t}$$

Por simplicidade (e razoabilidade), suponha que a variável *race* (raça) seja imutável no tempo para um mesmo indivíduo. A estimação *within* do parâmetro γ é factível? Em caso positivo, como você implementaria essa estimação e qual seria a interpretação do coeficiente γ da variável *race*? Em caso negativo, por que não poderíamos implementar essa regressão usando o estimador *within*?

- Considere ainda a equação do item anterior. Suponha adicionalmente que você pode acreditar que os efeitos fixos não são correlacionados com $[X_{it}, \text{race}_i]$ nem com o erro idiossincrático e_{it} . Por que a modelagem via efeitos aleatórios é neste caso mais eficiente que a estimação via efeitos fixos?
- Caso sua resposta ao item (a) tenha sido negativa no caso da factibilidade da estimação do parâmetro γ , explique quais os problemas que podemos incorrer ao utilizar efeitos aleatórios no caso de que suas hipóteses adjacentes a esse método não sejam respeitadas pelos efeitos fixos α_i .
- Vamos agora ampliar a equação dada no item (a). Considere assim a seguinte equação:

$$y_{it} = \alpha_i + f_t + \beta X_{i,t} + \gamma \text{race}_i + e_{i,t}$$

O modelo anterior com α_i e f_t é conhecido como *two-way fixed effects* ou TWFE. Uma estimação *within* adaptada também pode ser utilizada aqui para eliminar os termos α_i e f_t . Interprete o que f_t representa e dê algum exemplo de alguma variável que poderia ser modelada neste termo.

e Suponha agora uma nova equação para nosso modelo em questão, qual seja:

$$y_{it} = \alpha_i + f_t + \beta X_{i,t} + \gamma \text{race}_i + \kappa \text{crise_2008}_t + e_{i,t}$$

A variável `crise_2008` é uma variável indicadora que assume valor 1 caso o ano analisado seja 2008, 2009 ou 2010. Suponha que $t = 2000 - 2015$. A estimação do coeficiente κ é factível usando o estimador *within* de efeitos fixos? Explique o porquê de ela ser ou não factível?

f Até agora, consideramos efeitos fixos observacionais e temporais num contexto de dados em painel. Essa ideia de isolar efeitos fixos de quaisquer tipos é bastante generalista e pode ser utilizada em diferentes configurações. Vamos pensar agora em alguns problemas em não se modelar os diferentes tipos de efeitos fixos relevantes presentes num modelo de dados em painel. Assim sendo, considere que efeitos fixos temporais são relevantes no modelo populacional de interesse. Contudo, suponha que tenhamos utilizado a equação do item (a), invés da equação do item (d). Suponha que o vetor de variáveis observáveis $X_{i,t}$ contenha variáveis correlacionadas com os efeitos fixos temporais f_t . Que tipo de problema temos ao ignorarmos em nossa modelagem os efeitos fixos temporais? Em termos de consistência e eficiência, quão grave é esse problema?

Exercício 16. Foi noticiado em dois jornais locais que a política atual de redução da criminalidade de seu estado era efetiva. Em um dos jornais, a comprovação dessa efetividade foi associada à informação de que a taxa de roubos de veículos havia caído 10% nos últimos dois anos. Já no outro jornal, a efetividade da política foi consubstanciada por meio da comparação da taxa de roubo de veículos no último ano do seu estado com a média da mesma taxa para os demais estados do país. Como um conhecedor da área de avaliação de impacto, ao ler os jornais você considera ambas as matérias ingênuas e decide tentar estimar o verdadeiro impacto do programa. Responda aos itens a seguir:

a) Por que você acha a matéria do primeiro jornal ingênuas? E a segunda? Apresente as equações de regressão que você escreveria para representar cada uma das matérias dos jornais.

b) Você decide estimar o efeito da política usando o método de diferenças em diferenças (*DD*). Suponha que existem dados disponíveis para vários estados (incluindo o seu) ao longo do tempo para os períodos anterior e posterior à introdução da política. Apresente a equação de regressão que você usaria para estimar o efeito da política pelo método *DD*. Em que ela difere das equações de regressão do item anterior?

c) O que você deveria testar com dados anteriores ao início da política para averiguar se o grupo de controle parece adequado aos seus propósitos?

Exercício 17. Card e Krueger (1994) estudam o efeito do salário mínimo sobre emprego nos EUA. Para estimar este efeito, os autores usam um **experimento natural**: em Abril de 1992 o salário mínimo aumentou no estado de Nova Jersey, mas se manteve constante na Pensilvânia (estado vizinho). Considere ainda que os autores coletam dados de restaurantes *fast-foods* (principais empregadores ao salário mínimo) para os meses de Fevereiro e Novembro de 1992. Com base nessas informações, responda:

a) Você acredita que haveria problema de endogeneidade numa regressão simples caso? Justifique sua resposta.

b) Descreva como você estimaria este efeito utilizando o método de diferenças-em-diferenças.

c) Suponha agora que os autores verifiquem que no final de 1990 os níveis de emprego em Nova Jersey e Pensilvânia eram bastante semelhantes, todavia ao longo do ano de 1991 o nível de emprego na Pensilvânia caiu relativamente à Nova Jersey. Qual problema isto traria para a identificação do efeito causal de interesse?

Exercício 18. Suponha que a prefeitura de um município tenha criado um programa de treinamento para trabalhadores com baixa escolaridade. Considere que temos também informações sobre o grupo que não recebeu o tratamento. A Tabela abaixo apresenta dados sobre o salário médio dos grupos de tratamento e controle:

Ano	Tratamento	Controle
1994	650	525
1996	750	675
1998	950	825
2000	1100	950
2002	1200	1025

Considere ainda que o treinamento ocorreu no ano de 1999 e você foi contratado para estimar o impacto deste programa sobre os salários dos trabalhadores.

- Calcule o estimador de diferenças-em-diferenças do impacto do treinamento, definindo $t_0 = 1998$ e $t_1 = 2000$.
- Represente graficamente (i) as curvas de tendência dos salários do grupo de tratamento e controle; (ii) o momento da intervenção; e (iii) a magnitude do efeito estimado por diferenças-em-diferenças em 2000.
- Qual hipótese de identificação deve ser satisfeita para que o estimador de diferenças-em-diferenças seja uma boa estimativa causal do efeito do treinamento? Você considere que neste temos uma estimativa não-viesada do efeito causal? Justifique.
- Considere agora que ao invés de utilizar os dados na tabela acima, utilizaremos uma base de dados com informações para os dois grupos (Tratamento e Controle) e para dois períodos, com 5000 observações. Como você estimaria o efeito do programa por meio de uma regressão?

Exercício 19. Uma das grandes vantagens do método de diferenças em diferenças é que ele é capaz de lidar com a seleção decorrente de certo tipo de características não observadas dos indivíduos. Que tipo de características é esse, e como o método as incorpora na regressão?

Exercício 20. Considere o Modelo de Probabilidade Linear (MPL).

- Mostre que $E(Y_i|X_i) = \text{Prob}(Y_i = 1)$. Interprete.
- Mostre que $\text{Var}(u_i) = E(Y_i|X_i) [1 - E(Y_i|X_i)]$. Interprete.

Exercício 21. Prove que:

$$\frac{\partial P_i}{\partial x_{ijk}} = f(X\beta)\beta_k$$

Onde $f(\cdot)$ é a função densidade de probabilidade avaliada no ponto $y^* = X\beta$.

Exercício 22. Seja z_1 um vetor de variáveis contínuas, z_2 uma variável contínua e d_1 uma variável binária.

a) No modelo $Prob(y = 1|z_1, z_2) = F(z_1\delta_1 + \beta_1z_2 + \beta_2z_2^2)$, onde $F(\cdot)$ é a função de distribuição acumulada da Normal, derive o efeito marginal de z_2 sobre $Prob(y = 1)$. Explique em detalhes como você estimaria esse efeito parcial.

b) No modelo $Prob(y = 1|z_1, z_2) = F(z_1\delta_1 + \beta_1z_2 + \beta_2d_1)$, onde $F(\cdot)$ é a função de distribuição acumulada da Normal, derive o efeito marginal de d_1 sobre $Prob(y = 1)$. Explique em detalhes como você estimaria esse efeito parcial.

Exercício 23. Considere o modelo $y_i^* = \beta_0 + \beta_1x_{1i} + \beta_2x_{2i} + \dots + \beta_kx_{ki} + u_i$, onde y_i^* é uma variável latente, não observável. Observe-se a seguinte variável *dummy*:

$$y_i = \begin{cases} 1, & \text{se } y_i^* > 0 \\ 0, & \text{caso contrário} \end{cases}$$

a) Prove que $Prob(Y_i = 1) = 1 - F[-(\beta_0 + \beta_1x_{1i} + \beta_2x_{2i} + \dots + \beta_kx_{ki})]$, onde F é uma função de distribuição acumulada.

b) Como o modelo pode ser simplificado se a função de densidade f for simétrica?

c) Qual a diferença entre os modelos Probit e Logit?

Exercício 24. Mostre qual é a função de log-verossimilhança no modelo Probit. Prove. Obtenha as condições de primeira ordem para o modelo com constante e k variáveis explicativas.

Exercício 25. (i) Para uma variável de resposta binária y , seja $\bar{y}$ a proporção de uns na amostra (o que é igual a média amostral de y_j). Seja $\hat{q}_0$ o percentual corretamente predito para $y = 0$ e seja $\hat{q}_1$ o percentual corretamente predito para $y = 1$. Se $\hat{p}$ é o total corretamente predito, mostre que $\hat{p}$ é uma média ponderada de $\hat{q}_0$ e $\hat{q}_1$:

$$\hat{p} = (1 - \bar{y})\hat{q}_0 + \bar{y}\hat{q}_1$$

(ii) Numa amostra de 300, suponha que $\bar{y} = 0,70$, tal que temos 210 resultados com $y_i = 1$ e 90 com $y_i = 0$. Suponha que o percentual corretamente predito quando $y = 0$ é 80, e o percentual corretamente predito quando $y = 1$ é 40. Encontre a porcentagem total corretamente predita.

Exercício 26. O efeito parcial de um modelo PROBIT pode ser escrito pela seguinte fórmula, considerando que x_k é um vetor de variáveis contínuas.

$$\frac{\partial P(Y = 1)}{\partial x_k} = \phi(x'\beta)\beta_k$$

Em que $\phi(x'\beta) = \frac{\partial \Phi(x'\beta)}{\partial x'\beta}$ Responda aos seguintes itens:

- a Qual é o efeito parcial da variável x_k sobre Y no seguinte modelo linear que é estimado via MQO?

$$Y_i = \alpha + \alpha_1 x_1 + \cdots + \alpha_k x_k + \varepsilon_i$$

Esse efeito é constante no domínio de x_k ou ele é variável?

- b Considerado que $\phi(x'\beta)$ é a derivada da distribuição normal em seu domínio, ou seja, a derivada de $\Phi(x'\beta)$, demonstre que $\phi(x'\beta)$ neste caso não é constante em seu domínio. Confira uma intuição a essa derivação.
- c Considere agora duas alternativas para a estimação dos efeitos parciais. Dado que β_k , quando estimado de forma consistente, é simplesmente um parâmetro, foquemos na função $\phi(x'\beta)$. Uma escolha precisa ser feita em relação a que vetor x considerar na análise. Considere que usemos $x = \bar{x}$, ou seja, calculemos os efeitos parciais no modelo PROBIT na média (também conhecido como Partial Effects at Average-PEA). Considere que x_k seja uma variável dicotômica. Se $\bar{x}_k \in (0, 1)$, esse efeito parcial calculado de fato existe para alguma observação da amostra?
- d Agora, considere que ao invés de considerarmos algum vetor x específico para o cálculo do efeito parcial, façamos: $\frac{\partial P(Y=1)}{\partial x_k} = E(\phi(x'\beta))\beta_k$. Neste caso, ainda temos o problema de calcular algum efeito parcial para alguma observação que não existe em nossa amostra? Qual o revés de usar um valor único para representar todo o efeito parcial da variável explicativa sobre a variável dependente?

Exercício 27. As seguintes observações foram sorteadas de forma aleatória de uma distribuição normal censurada:

3,83	7,20	0,00	0,00
4,41	8,02	5,78	7,08
0,00	0,80	13,07	4,32
0,00	8,68	5,46	0,00
8,10	0,00	1,25	5,60

Considere o modelo Tobit a seguir:

$$\begin{aligned} y_i^* &= \mu + \epsilon_i \\ y_i &= y_i^* \text{ se } \mu + \epsilon_i > 0 \\ y_i &= 0 \text{ caso contrário} \\ \epsilon_i &\sim N(0, \sigma^2) \end{aligned}$$

- a) Qual seria o estimador de MQO nesse contexto? Ele estaria subestimando ou superestimando μ ? Discuta.
- b) E se usássemos apenas as observações positivas, qual problema teríamos? Qual seria o estimador de MQO? Estaríamos subestimando ou superestimando?
- c) Monte a função verossimilhança para estimar um modelo Tobit, explicando cada termo.

Exercício 28. Considere o seguinte modelo Tobit:

$$\begin{aligned} y &= \max\{y^*, 0\} \\ y^* &= x\beta + \sigma u \\ u &\sim N(0, 1) \end{aligned}$$

- Obtenha $E[y^*|x]$ e $\frac{\partial E[y^*|x]}{\partial x}$.
- Obtenha $E[y|x]$.
- Obtenha $\frac{\partial E[y|x]}{\partial x}$. O resultado obtido nesse item é igual ao obtido no item anterior?
- Faz sentido aplicar MQO para estimar β ?

Exercício 29. (i) Suponha um modelo Tobit onde $x_1 = \log(z_1)$, e este é o único lugar onde z_1 aparece em x . Mostre que:

$$\frac{\partial E(y|y > 0, x)}{\partial z_1} = (\beta_1/z_1)\{1 - \lambda(x\beta/\sigma)[x\beta + \lambda(x\beta/\sigma)]\}$$

onde β_1 é o coeficiente de $\log(z_1)$.

(ii) Se $x_1 = z_1$ e $x_2 = z_1^2$, mostre que:

$$\frac{\partial E(y|y > 0, x)}{\partial z_1} = (\beta_1 + 2\beta_2 z_1)\{1 - \lambda(x\beta/\sigma)[x\beta + \lambda(x\beta/\sigma)]\}$$

onde β_1 é o coeficiente de z_1 e β_2 é o coeficiente de z_1^2 .

Exercício 30. Sobre o modelo Poisson dado pela equação $\frac{e^{-\lambda} \lambda_i^{y_i}}{y_i!}$:

- Explique por que o modelo de Poisson é associado a análise de eventos raros.
- Derive a função de máxima verossimilhança desse modelo e encontre o estimador de λ .
- Assuma que $\lambda = e^{\beta_0 + \beta_1 x_{1i}}$. Substitua essa expressão no modelo e ache as CPOs para estimar os β s.
- Qual a premissa referente a λ ? Se essa premissa não for válida, como devemos proceder?

Exercício 31. Seja *patentes* o número de patentes desenvolvidas por uma firma num dado ano. Assuma que a esperança condicional de *patentes* dado *vendas* e *PD* é:

$$E(\text{patentes}|\text{vendas}, PD) = \exp[\beta_0 + \beta_1 \log(\text{vendas}) + \beta_2 PD + \beta_3 PD^2]$$

onde *vendas* são as vendas anuais da firma e *PD* é o gasto total em pesquisa e desenvolvimento nos últimos 10 anos.

- Como você estimaria β_j ? Justifique sua resposta discutindo a natureza de *patentes*.
- Como você interpreta β_1 ?
- Encontre o efeito parcial de *PD* em $E(\text{patentes}|\text{vendas}, PD)$.

Exercício 32. Considere um modelo de Poisson dado pela seguinte função de densidade condicional: $f(y_i|x_i, \beta) = \frac{e^{-\lambda} \lambda_i^{y_i}}{y_i!}$:

Uma hipótese relevante à distribuição de Poisson é que a média de uma variável que segue a distribuição de Poisson é idêntica à sua variância. Vamos neste item pensar em termos de consistência e eficiência de uma regressão de máxima verossimilhança no contexto de um modelo de Poisson condicional.

- Argumente pela consistência do estimador de máxima verossimilhança para o modelo de Poisson dado no enunciado do exercício (dica: na derivação do estimador, você invoca em algum momento a propriedade de equidispersão da distribuição condicional de Poisson?)

- b Agora, em termos de consistência. Salvaguardadas algumas hipóteses mais técnicas, sabemos que o estimador de máxima verossimilhança é eficiente, dada a validade igualdade informacional de Fisher. Em grosso modo, a distribuição condicional pela qual modelamos os dados é de fato verdadeira no sentido de os dados terem sido gerados a partir dela. Caso haja overdispersion nos dados do nosso modelo de Poisson, isso nos impede de utilizar a regressão de máxima verossimilhança com o modelo de Poisson? Se sim (ou não), por que isso ocorre?