Macroeconomia Aberta e DSGE: Fundamentos, Estimação e Aplicações

Introdução à Econometria Bayesiana

João Ricardo Costa Filho

Sobre modelos

Good ideas shine far more brightly when supported by good models
Avinash Dixit (“The making of Economic Policy", 1996, p. 17)

All models are wrong.
George Box

Models are to be used, not believed.
Henri Theil (“Principles of Econometrics", 1971, p. vi)

The Bayesian way of thinking

Probabilidade

"The most important questions of life are, for the most part, really only problems in probability."

Laplace (1812)

Probabilidade na. . . vida?

Probabilidade na. . . vida?

- George Harrison, Ernest Hemingway, David Hume e Han Solo.

Probabilidade na... vida?

- George Harrison, Ernest Hemingway, David Hume e Han Solo.
- Qual é a probabilidade do sol nascer amanhã?

Probabilidade na... vida?

- George Harrison, Ernest Hemingway, David Hume e Han Solo.
- Qual é a probabilidade do sol nascer amanhã?
- Qual é a probabilidade do sol nascer amanhã, dado que ela nasceu hoje?

Probabilidade na. . . vida?

- George Harrison, Ernest Hemingway, David Hume e Han Solo.
- Qual é a probabilidade do sol nascer amanhã?
- Qual é a probabilidade do sol nascer amanhã, dado que ela nasceu hoje?
- Qual é a probabilidade de conseguir atravessar um campo de asteróides?

Probabilidade na... vida?

- George Harrison, Ernest Hemingway, David Hume e Han Solo.
- Qual é a probabilidade do sol nascer amanhã?
- Qual é a probabilidade do sol nascer amanhã, dado que ela nasceu hoje?
- Qual é a probabilidade de conseguir atravessar um campo de asteróides? *Don't tell me the odds!*.

Probabilidade

Da definição de probabilidade condicional,

Da definição de probabilidade condicional,

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

Da definição de probabilidade condicional,

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

temos as seguintes propriedades:

Probabilidade

Da definição de probabilidade condicional,

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

temos as seguintes propriedades:

- $P(A|B)P(B) = P(B|A)P(A)$;

Da definição de probabilidade condicional,

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

temos as seguintes propriedades:

- $P(A|B)P(B) = P(B|A)P(A)$;
- $\forall$ eventos mutualmente excludentes e exaustivos $A_1, A_2, \ldots, A_n$, temos:

Probabilidade

Da definição de probabilidade condicional,

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

temos as seguintes propriedades:

- $P(A|B)P(B) = P(B|A)P(A)$;
- $\forall$ eventos mutualmente excludentes e exaustivos $A_1, A_2, ..., A_n$, temos:
  - $P(A_1|B) + P(A_2|B) + ... + P(A_n|B) = 1$

Probabilidade

Da definição de probabilidade condicional,

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

temos as seguintes propriedades:

- $P(A|B)P(B) = P(B|A)P(A)$;
- $\forall$ eventos mutualmente excludentes e exaustivos $A_1, A_2, ..., A_n$, temos:
  - $P(A_1|B) + P(A_2|B) + ... + P(A_n|B) = 1$
  - $\sum_{i=1}^{n} P(B|A_i)P(A_i) = P(B)$

Probabilidade

Dois eventos A e B são independentes se $P(A|B) = P(A)$.

Probabilidade

Dois eventos A e B são independentes se $P(A|B) = P(A)$.

- Se A e B são independentes, então $P(A \cap B) = P(A) \cdot P(B)$;

Probabilidade

Dois eventos A e B são independentes se $P(A|B) = P(A)$.

- Se A e B são independentes, então $P(A \cap B) = P(A) \cdot P(B)$;
- Se A e B são independentes, então A e B^c também são independentes, i.e., $P(A \cap B^c) = P(A) \cdot P(B^c)$.

Teorema de Bayes

Teorema de Bayes

Sejam os eventos $A_1, A_2, ..., A_n$ mutualmente excludentes e exaustivos no espaço amostral S e B a realização de um evento no espaço amostral Z a realização de um evento.

Teorema de Bayes

Sejam os eventos $A_1, A_2, ..., A_n$ mutualmente excludentes e exaustivos no espaço amostral S e B a realização de um evento no espaço amostral Z a realização de um evento. Pela regra da multiplicação, sabemos que $P(A_i \cap B) = P(B|A_i) \cdot P(A_i)$.

Teorema de Bayes

Sejam os eventos $A_1, A_2, ..., A_n$ mutualmente excludentes e exaustivos no espaço amostral S e B a realização de um evento no espaço amostral Z a realização de um evento. Pela regra da multiplicação, sabemos que $P(A_i \cap B) = P(B|A_i) \cdot P(A_i)$. Substituindo na definição da probabilidade condicional, temos:

Teorema de Bayes

Sejam os eventos $A_1, A_2, ..., A_n$ mutualmente excludentes e exaustivos no espaço amostral S e B a realização de um evento no espaço amostral Z a realização de um evento. Pela regra da multiplicação, sabemos que $P(A_i \cap B) = P(B|A_i) \cdot P(A_i)$. Substituindo na definição da probabilidade condicional, temos:

$$P(A|B) = \frac{P(B|A_i) \cdot P(A_i)}{P(B)}$$

Teorema de Bayes

Teorema de Bayes

Pelos axiomas da probabilidade e pela regra da multiplicação, sabemos que

Teorema de Bayes

Pelos axiomas da probabilidade e pela regra da multiplicação, sabemos que

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + ... + P(B \cap A_n)$$
$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + ... + P(B|A_n)P(A_n)$$
$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Teorema de Bayes

Pelos axiomas da probabilidade e pela regra da multiplicação, sabemos que

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + ... + P(B \cap A_n)$$
$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + ... + P(B|A_n)P(A_n)$$
$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Assim, obtemos:

Teorema de Bayes

Pelos axiomas da probabilidade e pela regra da multiplicação, sabemos que

$$
\begin{aligned}
P(B) &= P(B \cap A_1) + P(B \cap A_2) + ... + P(B \cap A_n) \\
P(B) &= P(B|A_1)P(A_1) + P(B|A_2)P(A_2) + ... + P(B|A_n)P(A_n) \\
P(B) &= \sum_{i=1}^{n} P(B|A_i)P(A_i)
\end{aligned}
$$

Assim, obtemos:

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

Aplicação – Princípio da Escolha Restrita

(Adaptado de Miller and Sanjurjo 2019) Você possui um restaurante com dois clientes habituais: Ann e Bob. Sabe-se que

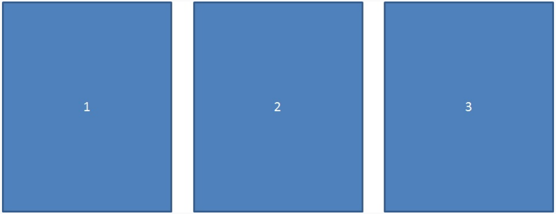
- Ann é indiferente entre as dez opções do cardápio.
- Bob prefere o hambúrguer e o pede sempre.

A cozinha recebe um pedido de hambúrguer. Qual é a probabilidade de ser o Bob? E de ser a Ann?

Aplicação – Monty Hall Problem

O apresentador de um programa na televisão te oferece a oportunidade de ganhar um carro novo. Para isso, basta acertar atrás de qual das três portas ele está. Após você escolher uma das portas, o apresentador abre um das outras duas portas, revelando que atrás dela havia um bode e te oferece a seguinte escolhe: trocar de porta ou manter a escolha. O que você faz?

Exercício – Monty Hall Problem



Aplicação: bull vs bear market

Trabalhemos com o exemplo 11.18 de Bussab and Morettin (2015). Seja y a variação percentual mensal do Ibovespa. Assuma que o mercado possa ter dois "estados": alta (bull market) e baixa (bear market).

Aplicação: bull vs bear market

Trabalhemos com o exemplo 11.18 de Bussab and Morettin (2015). Seja y a variação percentual mensal do Ibovespa. Assuma que o mercado possa ter dois "estados": alta (bull market) e baixa (bear market). O primeiro estado será designado por θ_1 ao passo que segundo será por θ_2.

Aplicação: bull vs bear market

Assuma que tenhamos a seguinte informação (*a priori*) sobre as probabilidade de cada um dos eventos:

Aplicação: bull vs bear market

Assuma que tenhamos a seguinte informação (*a priori*) sobre as probabilidade de cada um dos eventos:

priori	θ_1	θ_2
$p(\theta)$	$\frac{3}{5}$	$\frac{2}{5}$

Aplicação: bull vs bear market

As funções de verossimilhança são dadas por: $P(y > 0|\theta)$ e $P(y < 0|\theta)$.

Aplicação: bull vs bear market

As funções de verossimilhança são dadas por: $P(y > 0|\theta)$ e $P(y < 0|\theta)$. Assuma também que conheçamos as funções de verossimilhança e que elas sejam:

Aplicação: bull vs bear market

As funções de verossimilhança são dadas por: $P(y > 0|\theta)$ e $P(y < 0|\theta)$. Assuma também que conheçamos as funções de verossimilhança e que elas sejam:

$y \backslash \theta$	$p(y\|\theta)$ θ_1	θ_2
$y > 0$	$\frac{2}{3}$	$\frac{1}{3}$
$y < 0$	$\frac{1}{3}$	$\frac{2}{3}$

Aplicação: bull vs bear market

Assim, podemos calcular as probabilidades conjuntas $P(y, \theta) = P(\theta) \cdot P(y|\theta)$.

Aplicação: bull vs bear market

Assim, podemos calcular as probabilidades conjuntas $P(y,\theta) = P(\theta) \cdot P(y|\theta)$. Por exemplo, $P(y > 0, \theta = \theta_1 = P(\theta = \theta_1) \cdot P(y > 0|\theta = \theta_1) = 3/5 \cdot 2/3 = 6/15$.

Aplicação: bull vs bear market

Assim, podemos calcular as probabilidades conjuntas $P(y,\theta) = P(\theta) \cdot P(y|\theta)$. Por exemplo, $P(y > 0, \theta = \theta_1 = P(\theta = \theta_1) \cdot P(y > 0|\theta = \theta_1) = 3/5 \cdot 2/3 = 6/15$.

θ / y	$p(y,\theta)$		$p(y)$
	θ_1	θ_2	
$y > 0$	$\frac{6}{15}$	$\frac{2}{15}$	$\frac{8}{15}$
$y < 0$	$\frac{3}{15}$	$\frac{4}{15}$	$\frac{7}{15}$
$p(\theta)$	$\frac{9}{15}$	$\frac{6}{15}$	1

Aplicação: bull vs bear market

Podemos calcular também:

Aplicação: bull vs bear market

Podemos calcular também:

$$P(y > 0) = \sum_{\theta} P(y, \theta) = \sum_{\theta} P(\theta) \cdot P(y|\theta)$$

Aplicação: bull vs bear market

Podemos calcular também:

$$P(y > 0) = \sum_{\theta} P(y, \theta) = \sum_{\theta} P(\theta) \cdot P(y|\theta)$$

Por exemplo,

$$P(y > 0) = P(\theta_1) \cdot P(y > 0|\theta_1) + P(\theta_2) \cdot P(y > 0|\theta_2)$$
$$P(y > 0) = 3/5 \cdot 2/3 + 2/5 \cdot 1/3 = 8/15$$

Aplicação: bull vs bear market

As probabilidades marginais são dadas por:

y	$p(y)$
$y > 0$	$\frac{8}{15}$
$y < 0$	$\frac{7}{15}$

A partir da definição de probabilidade condicional, temos:

Aplicação: bull vs bear market

A partir da definição de probabilidade condicional, temos:

$$P(\theta = \theta_1 | y > 0) = \frac{P(\theta_1) \cdot P(y > \theta | \theta_1)}{P(y > 0)} = \frac{3/5 \cdot 2/3}{8/15} = 3/4$$

$$P(\theta = \theta_2 | y > 0) = \frac{P(\theta_2) \cdot P(y > \theta | \theta_2)}{P(y > 0)} = 1/4$$

Aplicação: bull vs bear market

A partir da definição de probabilidade condicional, temos:

$$P(\theta = \theta_1 | y > 0) = \frac{P(\theta_1) \cdot P(y > \theta | \theta_1)}{P(y > 0)} = \frac{3/5 \cdot 2/3}{8/15} = 3/4$$
$$P(\theta = \theta_2 | y > 0) = \frac{P(\theta_2) \cdot P(y > \theta | \theta_2)}{P(y > 0)} = 1/4$$

Da mesma forma, temos que $P(\theta = \theta_1 | y < 0) = 3/7$,
$P(\theta = \theta_2 | y < 0) = 4/7$.

Aplicação: bull vs bear market

θ \ y	$p(\theta \mid y)$	
	θ_1	θ_2
$y > 0$	$\frac{3}{4}$	$\frac{1}{4}$
$y < 0$	$\frac{3}{7}$	$\frac{4}{7}$

Estatística Bayesiana

"In God we trust. All others must bring data."
William Edwards Deming

Distribuições a priori e a posteriori

Distribuições a priori e a posteriori

Exemplo de Albert (2009):

Exemplo de Albert (2009):

- Os médicos recomendam oito horas de sono, em média para um adulto, por noite. Qual será a proporção (p) dos alunos de graduação que dormem pelo menos oito horas por noite?

Distribuições a priori e a posteriori

Exemplo de Albert (2009):

- Os médicos recomendam oito horas de sono, em média para um adulto, por noite. Qual será a proporção (p) dos alunos de graduação que dormem pelo menos oito horas por noite?
- Há incerteza sobre o valor populacional de p e através de uma amostra aleatória de estudantes podemos abordar essa questão.

Exemplo de Albert (2009):

- Os médicos recomendam oito horas de sono, em média para um adulto, por noite. Qual será a proporção (p) dos alunos de graduação que dormem pelo menos oito horas por noite?
- Há incerteza sobre o valor populacional de p e através de uma amostra aleatória de estudantes podemos abordar essa questão.
- Pesquisas prévias: a proporção está entre 0 e $0,5$.

Exemplo de Albert (2009):

- Os médicos recomendam oito horas de sono, em média para um adulto, por noite. Qual será a proporção (p) dos alunos de graduação que dormem pelo menos oito horas por noite?
- Há incerteza sobre o valor populacional de p e através de uma amostra aleatória de estudantes podemos abordar essa questão.
- Pesquisas prévias: a proporção está entre 0 e $0,5$.
- O pesquisador coleta uma amostra com 27 alunos, 11 deles dormem ao menos oito horas.

Distribuições a priori e a posteriori

Exemplo de Albert (2009):

- Os médicos recomendam oito horas de sono, em média para um adulto, por noite. Qual será a proporção (p) dos alunos de graduação que dormem pelo menos oito horas por noite?
- Há incerteza sobre o valor populacional de p e através de uma amostra aleatória de estudantes podemos abordar essa questão.
- Pesquisas prévias: a proporção está entre 0 e $0,5$.
- O pesquisador coleta uma amostra com 27 alunos, 11 deles dormem ao menos oito horas.
- Qual é o número esperado, em uma nova amostra com 20 alunos, de alunos que dormem pelo menos oito horas por noite?

Distribuições a priori e a posteriori

Distribuições a priori e a posteriori

A função de de verossimilhança $(L(p))$ é proporcional a $(\propto)$:

$$L(p) \propto p^s(1-p)^f, 0 < p < 1$$

onde s é o número de sucessos (alunos que dormem ao menos oito horas por noite) e f o número de fracassos. A distribuição a priori de p é dada por $g(p)$.

Distribuições a priori e a posteriori

A função de de verossimilhança ($L(p)$) é proporcional a ($\propto$):

$$L(p) \propto p^s(1-p)^f, 0 < p < 1$$

onde s é o número de sucessos (alunos que dormem ao menos oito horas por noite) e f o número de fracassos. A distribuição a priori de p é dada por $g(p)$. A distribuição a posteriori, a partir do Teorema de Bayes, com base nos dados (g(p | dados)) é dada por:

Distribuições a priori e a posteriori

A função de de verossimilhança ($L(p)$) é proporcional a ($\propto$):

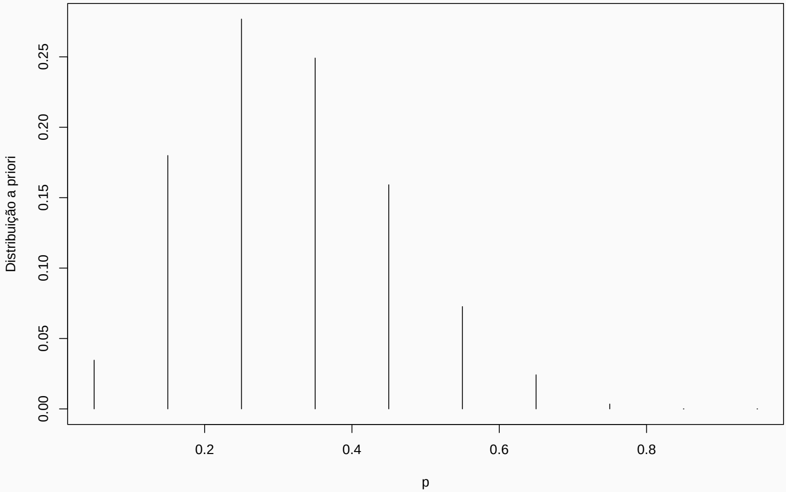
$$L(p) \propto p^s(1-p)^f, 0 < p < 1$$

onde s é o número de sucessos (alunos que dormem ao menos oito horas por noite) e f o número de fracassos. A distribuição a priori de p é dada por $g(p)$. A distribuição a posteriori, a partir do Teorema de Bayes, com base nos dados (g(p | dados)) é dada por:

$$g(p|\text{dados}) \propto g(p)L(p)$$

Como definir a distribuição a priori?

Prior Discreta



Função de verossimilhança

Na amostra, 11 dos 27 alunos dormem horas suficientes (assim, $s = 11$ e $f = 16$).

Função de verossimilhança

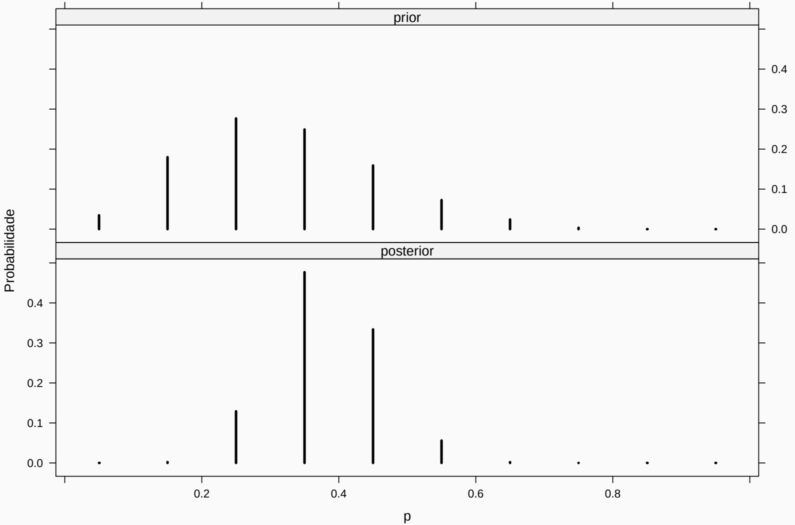
Na amostra, 11 dos 27 alunos dormem horas suficientes (assim, $s = 11$ e $f = 16$). A função de verossimilhança pode ser escrita como:

Função de verossimilhança

Na amostra, 11 dos 27 alunos dormem horas suficientes (assim, $s = 11$ e $f = 16$). A função de verossimilhança pode ser escrita como:

$$L(p) \propto p^{11}(1-p)^{16}, 0 < p < 1$$

Prior Discreta



Prior Discreta

Tomemos o caso quando $p = 0,25$:

Tomemos o caso quando $p = 0,25$:

$$P(p = 0,25|dados) = \frac{P(p = 0,25) \cdot P(dados|p = 0,25)}{\sum_i P(dados|p = p_i) \cdot P(p = p_i)}$$

$$P(p = 0,25|dados) = \frac{0,28 \cdot [\binom{27}{11} \cdot 0,25^{11} \cdot (1 - 0,25)^{27-11}]}{0,066969} = 0,13$$

Beta Prior

Beta Prior

Como $p \in [0, 1]$, podemos reescrever a função de densidade $g(p)$ como:

Beta Prior

Como $p \in [0, 1]$, podemos reescrever a função de densidade $g(p)$ como:

$$g(p) \propto p^{a-1}(1-p)^{b-1}, 0 < p < 1$$

onde $p \sim Beta(a, b)$ e os parâmetros a e b são escolhidos com base nas crenças a priori do pesquisador.

Beta Prior

Como $p \in [0,1]$, podemos reescrever a função de densidade $g(p)$ como:

$$g(p) \propto p^{a-1}(1-p)^{b-1}, 0 < p < 1$$

onde $p \sim Beta(a, b)$ e os parâmetros a e b são escolhidos com base nas crenças a priori do pesquisador. No exemplo, a pesquisadora acredita que a mediana da distribuição está em 0,3 e tem 90% de confiança de que o valor é menor do que 0,5.

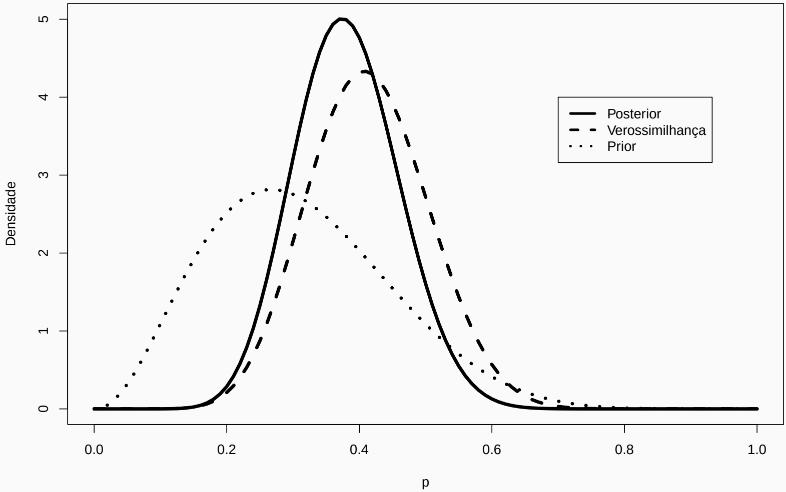
Beta Prior

$$\int_0^{0,3} \frac{1}{Beta(a,b)} x^{a-1} \cdot (1-x)^{b-1} dx = 0,5$$
$$\int_0^{0,5} \frac{1}{Beta(a,b)} x^{a-1} \cdot (1-x)^{b-1} dx = 0,9$$

Podemos combinar a *beta prior* com a função de verossimilhança e escrever a função de densidade posterior como

$$g(p|\text{ data }) = \frac{\left[\frac{1}{Beta(a,b)} p^{a-1}(1-p)^{b-1}\right] \cdot p^s(1-p)^f}{\int_0^1 \left[\frac{1}{Beta(a,b)} p_i^{a-1}(1-p_i)^{b-1}\right] \cdot \left[p_i^s(1-p_i)^f\right]}, 0 < p < 1$$
$$g(p|\text{ data }) \propto p^{a+s-1}(1-p)^{b+f-1}, 0 < p < 1$$

Beta Prior



Posterior
Verossimilhança
Prior
Densidade
p
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5

OK, tell me odds.

Econometria Bayesiana e DSGE

Uma muito (muito) breve introdução

Tópicos que abordaremos nas próximas aulas (An and Schorfheide 2007; Herbst and Schorfheide 2016):

- Modelos lineares vs modelos não-lineares.
- Representação espaço-estado (veja Costa-Filho (2022) sobre como recuperar essa representação no Dynare).
- Priors: distribuições diferentes para parâmetros diferentes (e.g.: Beta, Gama, Normal, Gama Inversa e Uniforme).
- Métodos númericos.

Uma muito (muito) breve aplicação

Trabalhemos com uma versão simplificada do modelo com dois países idênticos de Kim and Kim (2003) feita por Barillas et al. (2019):

Uma muito (muito) breve aplicação

Trabalhemos com uma versão simplificada do modelo com dois países idênticos de Kim and Kim (2003) feita por Barillas et al. (2019):

$$
\begin{aligned}
C_{1,t} &= C_{2,t} \\
C_{1,t}^{-\gamma} &= \beta \mathbb{E}_t \left[C_{1,t+1}^{-\gamma} \left(\alpha A_{1,t+1} K_{1,t+1}^{\alpha-1} + 1 - \delta \right) \right] \\
C_{2,t}^{-\gamma} &= \beta \mathbb{E}_t \left[C_{2,t+1}^{-\gamma} \left(\alpha A_{2,t+1} K_{2,t+1}^{\alpha-1} + 1 - \delta \right) \right] \\
A_{1,t} K_{1,t}^{\alpha} + A_{2,t} K_{2,t}^{\alpha} &= C_{1,t} + C_{2,t} + K_{1,t-1} - (1-\delta) K_{1,t} + K_{2,t-1} - (1-\delta) K_{2,t} \\
\ln A_{1,t} &= \rho \ln A_{1,t-1} + \epsilon_{1,t} \\
\ln A_{2,t} &= \rho \ln A_{2,t-1} + \epsilon_{2,t}
\end{aligned}
$$

Referências i

Albert, J. 2009. "Bayesian Computation with r Springer Science & Business Media." *New York*.

An, Sungbae, and Frank Schorfheide. 2007. "Bayesian Analysis of DSGE Models." *Econometric Reviews* 26 (2-4): 113–72.

Barillas, Francisco, Anmol Bhandari, Riccardo Colacito, Sagiri Kitao, Christian Matthes, Thomas J. Sargent, and Yongseok Shin. 2019. *Practicing Dynare 4.5.6*. New York University, Department of Economics; University of North Carolina, Department of Finance; University of Southern California, FBE Department; Washington University, Department of Economics. https://www.cimers.org/dynare-handbook.

Bussab, Wilton de Oliveira, and Pedro Alberto Morettin. 2015. *Estatística Básica*. 8th ed. São Paulo: Saraiva Educação.

Costa-Filho, João. 2022. "Retrieving the State-Space Representation from Dynare." ISEG-Lisbon School of Economics; Management, REM, Universidade de Lisboa.

Herbst, Edward P, and Frank Schorfheide. 2016. *Bayesian Estimation of DSGE Models*. Princeton University Press.

Kim, Jinill, and Sunghyun Henry Kim. 2003. "Spurious Welfare Reversals in International Business Cycle Models." *Journal of International Economics* 60 (2): 471–500.

Referências iii

Miller, Joshua B, and Adam Sanjurjo. 2019. "A Bridge from Monty Hall to the Hot Hand: The Principle of Restricted Choice." *Journal of Economic Perspectives* 33 (3): 144–62.