



Código:

2

1. INTRODUÇÃO

A Economia é a ciência que visa obter da melhor forma possível recursos da sociedade como um todo. Uma importante área de estudo da Economia é a Econometria. Isso ocorre porque as ~~(teorias)~~ teorias econômicas são criadas a partir da Filosofia e ~~(matemática econômica)~~ Matemática econômica. Ou seja, a Econometria deve testar empiricamente as teorias criadas de forma "a priori".

A Econometria é, portanto, fruto de uma junção entre Economia, Estatística, Matemática e Probabilidade, em que, muitas vezes, a Teoria econômica "diz" o que será realizado por meio das técnicas matemáticas. Sir Ronald Fisher prescreve que a estatística tenha três objetivos: 1) Estudo de variabilidade; 2) ~~Diagnóstico~~ Redução de variabilidade; e 3) Inferência Estatística.

Neste contexto, este documento visa discutir sobre a Inferência. A inferência é importante uma vez que permite fazer declarações confiáveis sobre alguma característica populacional de interesse por meio de amostras. Como nem sempre se consegue trabalhar com toda a população (seja pela dificuldade de acesso, seja pelo custo), este fato é de suma importância para a ciência de modo geral.

A partir deste ponto este documento está dividido em 4 seções: Background, Distribuições Amostrais e TLC, Inferência e Testes de Hipóteses.

2. BACKGROUND

Antes de entrarmos de fato na teoria deste documento, é importante definirmos alguns conceitos importantes. O primeiro é Experimento Aleatório. Um Experimento Aleatório é qualquer experimento que pode ter quaisquer valores diferentes no momento que seja repetido da mesma forma. Espaço Amostral são todos os possíveis resultados de um Experimento Aleatório.

Define-se também variável Aleatória Discreta ~~(ou seja uma variável que assume valores discretos)~~ variáveis contínuas. (Ou seja, ~~as~~ definidas em espaços finitos). Uma função de probabilidade é a descrição de uma variável Aleatória Discreta, e é definida como $f(x) = P(X=x)$; $f(n) \geq 0$; $\sum f(n) = 1$.

Uma variável Aleatória Contínua é definida sobre o conjunto $\mathbb{R}$ (ou $\mathbb{R}^n$). Nunca ocorre, tendo as propriedades: ~~$f(x) \geq 0$~~ ~~$f(x) \geq 0$~~ ~~$f(x) \geq 0$~~

EM BRANCO



Código:

2

$f(x) \geq 0$; $\int_{-\infty}^{\infty} f(x) dx = 1$; $P(a \leq x \leq b) = \int_a^b f(x) dx$. Define-se $E(X) = \int_{-\infty}^{\infty} x f(x) dx$ e $VAR(X) = \int_{-\infty}^{\infty} x^2 f(x) dx$.

3. DISTRIBUIÇÕES AMPSTRAIS E TEOREMA CENTRAL DO LIMITE

Um importante conceito dentro da ~~teoria da~~ Teoria da estatística é que as amostras que retiramos são aleatórias. De maneira, as estatísticas (calculadas) amostrais calculadas são aleatórias, possuindo suas respectivas distribuições de probabilidade. Seja $(x_1, x_2, \dots, x_n)$ uma amostra retirada de uma população com média μ e variância σ^2 . Pode-se demonstrar que a média amostral possui uma distribuição amostral (distribuição de estatísticas amostrais), quando n é suficientemente grande, normal. Ou seja, $\bar{X} \sim N(\mu, EP(\bar{X}))$, onde $EP(\bar{X}) = \frac{\sigma^2}{n}$.

O TLC é de uma importância para a estatística inferencial e todos de hipótese a serem vistos neste documento. Padronizando, temos $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$.

4. INFERÊNCIA

Como exposto, a inferência busca realizar conclusões sobre algum parâmetro populacional de interesse, levando em consideração as evidências apontadas pela amostra.

Quando falamos de inferência, geralmente estamos interessados, no contexto analisado, em estimação e testes de hipótese. No primeiro caso, podemos dividir os métodos de estimação entre estimação pontual e inferencial. Já os ~~testes de hipótese~~ testes de hipótese buscam tomar decisões sobre os parâmetros populacionais.

4.1. ESTIMADOR PONTUAL

Um estimador pontual é qualquer função $W(X)$ de uma amostra, cujo o interesse é realizar uma estimação sobre algum parâmetro θ . É importante dizermos um estimador de uma estatística, que é um valor $\hat{\theta}$, resultado de $W(X)$.

Há estimadores óbvios para alguns parâmetros populacionais, como por exemplo a média amostral para a média populacional, ou a variância amostral para a variância populacional. Entretanto, há casos que confiam na inferência (pela) para

Código: _____

EM BRANCO



Código:

2

levar a estimadores com propriedades sérias em suas propriedades. Assim, é útil ter as técnicas comuns para definir esses estimadores. Dentre os mais importantes estão: método dos momentos; máxima verossimilhança e estimadores bayesianos. Um outro estimador importante é o de Mínimos Quadrados Corridos, que não será exposto aqui, uma vez que é muito utilizado no contexto de regressão.

4.1.1. MÉTODO DOS MOMENTOS

A ideia do método dos momentos é igualar os momentos populacionais com momentos amostrais, e ~~resolver~~ resolver, para o parâmetro, o sistema de equações resultante. Define-se (momento populacional k -ésimo) o k -ésimo momento populacional como:

$$\mu_k = \int_{-\infty}^{\infty} x^k f(x) dx$$

e o k -ésimo momento amostral como:

$$m_k = \frac{1}{n} \sum x_i^k, \quad M = \text{tamanho amostral}$$

4.1.2. MÁXIMA VEROSSIMILHANÇA

A ideia de máxima verossimilhança é ajustar os parâmetros de forma a maximizar a probabilidade de que os dados amostrais tenham vindo de uma determinada população. Define-se a função verossimilhança como:

$$L(\theta) = \prod f(x_i|\theta).$$

Para deixar mais manipulável, define-se a logverossimilhança como:

$$\ell(\theta) = \sum \log [f(x_i|\theta)].$$

Calculando o vetor gradiente sobre $\ell(\theta)$, tem-se:

$$\nabla \ell(\theta) = \left(\frac{\partial \ell(\theta)}{\partial \theta_i} \right)$$

Folha nº: _____

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Código: _____

EM BRANCO



Generalmente, não sendo a solução fechada, aplicam-se algoritmos iterativos para a obtenção da soma.

4.1.3. MÉTODO BAYESIANO

Na continuação da estatística clássica, a estatística bayesiana afirma que os parâmetros possuem distribuições de probabilidade, podendo gerar inferências com a ajuda da estatística. Assim:

$$P(\theta/x) = \frac{P(x|\theta) \cdot P(\theta)}{P(x)}$$

Em que $P(\theta/x)$ é a distribuição a posteriori, $P(x|\theta)$ a verossimilhança, $P(\theta)$ é a distribuição a priori e $P(x)$ é a distribuição marginal. (Então) Então:

$$\hat{\theta}_{\text{bayes}} = \frac{\int_{-\infty}^{\infty} \theta f(\theta/x) d\theta}{\int_{-\infty}^{\infty} f(\theta/x) d\theta}$$

4.2. ESTIMAÇÃO INTERVALAR

Muitas vezes temos interesse em fazer inferências do tipo: $\theta \in \text{ME}$, em que ME é um intervalo. Então, tem-se que:

4.2.1. PARA MÉDIA (σ^2 conhecida)

$$P\left[\bar{x} - z_{\alpha/2} \cdot EP(\bar{x}) \leq \mu \leq \bar{x} + z_{\alpha/2} \cdot EP(\bar{x})\right] = 1 - \alpha,$$

em que α é o nível de confiança. Então:

$$IC_{1-\alpha} = \bar{x} \pm z \cdot EP(\bar{x}), \text{ em que } z \text{ é o valor correspondente da tabela } z.$$

4.2.2. PARA MÉDIA (σ^2 desconhecida)

→ AMOSTRAS GRANDES

$$IC_{1-\alpha} = \bar{x} \pm z \cdot \frac{s}{\sqrt{n}}, \text{ em que } z = \frac{z_{\alpha/2}}{\sqrt{n-1}}, \text{ a estatística } z \text{ é}$$

enunciada da variância.

Folha nº:

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Código:

EM BRANCO



4.1. PEQUENAS AMOSTRAS

IC_(μ) = $\bar{x} \pm t_{n-1}$, $EP(\bar{x})$, $EP(\bar{x}) = \frac{s}{\sqrt{n}}$. A distribuição t tem caudas mais pesadas, por isso a crítica repete-se.

4.2.3. PARA A VARIÂNCIA

Tem-se que $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2$. Então:

$$P\left(\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}; \sigma^2; \frac{(n-1)s^2}{\chi^2_{\alpha/2}}\right) = 1 - \alpha$$

4.2.4. PROPORÇÃO

Condição (P) e TLC, pode-se dizer que para $n\hat{p}$ e $n\hat{q} > 10$, para $\hat{p} = \frac{\sum X_i}{n}$ então $\hat{p} \sim N(\hat{p}; \sqrt{\frac{\hat{p}\hat{q}}{n}})$. Então:

$$IC_p = \hat{p} \pm z \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

5. PROPRIEDADES DOS ESTIMADORES

5.1. CONSISTÊNCIA

O estimador converge em probabilidade para o parâmetro. Ou seja:

$$\hat{\theta} \xrightarrow{p} \theta$$

5.2. AUSÊNCIA DE VIÉS

$E(\hat{\theta}) - \theta = 0$, ou seja, o estimador não se desvia do parâmetro.

5.3. EFICIÊNCIA

$VAR(\hat{\theta})$ deve ser a menor entre todos os estimadores.

6. TESTES DE HIPÓTESE

Testes de hipótese são baseados em afirmações sobre os parâmetros. Geralmente tem-se a hipótese nula e a hipótese alternativa. Formalmente, pode-se fazer:

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_0; \theta > \theta_0; \theta < \theta_0$$

Folha nº: _____

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Código: _____

EM BRANCO



UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Código:

2

Há quatro casos que se pode inferir ao testar uma hipótese:

	H_0 verdadeira	H_0 falsa
REJEITAR H_0	ERRO TIPO I	SEM ERRO
NÃO REJEITAR H_0	SEM ERRO	ERRO TIPO II

O Erro TIPO I é mais grave que o tipo II. Defina-se a probabilidade de incorrer em um Erro TIPO I como:

$$P(\text{ERRO TIPO I}) = P(\text{REJEITAR } H_0 / H_0 \text{ VERDADEIRA}) = 1 - \alpha$$

Assim, ao definir o nível de confiança, define-se o $P(\text{ERRO TIPO I})$, já o erro tipo II, tem-se:

$$P(\text{ERRO TIPO II}) = P(\text{NÃO REJEITAR } H_0 / H_0 \text{ FALSA}) = 1 - \beta$$

E $\beta [1 - P(\text{ERRO TIPO II})]$ é definido como Poder do Teste, ou seja:

$$\beta = P(\text{REJEITAR } H_0 / H_0 \text{ é FALSA})$$

Um procedimento geral de teste é: a) define-se a hipótese nula e alternativa; b) define-se a distribuição amostral; c) calcula a estatística de teste; d) compara a estatística de teste com o valor estatístico crítico. Se $K_{\text{calc}} > K_{\text{crit}}$, rejeita-se a hipótese nula.

Pode-se realizar testes de hipótese sobre quaisquer parâmetros, tendo como pressuposto a definição da distribuição amostral do teste. Para a maioria

A estatística de teste, de maneira geral, é a distância em termos de desvios-padrão (ERRO PADRÃO ou DISTURBIO AMOSTRAL). Então, o que se faz é comparar o nível de significância com a distância da estatística de teste a hipótese nula.

Folha nº:

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

Código:

EM BRANCO



Código:

2

6.1. LEMMA DE NEYMAN-PEARSON

Seja H_0 a estatística $L(x/\theta_0)$ s.r. Pode-se provar que este é o teste mais

potente, de forma que se a razão é menor que α rejeita-se a hipótese H_0 .
Este teste é a base para teste-t; teste-z e teste-F.

7. MÉTODOS NÃO PARAMÉTRICOS

Quando a distribuição é fortemente assimétrica e/ou temos poucas observações, podemos usar testes não-paramétricos, que não impõem uma distribuição específica às estatísticas do teste. Um exemplo é o ROSTRAP, que com amostras amostradas com reposição simula o processo de amostragem real e pode-se estimar os parâmetros por meio da distribuição amostral das estatísticas amostradas.

8. CONSIDERAÇÕES FINAIS

Este documento tem como objetivo apresentar sobre os métodos inferenciais. Importante mencionar que estes métodos são de grande importância para as Ciências Econômicas, uma vez que por se tratar de testes em situações reais de uma Ciência Social, o seu laboratório é a sociedade e suas amostras de teste a sociedade. Portanto, frequentemente é impossível fazer experimentos.

REFERÊNCIAS

CAVALA; BERGER. INFERÊNCIA ESTATÍSTICA.
BUSIARI; MOREIRA. ESTATÍSTICA BÁSICA.

Código:

EM BRANCO