

DA MATEMÁTICA AOS MODELOS ECONOMETRICOS: APLICAÇÃO AO ENSINO SUPERIOR ANGOLANO

Josefa Ukalango dos Santos Chiquete

Mestrado em Finanças

Professor Doutor Paulo Alexantre Botelho Rodrigues Pires

Junho de 2014



UNIVERSIDADE PORTUGALENSE

DEDICATÓRIA

Dedico este trabalho a Deus, aos meus filhos como um incentivo a formação acadêmica, aos meus pais e irmãos pela ajuda e incentivo prestados a quando da elaboração desta tese e em memória de meu querido avó Samuel Arnaldo (que faleceu dois dias antes da minha defesa) pelo incentivo que me foi prestado.

Agradecimentos

Agradeço a DEUS todo-poderoso que me deu saúde sabedoria e vida para que pudesse efetuar esta pesquisa, ao meu orientador o Dr. Paulo Pires que teve de deixar os seus afazeres para acompanhar de perto a realização desta tese, aos meus familiares em especial aos meus pais que enriqueceram a minha pesquisa concedendo-me diverso material bibliográfico, bem como ao seu apoio moral; aos irmãos e amigos Elias António, Pedro Chiquete e Victorina Naquela que me foram ajudando na pesquisa, bem como motivando sempre que eu me sentia cansada, pelo seu apoio moral e material; ao meu esposo e filhos pela compreensão que tiveram sempre que tivesse de me ausentar de casa para realizar as pesquisas necessárias e a todos que direta ou indiretamente contribuíram para esta pesquisa, a todos o meu muito obrigado.

Resumo

Com a paz efetiva no país surge uma grande necessidade de se encontrar pessoas capazes de gerir uma empresa contribuindo para o desenvolvimento económico do país. Para que haja uma correspondência positiva neste sentido, é necessário que haja pessoas formadas no campo da gestão económica e financeira. O crescente número de instituições superiores no país faz com que estas adequem os cursos a lecionar às necessidades reais do território. A fim de se saber o que se tem ensinado aos estudantes neste campo foi feita esta pesquisa que tem como foco saber como têm sido ministrados os conteúdos de econometria referentes à aplicação dos modelos econométricos para que se venha a entregar ao mercado financeiro angolano gestores capacitados a acompanhar de perto o desenvolvimento económico do país. Esta pesquisa ainda serve de base representativa daquilo que é o ensino dos modelos econométricos no ensino superior angolano.

Para tal a metodologia utilizada foi a análise comparativa dos programas das distintas instituições superiores da região em estudo, que é a região académica nº 2, de acordo com a distribuição feita pelo ministério do ensino superior angolano, tendo em conta o desenvolvimento académico nos diversos ramos do saber de cada região académica. Utilizou-se também o método da entrevista através de questionários dirigidos aos coordenadores da disciplina de Econometria. Constatou-se que na formação de um bom econometrista é fundamental o conhecimento profundo dos modelos de regressão linear simples e múltiplo, bem como os problemas que podem surgir ao fazer um estudo econométrico e como estes podem ser detetados e corrigidos para que os resultados da pesquisa seja eficaz. Fez-se, em seguida, um estudo conclusivo indicando os aspetos positivos e negativos verificados no decorrer da pesquisa com relação ao conteúdo transmitido aos estudantes para que se venha lançar no mercado de trabalho gestores de qualidades bem definidas.

Palavras – chaves: Econometria; modelos econométricos; Ensino superior; região académica nº 2.

Abstract

With effective peace in the country is a great need to find people able to manage a company contributing to the economic development of the country. So there is a positive match in this sense, there needs to be people formed in the field of economic and financial management. The increasing number of higher institutions in the country causes these suits the courses teaching to the real needs of the territory. In order to know what is taught to students in this field was made this research that focuses on how have been taught the contents of econometrics regarding application of Econometric models to come to hand over to the Angolan financial market managers able to closely monitor the country's economic development. This research also forms the basis of representative of the teaching of the Econometric models in higher education in Angola. The methodology used was the comparative analysis of the programs of the different higher institutions of the region under study, which is the academic region in 2, according to the distribution made by the Angolan Ministry of higher education, taking into account the academic development in the various branches of knowledge of each academic area. Also was used the method of interview through questionnaires addressed to chairpersons of Econometrics. It was noted that in the formation of a good econometrician is fundamental deep knowledge of simple linear regression models and multiple, as well as the problems that can arise when making an econometric study and how these can be detected and corrected so that the search results to be effective. There was then a conclusive study indicating the positives and negatives scanned in the course of research in relation to content transmitted to students so that they launch on the market of work clearly defined qualities managers.

Keywords: Econometrics; econometric models; Higher education; academic region nº 2.

Índice

Agradecimentos	
Resumo	
Abstract	
Lista de tabelas	
1. Introdução	1
2. Revisão da bibliográfica	3
2.1 Modelos econométricos	5
2.1.1 Modelos determinísticos	7
2.1.2 Modelos estocásticos	8
2.2 Fases de um estudo econométrico	10
2.3 Como usar modelos econométricos	16
2.4 Econometria e Economia Matemática	17
2.5 Problemas de uma análise de regressão	18
2.5.1 Multicolinearidade	18
2.5.2 Heteroscedasticidade	20
2.5.3 Autocorrelação	32
2.6 Resumo	47
3. Aplicação dos modelos econométricos no ensino superior angolano	49
3.1 Análise da entrevista aos coordenadores de disciplina	51
3.2 Análise dos programas de Econometria	53
4. Resultados	74
4.1 Resultado da revisão bibliográfica	74
4.2 Resultado da entrevista aos coordenadores e aos programas	75
5. Conclusões	77
Bibliografia	80
Anexos	82

Lista de tabelas

Tabela I- Cursos de gestão que lecionam a disciplina de Econometria-----	53
Tabela II- Ano curricular em que se leciona a disciplina de Econometria-----	54
Tabela III- Disciplinas inter-relacionadas ao ensino da Econometria-----	55
Tabela V- Objectivos a alcançar no ensino da Econometria por cursos-----	56
Tabela V- Bibliografia indicada no ensino da Econometria-----	60
TabelaVI-Programas de Econometria das distintas Universidades/Instituições-----	63
Tabela VII- Conteúdo matemático, estatístico e econométrico programado para os cursos de Economia e Gestão de Empresas-----	67

1. Introdução

A crescente necessidade de informações que os agentes económicos, investidores e, até mesmo, o governo do país precisa, e com a paz efetiva no país fazem com que investidores estrangeiros queiram desenvolver seus negócios em Angola devido ao crescente desenvolvimento económico e financeiro a que este país vem tendo nos últimos anos. E também com o constante surgimento de instituições do ensino superior em Angola, torna-se pertinente que se analise que tipo de quadros superiores às instituições de ensino vão fornecendo ao mercado económico e colocando no mercado de trabalho. Surge então a necessidade de se organizar a atividade comercial a todos os níveis, seja em pequenas ou grandes empresas, originando uma grande procura de indivíduos que tenham capacidade de organizar e gerir uma atividade comercial com zelo e eficácia. Contudo, o ministério do ensino superior tem procurado acompanhar o surgimento de novas instituições do ensino superior e, para tal dividiu o território em regiões.

Debruçamo-nos sobre o que se ensina na região nº2 quanto à formação de gestores capazes de responder as necessidades económicas e financeiras desta região académica, surge o tema: Da Matemática aos modelos econométricos aplicação ao ensino superior angolano “caso da região académica nº 2”. Que levou-me a uma profunda reflexão sobre o que vem a ser ensinado em termos de Econometria aos estudantes dos cursos Superiores de Gestão. Será que têm recebido bagagem suficiente para contribuírem positivamente no desenvolvimento económico da região? Como se têm preparado para encarar as dificuldades que o mercado lhes apresenta?

Em vista de dar resposta a estas questões a autora traçou os seguintes **objectivos de investigação**:

- 1- Verificar que conteúdo econométrico é ministrado aos estudantes dos cursos Superiores de Gestão;
- 2- Analisar se a bibliografia e o sistema de ensino são suficientes para uma boa formação e melhor compreensão daquilo que é o papel do econometrista na vida económica da região;
- 3- Averiguar se docentes possuem formação e conhecimentos suficientes para um ensino de qualidade.

Para a concretização desta pesquisa a autora utilizou as seguintes **metodologias**: Elaborar um plano estruturado a partir do qual deu início a recolha de dados bibliográficos em livros, revistas especializadas no ramo da gestão económica e na internet, com vista a inteirar-se do ponto de vista dos diversos autores, no que diz a teoria económica clássica e moderna, seus pontos de concordância e discórdia; conhecer as técnicas econométricas aplicáveis a uma análise económica financeira, bem como o papel e a importância de um econometrista na vida económica de uma região ou país. Foram também entrevistados os coordenadores de cursos de gestão económica afim se de obter mais informações acerca do processo de ensino da disciplina de econometria, bem como informações sobre o nível e capacidades dos docentes das instituições e ainda pesquisou-se e analisaram-se os programas de todas as disciplinas que de certa forma influenciam no ensino dos modelos econométricos como as Matemáticas e a Estatística, para um estudo comparativo em termos do que é lecionado nestas disciplinas e que influencia diretamente o ensino dos modelos econométricos. Por último, a autora formulou algumas conclusões resultantes da pesquisa feita.

2. Revisão da bibliografia

De acordo com a sociedade portuguesa de estatística (2009): a Econometria é um ramo relativamente novo da economia, foi fundada em 1930 e nas últimas décadas, tem evoluído consideravelmente. Vários fatores tem contribuído para esta evolução. Entre eles, temos como principal fator os avanços tecnológicos, os avanços da informática que permitiram a existência de bases de dados impensáveis nos anos 30, bem como os meios de cálculo necessários ao seu tratamento adequado e rápido.

As características específicas dos dados económicos levaram a Econometria a desenvolver várias técnicas estatísticas novas, contribuindo assim para a sua crescente automatização em relação a outras áreas da ciência em que o recurso à estatística matemática é igualmente intenso, com o rápido desenvolvimento técnico da segunda metade do século XX.

A Econometria é um recurso indispensável na acumulação do saber e na validação das hipóteses avançadas.

Importante como é para a formação de um economista, no ensino da econometria requer-se a conjugação de conhecimentos nas fronteiras da matemática, da estatística e da economia; requer disponibilidade de base de dados com dados quantitativos extensos e coerentes; e meios informáticos de alguns vultos.

Apesar da imensidão do domínio contemporâneo da Econometria, tudo começou numa ideia simples, fruto de trabalhos precursores que, ao longo do século XIX foram burilando os alicerces daquilo que conhecemos como modelo clássico de regressão linear.

No estudo da Econometria como disciplina tem-se sempre o modelo de regressão linear como ponto de partida, apresentando os fundamentos teóricos subjacentes, enunciando as hipóteses que o formam. Estudam-se os modelos de estimação apropriados e a análise estatística dos resultados, analisam-se também os casos de violação das hipóteses como as questões de heteroscedasticidades comum e autocorrelação. A compreensão deste modelo base leva ao entendimento de várias aplicações empíricas, facilita a apreensão das extensões e o desenvolvimento do modelo de forma geral.

Contudo podemos dizer que a **Econometria** é um conjunto de ferramentas estatísticas com o objectivo de entender a relação entre variáveis económicas através da aplicação de um modelo matemático.

Nos primeiros anos, a maioria das aplicações lidava com questões macroeconómicas para ajudar governos e grandes empresas a tomar suas decisões de longo prazo. Atualmente a econometria é uma ferramenta indispensável para modelar a realidade em quase todas as áreas económicas e de negócios.

De acordo com o boletim SPE (Sociedade Portuguesa de Estatística, 2009), a investigação em econometria financeira realiza-se em duas grandes áreas: uma que se preocupa fundamentalmente com o desenvolvimento de métodos econométricos adequados a dados financeiros; a outra, mais empírica, que aplica os métodos para testar hipóteses e teorias da economia financeira. Alguns dos temas genéricos de interesse na área do desenvolvimento dos métodos econométricos são a estimação, a construção de modelos econométricos e a previsão. De todo o modo, a área proeminente em econometria financeira é a das séries temporais. Onde se analisa, por exemplo, a evolução temporal das cotações, das taxas de câmbio, das taxas de juro, etc. É conveniente distinguir séries temporais de natureza macroeconómica e as de natureza financeira. Dados de natureza macroeconómica (consumo, produto, taxa de desemprego) podem ser observados com periodicidade mensal, trimestral ou anual; dados financeiros, como por exemplo, retornos de ações ou taxas de câmbio, podem ser observados com uma frequência muito superior; nalguns casos, com intervalos de minutos ou segundos entre duas observações consecutivas. Assim, o número de observações disponíveis de dados financeiros pode situar-se na ordem das centenas de milhares, ou ainda mais. Normalmente, prefere-se trabalhar com dados diários (evitando-se os problemas de microestrutura de mercado). Com as séries macroeconómicas raramente se passam das poucas centenas de observações (quando, na melhor das hipóteses, se têm observações mensais). Os dados macroeconómicos são menos fiáveis, estão mais sujeitos a erros de medição.

2.1 Modelos econométricos

Segundo Mansfield (1995): um modelo econométrico é um sistema de equações ou uma simples equação calculada a partir de dados do passado que é usado para prever variáveis económicas e empresariais. A essência é combinar a teoria económica com modelos estatísticos modernos.

Para Peixoto (2005 USP): a aplicação de métodos matemáticos é o conjunto de dados económicos com o objetivo de prover suporte empírico às teorias económicas, constituindo o ponto fundamental da econometria. Pode-se definir a econometria como um método de análise económico que agrega a estatística, a matemática e a teoria económica.

Numa abordagem clássica, um modelo é uma representação simplificada de um processo do mundo real. Nos anos 50 e 60 a preocupação dos econometristas centrava-se na construção de modelos econométricos completos que contivessem implicações econométricas testáveis, ou seja, os econometristas desta época preocupavam-se com a formulação de modelos económicos cuja representação econométrica produzisse hipóteses testáveis relevantes à ciência económica, no que diz respeito aos resultados na previsão e na avaliação dos resultados de políticas económicas.

Neste contexto dizemos que um modelo econométrico é um conjunto de equações comportamentais derivadas do modelo económico, envolvendo variáveis observáveis e um termo aleatório ou errático, com todos os factores que não foram incorporados ao modelo em análise (com afirmação sobre a existência de erros de observação em variáveis do modelo e sobre a especificação distribuição de probabilidades do termo aleatório, com o intuito de prover formas representativas passíveis de teste empírico, por meio de estimação, teste e chocagem do diagnóstico produzido). Em suma, a metodologia econométrica aqui analisada como modelo de análise tem como ponto de partida a teoria económica, permitindo ao pesquisador a construção de um modelo teórico que engloba as principais conclusões teóricas.

Ao caracterizarmos a distribuição de probabilidades do termo aleatório produzimos um modelo econométrico pronto a ser estimado e testado envolvendo métodos estatísticos de estimação de um modelo econométrico, cujo suporte é o conjunto de dados empíricos observáveis, bem como a

inferência estatística. Enfim, o modelo econométrico analisado caso seja aprovado é usado na previsão do comportamento de variáveis económicas ou na elaboração de cenários, aquando da formulação de políticas económicas.

Já a metodologia econométrica da escola inglesa apresentada por Hendry (1987) e Spanos (1989), resulta de uma crítica feita à modelagem clássica. Dentre as várias críticas destacamos aqui, a diferença entre o modelo econométrico proposto e o processo gerador dos dados produzidos pela realidade, ou seja, a associação entre uma variável teórica e a sua representação estatística, não é nada mais do que uma forma de ajustamento do modelo aos dados. Pode ir além do previsto por um arcabouço qualquer, logo a posição clássica que tem como ponto de partida o enfoque teórico para a modelagem econométrica sofre sérias críticas, porque se deve deixar espaço para informações adicionais contidas nos dados e que permitem um espaço científico mais amplo ao pesquisador. O que não significa que as técnicas usadas pela econometria tradicional estejam ultrapassadas. A diferença mais significativa está na maneira como o modelo de regressão linear é inicialmente formulado, se de acordo a abordagem tradicional ou pela via da análise estatística (mais genérico possível) para daí, extraídos os excessos, chegarmos a um modelo representativo da realidade.

Daí não poder dizer que, um modelo económico é modelo econométrico, a não ser que ao modelo económico se lhe acrescente um termo aleatório, qualificando assim o modelo estatístico numa base de dados experimental.

Segundo Peixoto (2005), a identificação entre dados experimentais e os gerados pela natureza não são evidentes por não haver correspondência entre um modelo estatístico e os dados observáveis escolhidos para representar as variáveis teóricas neles contidas.

A teoria económica, como arcabouço lógico de ideias, não pode estar totalmente deslocada do processo gerador que caracteriza o comportamento dos dados na natureza. O que nos leva à construção de um modelo económico estimável, que não admite a simplicidade como ponto de partida, mas sim a generalização.

Portanto existe inter-relação clara entre os vários estágios decisivos do método ao mesmo tempo em que se manifesta uma preocupação em seguir um processo analítico que parte da generalização e chega à simplificação. A modelagem econométrica inglesa deve ser construída com base em um modelo estatístico cuja especificação seja a mais generalizada possível, e ter um mecanismo que parametrize as informações contidas no conjunto de dados, tendo em vista os modelos estimáveis.

Nas ciências sociais a palavra modelo é usada para designar uma representação abstrata e simplificada de certo aspecto da realidade. Em econometria modelos são uma representação simbólica corporizada num conjunto de relações matemáticas estabelecidas entre grandezas quantificáveis, descrevendo hipóteses formuladas pelas teorias económicas como a forma típica de uma regressão linear:

$$y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k \quad (I)$$

Onde y é a variável dependente, $\beta_1 \dots \beta_k$ coeficientes de regressão e $x_2 \dots x_k$ variáveis independentes.

Para explicar porque em momentos diferentes as variáveis assumem valores diferentes numa certa população, recorreremos aos chamados modelos exactos ou determinísticos e aos modelos estocásticos ou aleatórios.

2.1.1 Modelos determinísticos

Um dos modelos muito usado no estudo dos modelos determinísticos é o modelo explicativo mais geral das funções algébricas.

Nem sempre o modelo determinístico é viável, para as medições económicas aplicadas à econometria, por não explicar corretamente a diferença existente nos resultados, recorrendo-se a outras formas de caracterizar as variáveis em questão como, por exemplo, no grupo de esferas da mesma espécie, a possibilidade de se retirar a mesma esfera várias vezes é ínfima, e só é calculada esta possibilidade de se retirar a mesma esfera tantas vezes quantas possíveis se recorrer aos indicadores como a média, a variância ou a distribuição de probabilidades, elementos pertencentes à estatística. Assim no lançamento de um dado não viciado a possibilidade de sair de forma

sucessiva cada uma das faces é de $\frac{1}{6}$. Logo este modelo explicativo não determina por cada variável o valor assumido apenas caracteriza a variável principal no momento da sua distribuição.

Logo os modelos exactos usados em física e química, pouco se aplicam em economia, por não dar a possibilidade de existência de uma variável aleatória, que influencie os resultados nas variáveis explicativas, somente a variável aleatória permite explicar as variações que ocorrem na variável em estudo, chamada de erro ou perturbação. O modelo determinístico afirma taxativamente a um dado nível esta associação.

Em economia a inclusão do termo de perturbação aleatória é justificada por várias razões. N um modelo explicativo para a procura de um bem, vê-se que, a teoria micro económica enfatiza o preço de um bem no rendimento dos consumidores, sem excluir outras determinantes da procura. O que nos mostra que existem determinantes que são relevantes para o caso de um bem ou outro, não generalizáveis.

A economia é uma ciência social. Não havendo outra razão para esperar a ausência de previsibilidade exata na determinação dos fenómenos económicos, bastaria a aleatoriedade intrínseca dos comportamentos humanos que plasmam a economia para justificar que nos modelos econométricos, seja deixado algum lugar para o acaso.

2.1.2 Modelo Estocástico

Para Andrade nos seus apontamentos sobre econometria aplicada (2001-2004), se pretendemos que uma proposição geral tenha carácter científico, ela deve ser posta à prova da experiência. Qualquer experiência é sempre singular e não podemos classificar uma experiência como “a experiência geral”. A econometria lida com numerosos resultados de experiências. E como, é certo que nenhum conhecimento científico deve ser tomado como absolutamente certo, a natureza estocástica dos resultados econométricos parece-nos adequada à análise científica em economia.

Apelando a teoria económica no que diz respeito à teoria do consumidor, sabe-se que o consumo de uma família num determinado intervalo de tempo,

um ano, por exemplo, depende do rendimento dessa família nesse mesmo período e depende ainda de outras variáveis, como o rendimento de períodos anteriores, as riquezas, o consumo efetuado em períodos anteriores, os gostos dos indivíduos, a influência da moda e da publicidade.

A teoria económica já não fornece muitas informações sobre outros aspectos relevantes para o estudo como a função que relaciona o consumo e o rendimento: que tipo de função é? É linear ou exponencial inversa? Da propensão marginal ao consumo sabe-se que se trata de um parâmetro entre 0 e 1, a teoria económica dá indicações da ordem de grandeza dos parâmetros que caracterizam uma função, mas não se preocupa com a quantificação destes parâmetros. Uma das questões centrais em econometria é a quantificação dos parâmetros que relacionam as variáveis económicas entre si.

Para já uma das funções matemáticas mais simples a considerar é a função linear. $C=a+bR$, onde b é a propensão marginal ao consumo ($0 < b < 1$), a é o consumo autónomo (consumo de uma família num dado período em que o rendimento fosse nulo; havendo um consumo mínimo chamado consumo de subsistência " $a > 0$ " conseguido à custa de poupanças anteriores ou endividamentos).

Os modelos normalmente utilizados em economia não são os determinísticos, mas os modelos estocásticos, dados pela fórmula:

$$C = \beta_1 + \beta_2 R + u.$$

Onde u representa uma variável aleatória (estocástica) que tem em consideração factores puramente aleatórios que influenciam o consumo. Com duas componentes explicativas do consumo: uma parte determinística $\beta_1 + \beta_2 R$, que traduz a influência do rendimento no consumo igual para todas as famílias com o mesmo nível de rendimento, a segunda componente $\beta_2 R$ variável de família para família cujo valor concreto dependerá do caso.

2.2 Fases de um estudo econométrico

Para se construir um modelo devemos primeiro saber que modelo se adequa à situação que se pretende estudar, que etapas ou fases se devem verificar neste estudo. Um estudo econométrico passa por determinadas fases que possibilitam uma melhor compreensão dos modelos econométricos:

1ª- **Fase de especificação do modelo:** os modelos básicos da economia não são geralmente apropriados para fins de quantificação, por limitarem-se a enunciar as principais determinantes de uma variável, explicando a dependência de y em relação à x .

$$y = f(x_1, x_2, \dots, x_k) \quad (II)$$

Por vezes, indicando o sentido do efeito *ceteris paribus* de uma dessas determinantes: $\frac{\partial y}{\partial x} < 0$, uma variação do preço de um bem induz uma variação de sentido contrário da procura de que é alvo. Sem mais informação, quanto à grandeza de um efeito, como ocorre na função – consumo, a respeito da propensão marginal ao consumo, $0 < \frac{\partial C}{\partial R} < 1$.

Raramente é atingido o ponto de precisar o efeito esperado de uma variação simultânea de várias das variáveis explicativas; um exemplo com alcance mais restrito, é o que se verifica numa função de produção de Cobb-Douglas, quando as quantidades de todos os fatores utilizadas por período variam por uma mesma proporção λ : se a tecnologia exibir rendimentos de escala constantes, o produto variará também pela proporção λ .

Quando se quer analisar a quantificação de relações económicas, a habitual parcimónia da teoria de Cobb-Douglas não é suficiente. Apenas ajuda na seleção das variáveis explicativas da variável que se pretende estudar, sendo também de pouco auxílio numa segunda tarefa esta fase: a de definição da forma funcional da relação entre variável explicada e variáveis explicativas. Não basta afirmar que Y “depende” de X , já que é verdade quer seja:

$$Y = \beta_1 + \beta_2 X ; Y = \beta_1 X^{\beta_2} \text{ ou } Y = \beta_1 e^{\beta_2 X} . \quad (III)$$

Para exemplificar (β_1 e β_2 duas constantes) apenas algumas de uma infinita variedade de formas que a dependência pode assumir. A especificação do modelo exige a opção por uma forma funcional concreta.

Quanto a isto é mais fácil encontrar na teoria prescrições de formas inapropriadas do que a prescrição da forma correta. Por exemplo, numa função de produção espera-se que corresponda um *output* de zero ao emprego de *inputs* em quantidades nulas; nas ilustrações acima, as implicações seriam as de obrigar a que β_1 fosse zero, em $Y = \beta_1 + \beta_2 X$, e a de afastar por completo a terceira função: $Y = \beta_1 e^{\beta_2 X}$. Tratando-se de uma função de custo total de curto prazo, prevendo a existência de custos fixos positivos ainda que o volume de produção por período se reduzisse a zero, seria a segunda das funções exemplificada na equação (III) que se tornaria inaceitável.

Perante o laconismo da teoria, não é surpreendente a frequência de recurso à forma funcional mais simples, a que supõe uma relação linear entre variável dependente e variáveis independentes. A abordagem colhe um suporte matemático no desenvolvimento da série de Taylor, que estabelece a possibilidade de aproximar por uma função linear uma função de maior complexidade. E, por isso o <<Y depende de X_2, X_3, \dots, X_k >> da teoria económica frequentemente se traduz na equação (I). Tantas vezes citada nos compêndios de econometria.

Selecionadas variável dependente, variáveis independentes e forma funcional da relação entre uma e outras, a tarefa seguinte na especificação de um modelo é a definição das hipóteses quanto ao processo de geração dos dados observados. A característica essencial do modelo econométrico é a de admitir que a observação da relação seja perturbada por erros aleatórios; que corresponde à variação de um termo, u , à equação $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k$. Sob que condições se admitem exercida a influência desses fatores aleatórios?

Do modelo de regressão linear:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \dots + \beta_k X_{ki} + u_i \quad (\text{IV})$$

Que define a existência de uma relação entre as variáveis Y , X_2, X_3, \dots, X_k e u válida para qualquer indivíduo i pertencente a uma determinada população, caracterizada pelos k parâmetros β desconhecidos; um dos objetivos dos métodos econométrico é precisamente o de estimar estes coeficientes, através de um método estatístico apropriado, tendo em conta a presença no modelo do termo de perturbação aleatória. Que apresenta como um método de estimação dos coeficientes de regressão do modelo, os estimadores de mínimos quadrados dos coeficientes que a partir de uma dada amostra, permitem deduzir a estimativa para os parâmetros. Quais serão as propriedades estatísticas desses estimadores? Serão estimadores cêntricos consistentes? As propriedades dos estimadores vão depender das características da variável Y (depende da variável aleatória u , que é uma das suas componentes) e das variáveis explicativas X_2, X_3, \dots, X_k . Quanto às propriedades estatísticas dos estimadores, é necessário conhecer a distribuição de probabilidade da variável u , ou pelo menos conhecer os principais momentos desta distribuição. Não sendo conhecida essa distribuição, nem a variável u observável, terão de estabelecer-se determinadas hipóteses relativas à sua distribuição de probabilidades. Essas hipóteses e outras relativas às variáveis explicativas são conhecidas como as hipóteses clássicas do modelo de regressão linear.

Esta fase de especificação do modelo é dominada pela teoria e pela especulação.

2ª- Fase da seleção da amostra: debruça-se sobre a realidade. Não se faz econometria sem dados, e a recolha de dados levanta novos problemas. Uma amostra é um conjunto de observações das variáveis de interesse. As amostras podem ser do tipo seccionais ou transversais, em que se observam diversas unidades estatísticas referidas a um mesmo momento ou período, de amostras temporais ou cronológicas em que, para a mesma unidade estatística, se dispõem de dados para diferentes momentos ou períodos. Um terceiro tipo de amostras são as ditas longitudinais ou em painel, onde se conjugam as perspetivas seccional e temporal, nelas são observadas diferentes unidades estatísticas, ao longo de vários períodos ou momentos. As unidades estatísticas poderão ser países, regiões, empresas, etc;

variáveis que reflitam um *stock* referem-se a um dado momento; variáveis representativas de fluxos podem ter periodicidade anual, trimestral, mensal, semanal ou diária.

Conforme o modelo terá de optar por algum destes tipos de amostras, mas há muitos temas susceptíveis de análise segundo uma ou outra perspectiva. Assim a função-consumo pode ser estudada com base numa amostra seccional de família, ou considerando o consumo agregado numa economia, com base numa amostra de observações anuais. Conceitos como rendimento, preço ou juro são entidades abstratas. Numa economia concreta, tem-se informação quanto ao rendimento nacional, rendimento disponível das famílias, receitas da administração pública, entre outros tipos de rendimentos. Há uma enorme variedade de preços variando de país para país ou, de loja para loja, preços no produtor e preços no consumidor, preços expressos em diferentes unidades monetárias ou referidos a momentos diversos.

Em economia, é frequente dispor-se de informação multifacetada sobre uma variável: os preços constantes ou preços variáveis, em termos reais ou nominais, corrigidos de oscilações sazonais ou não, com valores totais ou em capitações, expressa em nível ou por meio de taxas de variação. A teoria nada diz acerca disto, para ajudar na recolha de dados nada mais resta que o bom senso. E há várias áreas onde não se consegue informações estatísticas fiáveis.

Especificado o modelo e coligidos os dados estatísticos relevantes, segue-se a fase de estimação.

3ª- Fase de estimação ou ajustamento do modelo: há um conjunto de constantes desconhecidas, que são os parâmetros do modelo, que se supõe governarem o processo que gerou os dados observados. O econometrista tem o dever de produzir estimativas, isto é, valores numéricos concretos para constantes que aproximem plausivelmente os seus valores ignorados. Recorre-se a inúmeros métodos para este fim. Aqui interessam métodos que possuam propriedades desejáveis, e como se defenderá o método dos mínimos quadrados sobleva as alternativas em muitas situações. Importa ter uma medida de precisão das estimativas que se observe e, formar uma ideia de credibilidade dos resultados em que se chegou.

A questão da estimação dos parâmetros do modelo de regressão linear clássico relaciona uma variável Y explicada com um conjunto de variáveis explicativas e uma variável aleatória u. Admite-se que, numa certa população, existe uma relação entre essas variáveis descritas pela equação:

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + u. \quad (V)$$

Em que $\beta_1, \beta_2, \beta_3, \dots, \beta_k$ são constantes e u é uma perturbação aleatória. Aqui é suficiente ver na inclusão da variável u, efeitos aditivos e natureza aleatória, uma forma prática e conveniente de acomodar no modelo uma multiplicidade de outros factores com influência sobre a variável explicada e não adequadamente refletidos pelas variáveis explicativas.

4ª- Fase de avaliação dos resultados: um indicador sumário e elementar da qualidade do ajustamento é o coeficiente de determinação, que nos possibilita estimar o valor de Y para cada uma das observações de acordo a equação:

$$\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \dots + \hat{\beta}_k X_{ki} \quad (VI)$$

Em que $i=1,2,\dots,n$, ou, a notação matricial;

$$\hat{Y} = X \hat{\beta} \quad (VII)$$

Geralmente esses valores ajustados para a variável explicada coincidem com os observados. O grau de fidelidade com que os primeiros reproduzem os segundos pode ser objeto de uma primeira quantificação e avaliação, recorrendo a um indicador que se designa por coeficiente de determinação, cujo símbolo é R^2 . Assim temos:

$$\hat{Y}_i = \hat{Y}_i + \ell_i \quad (VIII)$$

Esta igualdade traduz a diferença entre o valor estimado de Y e o seu valor médio e o resíduo de estimação. Dispõe-se em econometria de uma bateria extensa de testes estatísticos que permitem sujeitar os resultados obtidos na estimação a um escrutínio mais rigoroso;

5ª- Fase utilização dos resultados: o interesse primordial de um modelo econométrico é o de averiguar como ocorre a confirmação dos dados, sua refutação ou modulação da teoria e a quantificação dos impactos de diferentes factores sobre a variável em análise. Muitos dos avanços da economia no século XX tiveram na gênese a constatação, obtida em estudos

econométricos, de que comportamentos preditos na teoria não eram plenamente corroborados pela realidade e as limitações evidenciadas forneceram frequentemente pistas sobre os aspectos que havia que melhorar ou aprofundar.

Uma utilização mais pragmática de um estudo econométrico é o que ocorre na previsão. Especificando o modelo, estimado, validado pela análise estatística, pode servir de base a antecipações do que acontecerá em situações futuras ou diferentes, desde que seja razoável admitir que as características fundamentais do processo que gerou os dados observados permaneçam inalteradas.

Modelos econométricos constituem um instrumento indisponível de avaliação, se forem utilizados como meios de prever os impactos esperados de políticas, programas ou ações específicas.

Existem alguns autores que consideram que as fases de construção de um modelo econométrico resumem-se em apenas três fases como descreve, Mahia (2004). Para este, na construção de um modelo econométrico cumprem-se as seguintes etapas: 1ª- Especificação do modelo; que é a etapa de identificação do modelo, é a mais comprometedora, visto que, escolhido o modelo inadequado teremos maus resultados, ou não teremos resultado nenhum. Nesta fase faz-se o exame do marco histórico no contexto de sua aplicação, seleciona-se uma amostra para análise bem como as variáveis mais relevantes, sua medição e transformação, incluindo a seleção da forma funcional. Uma vez identificado o modelo correto a ser aplicado, vamos para a segunda etapa; 2ª-Estimação do modelo; aqui faz-se uso do modelo já identificado na primeira etapa pela estimação dos dados das amostras, analisando os dados aplicados nas equações e verificando os resultados. Escolhe-se o modelo de estimação a usar onde se aplicará a análise das propriedades e probabilidades, obtenção da estimação de parâmetros e a variância de perturbação aleatória. Vemos que a construção de um modelo não termina com a fase de identificação e estimação dos parâmetros, o resultado da estimação inicial é apenas um ponto de partida para que haja um modelo final que deverá ser contestado e validado; 3ª- Contestação e validação do modelo, nesta fase o processo de contestação e validação deve ser feito de forma ordenada porque geralmente não consiste

num processo linear, mais sim volta atrás, planeando-se uma revisão de especificação e estimação. Analisa-se a situação e quantificam-se os dados obtidos, analisa-se a veracidade dos parâmetros individuais e colectivos, contestam-se as hipóteses básicas de estrutura e relativas às perturbações aleatórias.

Para que um dado modelo económico seja considerado como modelo econométrico e aplicável a um determinado caso é necessário que estas fases citadas acima estejam muito bem identificadas de modos a evitar uma obtenção de resultados errados e não aplicáveis à situação pesquisada.

2.3 Como usar modelos econométricos

Até às décadas de 70 e 80, muito do que se fazia e estudava em matéria de macro economia e política económica baseava-se fortemente na estimação de modelos econométricos de larga escala, mais especificamente a metodologia de equações simultâneas. Uma vez que as previsões dessas análises começaram a apresentar falhas, críticas ao método foram surgindo, em especial em Sims (1980). Neste artigo seminal para a macro econometria, o autor introduz um tipo de modelagem que mais tarde se consagraria no modelo de Vector Auto-Regressivo (VAR). Vários estudos clássicos sobre política monetária utilizam este modelo, por exemplo, Bernanke e Blinder (1992), Bernanke e Mihov (1998), Blanchard e Quah (1989), entre outros. Nesse contexto, tal metodologia apresenta a vantagem de não ser necessário determinar *à priori* as variáveis endógenas e exógenas no modelo, prática que segundo a crítica de Sims estaria escandalizando os modelos de larga escala, uma vez que para sua estimação seria preciso impor "restrições incríveis para identificação". No VAR, todas as variáveis são consideradas endógenas e permite-se que haja relações entre todas elas e suas defasagens.

Segundo Mansfield (1995): Administradores e analistas em anos recentes têm baseado cada vez mais as suas previsões em técnicas de regressão múltipla e modelos de equações múltiplas, como por exemplo, para estimar a quantidade de automóveis produzidos por uma empresa americana num trimestre.

De acordo com um estudo feito por Harris (1985) publicado pela Federal Reserve Bank de Nova Iorque a seguinte equação de regressão é útil para este caso;

$$A = -22\,302 + 12,9D - 97,8I - 19,9R + 230P + 6,0N,$$

Em que A é a quantidade de automóveis produzidos trimestralmente, D é o rendimento disponível real, I é a taxa de juro bonificada, R é o rácio stock/vendas, P o preço dos automóveis, e N o nível de preços para os outros bens.

Calculam-se os valores das variáveis independentes e substituem-se na equação.

O modelo de Wharton (segundo Edwin, 1996) pioneiro neste campo, continha centenas de equações destinadas a explicar o nível de gastos das famílias, o nível de investimento económico, o rendimento global e o emprego, salários, preços e taxas de juros. As previsões do modelo de Wharton foram seguidas cautelosamente por grandes empresas e agências governamentais. Algumas empresas construíram a partir daí modelos próprios de equações múltiplas, como a General Electric. O que não significa que estes grandes modelos não falhem. No entanto, estes modelos continuam a ser usados pelos governos e empresas.

Tanto o modelo de Wharton com as suas centenas de equações e o da equação simples usado na previsão da quantidade de automóveis, são exemplos de modelos econométricos.

Como podemos usar os modelos econométricos?

Os modelos econométricos são usados para análise de estrutura, simulações, previsões e na constatação da teoria económica.

2.4 Econometria e Economia Matemática

O termo *econometria* é às vezes confundido com economia matemática. O termo "metria" da última parte da palavra econometria está relacionado com a medição de dados económicos, abordando estudos de observações *empíricas* através de métodos estatísticos de estimação e testes de hipóteses. O ramo da economia matemática, por sua vez, se destina à aplicação da matemática a aspectos puramente teóricos da análise

económica, preocupando-se muito pouco, ou quase nada, com problemas estatísticos como erros de medição das variáveis que estão sendo investigadas.

A aplicação da economia matemática se concentra no raciocínio dedutivo e não no indutivo, lidando, portanto, com problemas teóricos e não empíricos o que, por sua vez, não torna a econometria menos relevante.

Uma vez que os estudos empíricos e análises teóricas são complementares e se reforçam mutuamente, a validade das teorias é testada em relação aos dados empíricos antes que elas possam ser aplicadas com confiança. Em contrapartida, o trabalho estatístico necessita da teoria económica como guia, para determinar a direção mais relevante e proveitosa da pesquisa.

Todavia, de certa forma, a economia matemática pode ser considerada a mais básica das duas, pois para desenvolver um estudo estatístico e econométrico significativos, é indispensável uma boa estrutura teórica - de preferência em formulação matemática.

2.5 Problemas de uma análise de regressão

No estudo dos modelos econométricos podemos encontrar vários problemas que podem surgir em uma análise de regressão, comprometendo a confiabilidade dos valores dos coeficientes estimados, assim como a inferência estatística. Entre eles os mais comuns são a multicolinearidade, a heteroscedasticidade e a autocorrelação.

2.5.1 Multicolinearidade

Dos compêndios de Econometria vemos que a multicolinearidade não passa de uma colinearidade entre mais de duas variáveis, não envolvendo necessariamente colinearidade entre as variáveis duas a duas. A existência de uma correlação linear muito forte entre as variáveis explicativas pode levar a uma situação problemática, nos resultados obtidos, ligados à instabilidade e falta de robustez desses resultados, chamada de colinearidade imperfeita, quasi-colinearidade, ou ainda de multicolinearidade. Frequentemente nos casos de multicolinearidade surgem situações de grande imprevisão das estimativas obtidas para os coeficientes do modelo, de problemas na implementação de testes estatísticos ou de obtenção de intervalos de

confiança para os coeficientes de regressão. No entanto a detenção da multicolinearidade não é fácil, por ser um problema amostral. A detenção da multicolinearidade está na existência de indícios que, quantos presentes levarão o econometrista a suspeitar de uma situação de multicolinearidade e, eventualmente a adopção de alguma medida corretora ou, pelo menos atenuadora do problema. Entre os indícios de eventual existência de multicolinearidade tem-se:

1º Os sinais das estimativas dos coeficientes diferentes dos que eram esperados, aqui vemos que a variância do estimador de qualquer coeficiente de regressão tende a ser muito grande, podendo levar a que a estimativa do coeficiente apresente um sinal contrário ao que seria esperado;

2º Regressão globalmente significativa, mas poucas, ou nenhuma, variáveis individualmente significativas, aqui se detetam a influência conjunta das variáveis explicativas, mas sem que seja possível, simultaneamente, detectar a sua significância individual;

3º Alteração na significância individual das variáveis explicativas com a introdução de uma nova variável, aqui a significância global da regressão praticamente mantém-se, mas as variáveis que eram estatisticamente significativas, muitas vezes para um nível de significância muito pequeno, passam a não o ser;

4º Grande sensibilidade das estimativas obtidas a pequenas modificações na amostra, aqui vemos que acrescentar ou retirar, uma ou duas observações é suficiente, em muitos casos, para que as estimativas obtidas para os coeficientes de regressão se alterem significativamente, por vezes mudando mesmo de sinal;

5º Elevados coeficientes de correlação amostrais entre as variáveis explicativas, aqui encontramos duas limitações, em que a primeira é a dificuldade em estabelecer um limite para o valor a partir do qual se considera que há problemas de quasi-colinearidade, e a segunda, de que aqueles coeficientes medem o grau de associação linear entre duas variáveis, não contemplando os casos de forte relação linear entre mais de duas variáveis;

6º Elevados valores dos coeficientes (R_f^2), aqui os coeficientes de determinação destas regressões elevados indicariam a existência de quasi-

colinearidade, embora mais uma vez se coloque a questão do limite para o valor que indicaria o problema;

7º Determinante da matriz $(X'X)$ próximo de zero, aqui o determinante tenderá a ser próximo de zero, pelo que um valor reduzido para aquele determinante será um possível indicador de problemas de multicolinearidade. No entanto sendo a multicolinearidade um problema amostral, terá de ser resolvido a nível amostral.

Embora existam outros processos na literatura para lidar com esta questão, os passos acima citados têm dado significativa ajuda na compreensão e resolução dos problemas de multicolinearidade envolvendo a análise estatística sob a hipótese da normalidade, respeitando-se todas as hipóteses.

2.5.2 Heterocedasticidade

Nas aplicações económicas a heterocedasticidade é a regra e a homocedasticidade a exceção. Uma hipótese importante do modelo clássico de regressão linear é que os termos de erro que aparecem na função de regressão linear são homoscedásticos. $E(u_i) = 0$ com $i = 1, 2, \dots, n$. A homocedasticidade do modelo de regressão com duas variáveis pode ser reduzida, a variância condicional de Y_i , condicionada ao X_i dado, permanece a mesma, independentemente dos valores de X . Do contrário a variância condicional de Y_i , aumenta à medida que X aumenta. Neste caso, as variâncias de Y_i não são as mesmas. Então há heterocedasticidade que é um problema muito frequente em amostras transversais. $E(u_i^2) = \sigma_i^2$. Um exemplo clássico na matéria é o da função-consumo numa população de agregados familiares, em que a margem de consumo é muito maior quando o rendimento é elevado do que quando mal chega para assegurar as necessidades básicas de sobrevivência. É possível que não haja grande disparidade de comportamento entre os dois grupos de famílias no que toca à satisfação das necessidades básicas. Repare-se que o que está em causa não é a média dos gastos de consumo por grupos familiares, mas sim a variação de família para família em torno dessas médias. Colocam-se dificuldades análogas quando coexistem, numa amostra, pequenas e médias empresas a par de empresas de grande dimensão, ou pequenos países, em

termos geográficos ou económicos, a par de potências à escala regional ou mundial.

Segundo Gujarati e Porter (2011), para se diferenciar a homocedasticidade da heterocedasticidade deve-se ter em conta os seguintes aspectos:

1-Segundo os modelos de erro-aprendizagem, comportamentos incorretos de pessoas diminuem com o tempo ou o número de erros torna-se mais consistente, neste caso espera-se que σ_i^2 diminua.

2- À medida que a renda aumenta, as pessoas têm mais renda discricionária e, mais opções para escolher como aplicarão sua renda. Com tudo é provável que σ_i^2 aumente com a renda.

3- À medida que as técnicas de recolha de dados aprimoram-se é provável que σ_i^2 diminua.

4- A heterocedasticidade pode ser resultado da presença de dados discrepantes, aquela que difere muito em relação às observações da amostra. Um dado discrepante é a observação de uma população que difere daquela que gera as observações amostradas remanescentes. A inclusão ou exclusão de tal observação pode alterar substancialmente os resultados da análise de regressão.

5- A heterocedasticidade também pode surgir da violação do modelo clássico de regressão linear, quando este deve ser especificado corretamente. Muitas vezes algumas variáveis importantes são omitidas do modelo e isso nos dá a impressão de tratar-se de heterocedasticidade.

6- A assimetria é também uma fonte de heterocedasticidade na distribuição de um ou mais regressores incluídos no modelo.

7- A heterocedasticidade surge da transformação incorreta de dados (transformações proporcionais ou de primeira diferença) e da forma funcional incorreta (modelos lineares *versus* log-lineares).

O que acontece com os mínimos quadrados ordinários (MQO) e suas variâncias se introduzirmos a heterocedasticidade $E(u_i^2) = \sigma_i^2$ e mantivermos todas as demais hipóteses do modelo clássico? Partindo do modelo de duas variáveis:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \quad (\text{IX})$$

Aplicando a fórmula usual, o estimador de MQO de β_2 é:

$$\beta_2 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2} \quad (X)$$

Mas a variância será dada pela expressão:

$$Var(\beta_2) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2} \quad (XI)$$

Que é diferente da fórmula usual de variância obtida supondo-se a homocedasticidade.

$$Var(\beta_2) = \frac{\sigma^2}{\sum x_i^2} \quad (XII)$$

Se $\sigma_i^2 = \sigma^2$ para cada i , então as duas fórmulas são iguais. Porque? β_2 será o melhor estimador linear não tendencioso se as hipóteses do modelo clássico, que incluem a homocedasticidade, forem válidas. Se excluirmos a hipótese de homocedasticidade e substituirmos pela de heterocedasticidade, ainda assim β_2 conservar-se-á linear e não tendencioso. Para estabelecer a não tendenciosidade de β_2 , é necessário que este seja um estimador consistente sob as hipóteses do modelo clássico de regressão linear. Apesar da heterocedasticidade, pode-se indicar que β_2 é um estimador consistente, ou seja, quando o tamanho da amostra aumenta indefinidamente, o β_2 estimado converge para o seu verdadeiro valor. Sob certas condições de regularidade, β_2 é assintoticamente normalmente distribuído. O que também é válido para outros parâmetros de modelo de regressão múltipla. Se β_2 for ainda linear, não tendencioso e consistente, ele deixa de ser o melhor e a variância mínima não é dada pela equação (XI). O método dos MQO não usa as informações contidas na variabilidade desigual da variável dependente Y . Mas um método de estimação, conhecido como mínimos quadrados generalizados (MQG), leva tais informações em consideração explicitamente e é capaz de produzir melhor estimador linear não tendencioso. Do modelo de duas variáveis da equação (IX). Para facilitar o cálculo algébrico temos:

$$Y_i = \beta_1 X_{0i} + \beta_2 X_i + u_i \quad (XIII)$$

Em que $X_{0i}=1$ para cada i . Estas duas equações são idênticas. Se conhecidas as variâncias heterocedásticas σ_i^2 e dividindo a equação anterior por σ_i obtemos:

$$\frac{Y_i}{\sigma_i} = \beta_1 \frac{X_{0i}}{\sigma_i} + \beta_2 \frac{X_i}{\sigma_i} + \frac{u_i}{\sigma_i} \quad (\text{XIV})$$

Simplificando a expressão vem:

$$Y_i^* = \beta_1^* X_{0i}^* + \beta_2^* X_i^* + u_i^* \quad (\text{XV})$$

Em que as variáveis com asterisco são as variáveis do modelo transformadas (variáveis originais divididas por σ_i) β_1^* e β_2^* são os parâmetros do modelo transformados e u_i^* o termo de erro transformado. Daí vem que:

$$Var(u_i^*) = E(u_i^*)^2 = E\left(\frac{u_i}{\sigma_i}\right)^2 = \frac{1}{\sigma_i^2} E(u_i^2) = \frac{1}{\sigma_i^2} (\sigma_i^2) = 1 \quad (\text{XVI})$$

A variância do termo u_i^* agora é homocedástica. Uma vez que ainda conservamos as outras hipóteses do modelo clássico, esta constatação sugere que, aplicando o MQO ao modelo transformado produzirá melhores estimadores lineares não tendenciosos. Os β_1^* e β_2^* estimados agora são melhor estimador linear não tendencioso e não os estimadores de MQO. Este procedimento é conhecido como método de mínimos quadrados generalizados (MQG). Em suma os MQG são os MQO nas variáveis transformadas que satisfazem as hipóteses padrão de mínimos quadrados.

Para estimar β_1^* e β_2^* , escrevemos a função de regressão amostral do modelo transformado da equação (XV). Para obtermos os estimadores dos parâmetros de MQG, minimizamos:

$$\begin{aligned} \sum \hat{u}_i^{2*} &= \sum (Y_i - \beta_1^* X_{0i} - \beta_2^* X_i)^2 \text{ ou} \\ \sum \left(\frac{\hat{u}_i}{\sigma_i}\right)^2 &= \sum \left[\left(\frac{Y_i}{\sigma_i}\right) - \beta_1^* \left(\frac{X_{0i}}{\sigma_i}\right) - \beta_2^* \left(\frac{X_i}{\sigma_i}\right) \right]^2 \end{aligned} \quad (\text{XVII})$$

O estimador de MQG para β_2^* é:

$$\beta_2^* = \frac{(\sum w_i)(\sum w_i X_i Y_i) - (\sum w_i X_i)(\sum w_i Y_i)}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \quad (\text{XVIII})$$

E a sua variância é dada por:

$$Var(\beta_2^*) = \frac{\sum w_i}{(\sum w_i)(\sum w_i X_i^2) - (\sum w_i X_i)^2} \quad (XIX)$$

Nos MQG minimizamos a soma ponderada dos quadrados dos resíduos com $w_i = 1/\sigma_i^2$ funcionando como peso, nos MQO minimizamos uma soma dos quadrados dos resíduos não ponderada ou igualmente ponderada. Nos MQG as observações vindas de uma população com σ_i maior obterão peso proporcional maior na minimização da soma dos quadrados dos resíduos. Pois ao estimarmos a função de regressão populacional de uma forma mais confiável, gostaríamos de dar mais peso às observações agrupadas em torno de sua média do que aquelas que estão bastante dispersas.

O que acontece com nosso intervalo de confiança, testes de hipóteses e outros procedimentos, se continuarmos a usar o estimador de MQO, β_2 ? Para responder a esta questão analisamos dois aspectos que são: 1º Estimação de MQO admitindo-se a heterocedasticidade, aqui se considera explicitamente a heterocedasticidade. Supondo que σ_i^2 é conhecido, é possível mostrar que $var(\beta_2^*) \leq var(\beta_2)$, o que significa que os intervalos de confiança baseados nos últimos serão desnecessariamente maiores. E os testes habituais t e f provavelmente nos darão resultados imprecisos, pois a variância de β_2 é excessivamente grande e o que parece ser um coeficiente estatisticamente insignificante pode ser significativo se os intervalos de confiança forem estabelecidos com base no procedimento MQG;

2º Estimação de MQO desconsiderando a heterocedasticidade, este é o caso mais provável de se encontrar dos dois aqui apresentados. Diante da suspeita de heterocedasticidade, usar o programa padrão de cálculo de regressões de MQO e desprezar a heterocedasticidade fornece à variância de β_2 em primeiro lugar um estimador tendencioso para a variância como mostra a equação (XI). Em média ele sobreestima ou subestima à variância, e, em geral, não se pode afirmar se é positiva ou negativa (se sobreestimação ou subestimação), porque depende da natureza da relação entre σ_i^2 e os valores assumidos pela variável explanatória X . O viés surge do facto de que o estimador convencional σ_i^2 , é um estimador tendencioso quando existe heterocedasticidade. Resulta de que não podemos contar com

os intervalos de confiança calculados de maneira convencional e com os testes t e F normais. Em suma, persistimos no uso dos procedimentos comuns de teste apesar de heterocedasticidade, quaisquer que sejam as conclusões a que chegarmos ou as inferências que fizermos poderão ser equivocados. Fica claro que a heterocedasticidade é potencialmente um problema grave, o analista precisa saber se ela está presente em determinada situação. Se detetada pode-se adotar medidas corretivas, como a regressão de mínimos quadrados ponderados ou alguma outra técnica. Antes de examinarmos os vários procedimentos corretivos, devemos descobrir se a heterocedasticidade está presente ou se provavelmente irá apresentar-se em algum caso.

Como saber se a heterocedasticidade está presente em alguma situação específica? Não há regras específicas para detectar a heterocedasticidade, apenas alguns procedimentos. Em estudos económicos há apenas um valor amostral Y , na maioria dos casos que envolvem investigação econométrica, a heterocedasticidade pode ser uma questão de intuição, inferências baseadas em informações, experiências empíricas anteriores ou pura especulação. Examinemos alguns dos métodos informais e formais para detetar a heterocedasticidade. A maioria desses métodos baseia-se no exame dos resíduos dos MQO, visto que estes é que são observados e não os termos de erro u_i .

Os métodos informais usados na deteção da heterocedasticidade são:

- 1º *Natureza do problema*, a natureza do problema em questão surge da probabilidade de encontrarmos heterocedasticidade. Na verdade, em dados de corte transversal envolvendo unidades heterogeneas, a heterocedasticidade pode ser a regra e não a exceção. Em uma análise de corte transversal que envolve despesas com investimentos em relação às vendas, taxas de juro, em geral espera-se encontrar heterocedasticidade se empresas de tamanho pequena, médias e grandes fizerem parte da amostra;
- 2º *Método gráfico*, não havendo informação empírica sobre a natureza da heterocedasticidade, na prática pode se fazer a análise de regressão supondo-se que não há heterocedasticidade e então fazer um exame *post mortem* dos resíduos elevados ao quadrado \hat{u}_i^2 para ver se existe um padrão

sistemático. Sendo $\hat{u}_i^2 \neq u_i^2$, podem ser substitutos um do outro, principalmente se o tamanho da amostra for suficientemente grande.

Entre os métodos formais, após obter os resíduos, a detecção da heterocedasticidade tem a citar os mais usuais que são:

1º O *teste de Park*, que é um procedimento que envolve duas etapas. Na primeira fazemos a regressão de MQO desconsiderando a questão da heterocedasticidade. Obtemos \hat{u}_i dessa regressão e na segunda etapa fazemos a regressão:

$$\ln \hat{u}_i^2 = \ln \sigma^2 + \beta \ln X_i + v_i = \alpha + \beta \ln X_i + v_i \quad (\text{XX})$$

Em que v_i é o termo de erro estocástico. O *teste de Park* apresenta alguns problemas. Goldfeld e Quandt alegaram que o termo de erro v_i da equação acima, pode não satisfazer as pressuposições dos MQO e ele mesmo pode ser heterocedástico. O método de Park pode ser usado como um método exploratório.

2º *Teste de Glejser*, semelhante ao de Park, após obter os resíduos da regressão dos MQO \hat{u}_i , Glejser sugere a regressão dos valores absolutos de \hat{u}_i , contra a variável X , associada a σ_i^2 . Glejser usa nas suas experiências as fórmulas funcionais:

$$\begin{aligned} |\hat{u}_i| &= \beta_1 + \beta_2 X_i + v_i \\ |\hat{u}_i| &= \beta_1 + \beta_2 \sqrt{X_i} + v_i \\ |\hat{u}_i| &= \beta_1 + \beta_2 \frac{1}{X_i} + v_i \\ |\hat{u}_i| &= \beta_1 + \beta_2 \frac{1}{\sqrt{X_i}} + v_i \\ |\hat{u}_i| &= \sqrt{\beta_1 + \beta_2 X_i} + v_i \\ |\hat{u}_i| &= \sqrt{\beta_1 + \beta_2 X_i^2} + v_i \end{aligned} \quad (\text{XXI})$$

Em que v_i é o termo de erro.

Goldfeld e Quandt ressaltam que o termo de erro v_i tem alguns problemas, uma vez que se espera que seu valor seja diferente de zero, esteja relacionado serialmente e, ironicamente heterocedástico. Uma dificuldade com o método de Glejser é que modelos como $|\hat{u}_i| = \sqrt{\beta_1 + \beta_2 X_i} + v_i$ e

$|u_i| = \sqrt{\beta_1 + \beta_2 X_i^2} + v_i$ não são lineares nos parâmetros e não podem ser estimados com o procedimento usual de MQO. Glejser constatou que, para amostras grandes, os primeiros quatro modelos anteriores em geral apresentam resultados satisfatórios quanto à detecção da heterocedasticidade. A técnica de Glejser pode ser usada para amostras grandes e, nas amostras pequenas, usa-se como um artifício qualitativo.

3º *Teste de correlação por ordem de Spearman*, a que o coeficiente de correlação de ordem precedente pode ser usado para detetar a heterocedasticidade. Supondo que:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \quad (\text{XXII})$$

Temos as seguintes etapas:

- I- Ajuste da regressão aos dados em Y e X, obtendo os resíduos \hat{u}_i ;
- II- Ignorando o sinal de \hat{u}_i , ou seja, tomando o seu valor absoluto $|\hat{u}_i|$, ordena-se $|\hat{u}_i|$ e X_i de acordo a uma ordem ascendente ou descendente e calcula-se o coeficiente de correlação pela ordem apresentada;
- III- Supondo que o coeficiente de correlação por ordem da população, seja zero e $n > 8$ a significância de r_s na amostra pode ser verificada pelo teste t a seguir:

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}}, \text{ com grau de liberdade igual a } n-2. \text{ Se o valor } t \text{ excede}$$

o valor crítico t, pode-se aceitar a hipótese da heterocedasticidade. Se o modelo de regressão envolver mais de uma variável X, r_s poderá ser calculado entre $|\hat{u}_i|$ e cada uma das variáveis X separadamente e poderá ser testado quanto à significância estatística pelo teste t da equação anterior.

4º *teste de Goldfeld-Quandt*, este é aplicável quando se supõe que a variância heterocedástica, σ_i^2 , relaciona-se positivamente com uma das variáveis explanatórias no modelo de regressão. Do modelo usual de duas variáveis e supondo que σ_i^2 relacione-se positivamente X_i como $\sigma_i^2 = \sigma^2 X_i^2$

com σ_i^2 constante. A hipótese postula que σ_i^2 é proporcional ao quadrado da variável X, considerada muito útil por Prais e Houthakker, no estudo sobre orçamentos de famílias. Se adequada a equação anterior, significa que σ_i^2 , será maior quanto maiores forem os valores de X_i , assim a heterocedasticidade estará presente no modelo. Para teste sugerem-se as seguintes etapas:

- I- Ordene ou classifique as observações de acordo com os valores de X_i , a começar pelo valor mais baixo de X;
- II- Omita c observações centrais, em que c é especificado, e divida as observações remanescentes em dois grupos com observações (n-c)/2 em cada um;
- III- Ajuste as regressões MQO separadas, para as primeiras observações (n-c)/2 e para as últimas (n-c)/2, e obtenha as respectivas somas dos quadrados dos resíduos, SQR1 e SQR2 em que SQR1 representa a soma dos quadrados dos resíduos a partir da regressão correspondente aos valores menores de X_i e SQR2 a partir do conjunto com maiores valores de X_i . Essas somas têm cada uma $\frac{(n-c)}{2} - k$ ou $\left(\frac{n-c-2k}{2}\right)$ graus de liberdade em que k é o número de parâmetros a serem estimados, incluindo o intercepto. Para o caso de duas variáveis k=2;
- IV- Calcule a razão:

$$\lambda = \frac{SQR_2 / gl}{SQR_1 / gl} \quad (XXIII)$$

Pressupondo que os u_i sejam distribuído normalmente, e se a hipótese da homocedasticidade for válida, então poderemos mostrar que λ da equação inicial segue a distribuição F, tendo o numerador e o denominador (n-c-2k)/2 graus de liberdade. Se $\lambda = (F)$ calculado for maior que F crítico ao nível de significância escolhido, poderemos rejeitar a hipótese de homocedasticidade e dizer que a heterocedasticidade é muito provável.

5° *Teste de Breusch-Pagan-Godfrey*, aqui o sucesso do teste de Goldfeld-Quandt depende não só do valor de c, mas também de identificar a variável X

correta com a qual se colocam as observações em ordem. O que pode ser evitado com o teste de Breusch-Pagan-Godfrey (BPG). Seja o modelo de regressão linear com k variáveis explicativas:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (\text{XXIV})$$

Se a variância do erro σ_i^2 descrito como:

$$\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi}) \quad (\text{XXV})$$

σ_i^2 é função das variáveis não estocásticas Z; os X's podem servir de Z's.

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi} \quad (\text{XXVI})$$

σ_i^2 é função linear dos Z, se $\alpha_2 = \alpha_3 = \dots = \alpha_m = 0$, $\sigma_i^2 = \alpha_1$, que é uma constante, para sabermos se é homocedástica podemos testar a hipótese acima, através das etapas seguintes:

I- Calcule a equação:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + u_i \quad (\text{XXVII})$$

Por MQO e obtenha os resíduos $\hat{u}_1, \hat{u}_2, \dots, \hat{u}_n$;

II- Obtenha:

$$\sigma^2 = \sum \frac{\hat{u}_i^2}{n} \quad (\text{XXVIII})$$

Que é o estimador de máxima verosimilhança de σ^2 ;

III- Construa variáveis p_i definidas:

$$p_i = \frac{\hat{u}_i^2}{\sigma^2} \quad (\text{XXIX})$$

Que são cada resíduo elevado ao quadrado dividido por σ^2 ;

IV- Faça a regressão de p_i assim construída sobre os Z's:

$$p_i = \alpha_1 + \alpha_2 Z_{2i} + \dots + \alpha_m Z_{mi} + v_i$$

Em que v_i é o termo residual da regressão;

V- Obtenha a soma dos quadrados explicados SQE da equação anterior e defina:

$$\Theta = \frac{1}{2}(\text{SQE}) \quad (\text{XXX})$$

Supondo que os u_i , sejam normalmente distribuídos, se há homocedasticidade e se o tamanho da amostra n aumenta

indefinidamente, então $\Theta_{ass} \sim \chi^2_{m-1}$ em que ass significa assintoticamente, segue a distribuição de qui-quadrado com (m-1) graus de liberdade. Se numa aplicação $\Theta(=x^2)$ calculado for maior que o valor crítico x^2 no nível escolhido de significância, pode-se rejeitar a hipótese de homocedasticidade, caso contrário não será rejeitada.

6° *Teste geral de heterocedasticidade de White*, este não requer a hipótese da normalidade e é de fácil implementação. Considere o modelo de regressão com três variáveis na equação (XXVII), para implementar o teste de White segue-se as etapas:

I- Com os dados, calculamos a equação anterior e obtemos os resíduos u_i ;

II- Fazemos a regressão auxiliar:

$$\hat{u}_i^2 = \alpha_1 + \alpha_2 X_{2i} + \alpha_3 X_{3i} + \alpha_4 X_{2i}^2 + \alpha_5 X_{3i}^2 + \alpha_6 X_{2i} X_{3i} + v_i, \quad (XXXI)$$

Os resíduos ao quadrado da regressão original são calculados por regressão contra as variáveis ou regressores X originais, seus valores elevados ao quadrado e os produtos cruzados dos regressores. Pode-se incluir regressores com expoentes mais altos;

III- Sob a hipótese nula de que não há heterocedasticidade, pode-se mostrar que o tamanho da amostra (n) multiplicado pelo R^2 da regressão auxiliar segue assintoticamente a distribuição de qui-quadrado com graus de liberdade iguais ao número de regressores na regressão auxiliar $n \cdot R^2 \sim \chi^2_{gl}$ em que os graus de liberdade são definidos como anteriormente;

IV- Se o valor de qui-quadrado obtido na equação anterior excede o valor crítico do qui-quadrado ao nível escolhido de significância, então há heterocedasticidade, se não excede não há heterocedasticidade e na regressão dos resíduos $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = \alpha_6 = 0$.

O teste de White pode ser um teste heterocedástico puro ou de erro de especificação ou ambos. Se não houver termos de produtos cruzados, ele será um teste de heterocedasticidade pura.

Há vários outros testes de heterocedasticidade, cada qual baseado em certas hipóteses. De que maneira decidimos qual o melhor? Ao compararmos os testes, devemos ter em atenção o seu tamanho, potência e a sensibilidade a discrepâncias. Assim o teste de White para heterocedasticidade, que é o mais fácil de ser aplicado apresenta algumas limitações, baixando a potência contra as alternativas, além de quase nada ajudar na identificação dos fatores ou variáveis que causam heterocedasticidade. Já o teste de Breusch-Pagan-Godfrey é sensível à hipótese da normalidade. No teste de Goldfeld-Quandt, se omitirmos muitas observações, poderá diminuir seu desempenho. Contudo para sabermos que teste aplicar a dado exercício, deve-se analisar bem que tipo de hipóteses apresenta, qual o seu nível de significância e a sensibilidade.

A heterocedasticidade não destrói as propriedades de não tendenciosidade e consistência dos estimadores de MQO, mais deixam de ser eficientes, mesmo em grandes amostras. Esta ineficiência torna duvidoso o valor dos procedimentos usuais dos testes de hipóteses. Portanto medidas corretivas podem ser necessárias. Existem duas maneiras de efetuar esta correção: I^a Quando σ_i^2 é conhecido, o método dos mínimos quadrados ponderados é o mais prático para corrigir a heterocedasticidade, visto que os estimadores obtidos são o melhor estimador linear não tendencioso;

II^a Quando σ_i^2 não é conhecido, haverá uma forma de obter estimativas consistentes das variâncias e covariâncias dos estimadores de MQO, mesmo quando há heterocedasticidade? Para White esta estimativa pode ser realizada de modo que inferências estatísticas válidas assintoticamente possam ser feitas sobre os verdadeiros valores dos parâmetros. Hoje, vários programas apresentam as variâncias de heterocedasticidade de White e erros padrão com as variâncias dos MQO e erros padrão usuais. Os erros padrão corrigidos para heterocedasticidade de White são chamados erros padrão brutos. Os erros padrão corrigidos para heterocedasticidade são maiores que os obtidos pelos MQO, com base nestes os regressores são estatisticamente significantes e para os estimadores de White não são. Uma desvantagem do método de White é que os estimadores obtidos podem não ser tão eficientes quanto os obtidos pelos métodos que transformam dados,

para refletir tipos específicos de heterocedasticidade. Todas as transformações estudadas especulam a natureza do σ_i^2 . Dependendo da natureza do problema e da gravidade da heterocedasticidade, determinaremos qual das transformações discutidas funcionará. Há alguns problemas com as transformações que se devem ter em conta: - Indo além do modelo de duas variáveis, podemos não saber quais das variáveis X e Y deve ser escolhida para transformar os dados; - A transformação logarítmica, não é aplicável se alguns dos valores de X e Y forem zero ou negativos; - Há o problema de correlação espúria em que se encontra correlação nas razões das variáveis, mesmo que as originais não estejam correlacionadas, ou seja, aleatórias. Assim no modelo linear as variáveis Y e X podem não estar correlacionadas, mas no modelo transformado $Y_i/X_i = \beta_1(1/X_i) + \beta_2$, Y_i/X_i e $1/X_i$, estão frequentemente correlacionadas; - Se σ_i^2 não são conhecidos e são estimados com base em uma ou mais transformações, todos os procedimentos de uso dos testes t e F, são válidos somente para amostras maiores. Deve-se ter cuidado na interpretação dos resultados com base nas várias transformações em amostras pequenas ou finitas.

Quando devemos ficar preocupados com o problema da heterocedasticidade? A heterocedasticidade nunca foi razão para descartar-se um modelo que, sob outros aspetos, é considerado bom. A gravidade da heterocedasticidade pode não ser grande o suficiente para gerar preocupações. Lembremo-nos de que os estimadores de MQO são lineares, não tendenciosos e têm distribuição normal assintoticamente em grandes amostras.

2.5.3 Autocorrelação

A autocorrelação pode ser definida como a correlação entre membros de séries de observações ordenadas no tempo (séries temporais) ou no espaço (dados de corte transversal), segundo Gujarati (terceira edição). No modelo clássico de regressão linear pressupõe-se que não existe autocorrelação entre as perturbações aleatórias do termo de erro u_i (que é nula a covariância entre duas quaisquer perturbações aleatórias), isto é: $E(u_i, u_j) = Cov(u_i, u_j) = 0, i \neq j$. O uso do termo correlação, quando se refere à covariância, é

justificado pelo facto de ser diferente de zero o coeficiente de correlação linear de duas variáveis aleatórias se também for diferente de zero a covariância entre elas, por outro lado, o uso do prefixo auto encontra justificação por estar em causa, a correlação entre realizações de uma mesma variável, a perturbação aleatória u . [O termo de erro relacionado a qualquer uma das observações não é influenciado pelo termo de erro de qualquer outra observação. Se empregamos dados de corte transversal em uma regressão das despesas de uma família sobre a renda familiar, o efeito de um aumento da renda da família nesses gastos não deverá afetar os gastos de outra família, se houver esta dependência, estaremos perante um problema de autocorrelação]. Embora hoje seja comum tratar os termos autocorrelação e correlação serial como sinónimos, alguns autores como Tintner os diferenciam. Para Tintner (1965), a autocorrelação não passa de uma correlação desfasada entre determinada série com ela mesma, com uma desfasagem de algumas unidades de tempo. Já a correlação serial é uma correlação desfasada entre duas séries diferentes. A autocorrelação ocorre mais frequentemente, mas não exclusivamente, quando se lida com amostras de natureza cronológica ou temporal. Poderá acontecer que alguns dos factores aleatórios que influenciam o comportamento de um mesmo agente ao longo do tempo tenham efeitos persistentes de um período ou momento para os seguintes. Nesse movimento o valor da série em um ponto do tempo é maior que o anterior. Há um impulso embutido nele que continua até que algo aconteça para desacelerá-lo. Logo em regressões que envolvem séries temporais, as observações sucessivas tendem a ser interdependentes: Há outras situações em que é expectável haver autocorrelação, em que esta é induzida por *erro de especificação do modelo*. O facto de se usar um modelo incorretamente especificado, pela não consideração de uma das variáveis explicativas relevantes, pode fazer com que nesse modelo exista autocorrelação; muitas vezes a inclusão de tais variáveis elimina o padrão desta autocorrelação. O mesmo poderá acontecer com erros de especificação no que diz respeito à *forma funcional do modelo*. Para além destes casos existe ainda o *fenómeno da teia de aranha*, que tem haver com a oferta de produtos agrícolas, em que a oferta reage ao preço com a desfasagem de um período, uma vez que as decisões de oferta levam tempo

a serem implementadas. No início do plantio da safra deste ano, os agricultores estão influenciados pelo preço vigente no ano anterior. Se os agricultores produzirem demais este ano, terão de reduzir a produção no ano seguinte, tendo em conta a lei da procura e da oferta. Também temos ainda o caso da *desfasagem*, em que verificamos que as despesas do período atual dependem, dentre outras coisas, das despesas do período anterior. Uma regressão deste tipo é conhecida como auto regressão, e que uma das variáveis explanatórias é o valor desfasado da variável dependente. A *manipulação dos dados* é outra forma de obter autocorrelação, quando numa análise aplicada são manipulados os dados brutos, o que suaviza os dados amenizando as flutuações dos dados, outra fonte de manipulação é a *interpolação* ou a *extrapolação* de dados. Na *transformação de dados*, a autocorrelação é induzida como resultado da transformação do modelo original. Há várias razões pelas quais o termo de erro em um modelo de regressão pode ser auto correlacionado. Mas autocorrelação propriamente dita ocorre no primeiro caso. A autocorrelação pode ser tanto positiva quanto negativa, embora a maior parte das séries temporais económicas apresentem – na positiva, na sua maioria evolui para cima ou para baixo por longos períodos e não apresenta oscilações constantes.

O que acontecerá aos estimadores de MQO e suas variâncias se introduzirmos autocorrelação nos termos de erro supondo que $E(u_t u_{t+s}) \neq 0$ ($s \neq 0$), mantendo todas as outras hipóteses do modelo clássico? (t nos indica que lidamos com séries temporais). Como ponto de partida ou de primeira aproximação, supomos que as perturbações são geradas por: $u_t = \rho u_{t-1} + \varepsilon_t$, $-1 < \rho < 1$, em que ρ é o coeficiente autocovariância, ε_t é o termo de erro estocástico, que atenda a hipótese padrão dos MQO: $E(\varepsilon_t) = 0$; $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$; $\text{Cov}(\varepsilon_t, \varepsilon_{t+s}) = 0, s \neq 0$.

Em engenharia um termo de erro com as propriedades anteriores é chamado de *ruído branco* (*white noise*). A equação $u_t = \rho u_{t-1} + \varepsilon_t$ postula que o valor do termo de erro no período t é igual a ρ vezes o seu valor no período anterior, acrescido de um termo de erro puramente aleatório. Conhecido como processo autorregressivo de primeira ordem de Markov ou de primeira ordem, designado como AR (1). Esta equação é chamada de autorregressiva

por ser interpretada como a regressão de u_t na primeira desfasagem de um período. De primeira ordem porque o valor imediatamente anterior e u_t estão envolvidos na sua desfasagem máxima de 1. ρ coeficiente de auto covariância, representa também o coeficiente de autocorrelação de primeira ordem (ou de defasagem 1). De AR(1), vem:

$$Var(u_t) = E(u_t^2) = \frac{\sigma_\varepsilon^2}{1-\rho^2} \quad (XXXII)$$

$$cov(u_t, u_{t+s}) = E(u_t, u_{t+s}) = \rho^s \frac{\sigma_\varepsilon^2}{1-\rho^2} \quad (XXXIII)$$

$$Cor(u_t, u_{t+s}) = \rho^s \quad (XXXIV)$$

Em que $cov(u_t, u_{t+s})$ [$Cor(u_t, u_{t+s})$] é a covariância [correlação] entre termos de erro separados por s períodos, logo $cov(u_t, u_{t+s}) = cov(u_t, u_{t-s})$ e $Cor(u_t, u_{t+s}) = Cor(u_t, u_{t-s})$. Como ρ é constante e de valor $(-1; +1)$, AR (1) mostra que a variância é homocedástica, embora u_t esteja relacionado com o seu valor passado e o de vários períodos anteriores. Se $|\rho| < 1$ o valor da covariância diminuirá à medida que se retrocede a um passado distante; Se $|\rho| > 1$ é estacionário, a média, variância e covariância não variam ao longo do tempo e se $\rho = 1$ não está definida.

O processo AR (1) é muito usado devido a sua simplicidade em relação aos outros processos AR de ordem superior e porque em muitas aplicações, é de grande utilidade, uma vez que quantidades consideráveis de trabalhos teóricos e práticos foram feitos usando este processo.

Para se determinar o estimador BLUE em presença de autocorrelação, tem-se em conta o modelo de regressão de duas variáveis e o processo AR(1), assim:

$$\beta_2^{MQG} = \frac{\sum_{t=2}^n (x_t - \rho x_{t-1})(y_t - \rho y_{t-1})}{\sum_{t=2}^n (x_t - \rho x_{t-1})^2} + C \quad (XXXV)$$

Com C factor de correlação (não considerado na prática), cuja variância:

$$Var\beta_2^{MQG} = \frac{\sigma^2}{\sum_{t=2}^n (x_t - \rho x_{t-1})^2} + D \quad (XXXVI)$$

Com D factor de correlação (desconsiderado na prática). Vemos que o estimador β_2^{MQG} é obtido dos MQG, onde encontramos informações adicionais disponíveis (sobre heterocedasticidade ou autocorrelação) no

processo direto de estimação por transformação de variáveis, já nos MQO estas não são diretas. Como se pode ver o estimador MQG de β_2 apresenta o parâmetro de autocorrelação ρ , fazendo uso de todas as informações disponíveis, já na fórmula do MQO é ignorado, o que torna o estimador de MQG num estimador BLUE. Contudo com a autocorrelação o estimador BLUE de β_2 e a variância mínima é dada pela equação $Var\beta^{MQG}_2$.

Na presença de autocorrelação, os estimadores de MQO são lineares, não tendenciosos, consistentes, com distribuição normal assintótica, mas não eficientes. No entanto o que acontece com os procedimentos habituais de teste de hipótese se continuarmos a utilizar os estimadores dos MQO? Aqui distinguem-se dois casos fundamentais:

1. *Estimação por meio de MQO considerando a autocorrelação;* Na presença de autocorrelação, os estimadores de MQO, embora não tendenciosos consistentes e assintoticamente normalmente distribuídos, não são eficientes. Os procedimentos de inferência com base nos testes t, F e χ^2 tornam-se inadequados. Griliches e Rao em um estudo de monte Carlos verificaram que, numa amostra relativamente pequena onde o coeficiente de auto correlação, $\rho < 0,3$, o MQO terá resultados satisfatórios. O que mostra que em pequenas amostras em que $\rho < 0,3$, os MQO são aplicados satisfatoriamente. Mostrando que β_2 não é MELNT, mesmo recorrendo a $var(\beta_2)_{AR1}$, os intervalos de confiança obtidos são susceptíveis de serem mais amplos que os dos MQG. Mesmo que o tamanho da amostra aumente indefinidamente, β_2 (estimado) não é assintoticamente eficiente. Estamos propensos a declarar que um coeficiente é estatisticamente insignificante (não diferente de zero), embora possa não ser (com base nos MQG, que é o procedimento correto). Logo para estabelecermos intervalos de confiança e testar hipóteses devemos usar os MQG e não os MQO, mesmo que os estimadores obtidos por estes últimos sejam não tendenciosos e consistentes.
2. *Estimação por meio de MQO não considerando a autocorrelação;* Se não utilizando β_2 (estimado), usarmos $var(\beta_2\text{estimado}) = \sigma^2 / \sum x_t^2$, que ignora na totalidade o problema da autocorrelação, ou seja, se

acreditarmos que as hipóteses habituais do modelo clássico são válidas, os erros surgirão pela seguinte razão:

- 1- A variância residual $\sigma^2 = \sum u_t^2 / (n-2)$ que subestimar o verdadeiro σ^2 ;
- 2- Como resultados serão obrigados a superestimar R^2 ;
- 3- Mesmo que σ^2 não esteja subestimado, a $Var(\beta^2)$ pode subestimar a $Var(\beta^2)_{AR1}$, sua variância sob a autocorrelação (da primeira ordem), apesar da ineficiência desta última comparada com $Var(\beta^2)^{MQG}$;
- 4- Os testes comuns de significância t e F deixam de ser válidos, e se aplicados, nos levarão a conclusões equivocadas sobre a significância estatística dos coeficientes de regressão estimados.

Existindo autocorrelação, dada por AR(1), demonstra-se que:

$$E(\sigma^2) = \sigma^2 \{n - [2/(1-\rho)] - 2\rho r\} / n - 2 \quad (XXXVII)$$

Em que $r = \sum_{t=1}^{n-1} x_t^2$, interpretado como coeficiente de correlação amostral entre valores sucessivos de x. Se ρ e r forem positivos evidencia-se que $E(\sigma^2) < \sigma^2$, a fórmula da variância residual subestimar em média o verdadeiro σ^2 . Não sendo subestimado σ^2 , a $var(\beta^2)$ é um estimador tendencioso da $var(\beta^2)_{AR1}$, se ρ é positivo e se os X forem positivamente correlacionados (o que é verdadeiro na maioria das séries temporais económicas), então: a variância de β^2 calculado por MQO subestima sua variância sob AR(1); $Var(\beta^2) < var(\beta^2)_{AR1}$. Usando $Var(\beta^2)$ subestimaremos o erro padrão do estimador β_2 . Logo ao calcularmos a razão $t = \beta_2 / ep(\beta_2)$, sob a hipótese de que $\beta_2 = 0$, estará superestimando o valor de t e, a significância estatística do β_2 estimado. A situação piora se, além disso, σ^2 for subestimado.

Para detenção da autocorrelação distinguem-se os seguintes métodos:

- 1- *Método gráfico*, a hipótese da ausência da correlação do modelo clássico refere-se aos termos de erro da população, u_t observados diretamente. Temos as suas proxies, os resíduos \hat{u}_t , obtidos pelo procedimento habitual dos MQO. Mesmo que os termos de erro sejam homoscedásticos e não correlacionados, seus estimadores, os resíduos, \hat{u}_t , são heterocedásticos e autocorrelacionados. Contudo, pode-se demonstrar que, à medida que a amostra aumenta indefinidamente, os resíduos tendem a convergir para seus

verdadeiros valores, os u_t . Um exame visual de u_t ou u_t^2 fornece informações úteis quanto à autocorrelação e não só. Os resíduos podem ser analisados de diferentes pontos de vista. Podemos plotá-los contra o tempo, uma **plotagem sequencial no tempo**, como podemos plotar os **resíduos padronizados** contra o tempo, que são os resíduos u_t divididos pelo erro padrão da regressão $\sqrt{\sigma^2}$, ou seja, u_t/σ . U_t e σ são medidos na mesma unidade que o regressando Y . Os valores dos resíduos padronizados são números puros e podem ser comparados com os resíduos padronizados de outras regressões. Além disso, têm média zero e variância aproximadamente unitária (ditos normalizados), o que ocorre também em grandes amostras. Este método apesar de poderoso e sugestivo, é de natureza subjetiva ou qualitativa, mas existem métodos quantitativos que podem complementar esta abordagem qualitativa.

- 2- *teste das carreiras*, também conhecido por *teste de Geary*, é um teste não paramétrico (onde não se fazem hipóteses sobre a distribuição de probabilidade das quais as observações são extraídas). Uma carreira é uma sequência ininterrupta de um símbolo ou atributo, como + ou -. A extensão de uma carreira é o número de elementos que a formam. Numa sequência de observações rigorosamente aleatórias, as carreiras são examinadas tendo como base um teste e a aleatoriedade das mesmas. As carreiras a observar são de mais ou de menos em relação ao número de carreiras esperada em uma sequência rigorosamente aleatória? Quando numa sequência de observações verificamos que há carreiras de mais, então, os resíduos frequentemente alteram o sinal, indicando uma correlação serial negativa, e se as carreiras forem poucas, então, ocorrerá uma correlação serial positiva dos resíduos. Seja: N =número total de observações $=N_1+N_2$; N_1 =número de sinais positivos; N_2 =número de sinais negativos; e R =número de carreiras.

Partindo da hipótese nula de que os resultados sucessivos dos resíduos são independentes e que $N_1 < 10$ e $N_2 > 10$, o número de carreiras é assintoticamente normalmente distribuído com Média:

$$E(R) = \frac{2N_1N_2}{N} + 1 \quad (\text{XXXVIII})$$

E Variância:

$$\sigma_R^2 = \frac{2N_1N_2(2N_1N_2 - N)}{(N)^2(N-1)} \quad (\text{XXXIX})$$

Observação: $N = N_1 + N_2$

Sendo a hipótese nula de aleatoriedade sustentável, de acordo com as propriedades da distribuição normal espera-se que:

$$\text{Prob}[E(R) - 1,96\sigma_R \leq R \leq E(R) + 1,96\sigma_R] = 0,95$$

O que nos dá a entender que existe 95% de probabilidade de R estar no intervalo de confiança desejado, recorrendo à seguinte regra: “Não rejeite a hipótese nula de aleatoriedade com 95% de confiança se R, o número de carreiras, ficar no intervalo de confiança citado; rejeite-a se o R estimado ficar fora desse limite”.

Como regra geral, havendo autocorrelação positiva, o número de carreiras será pequeno; e sendo negativa a autocorrelação, o número de carreiras será grande.

3- *Teste d de Durbin-Watson*, o teste desenvolvido por Durbin e Watson é o mais eficaz para a detenção da autocorrelação serial, este teste é também conhecido como estatística d de Durbin-Watson, definido por: $d = \frac{\sum_{t=2}^{t=n} (u_t - u_{t-1})^2}{\sum_{t=1}^{t=n} u_t^2}$, que é a razão da soma das diferenças elevadas ao quadrado, entre resíduos sucessivos e a SQR. No numerador da estatística d perde-se uma observação no cálculo das diferenças sucessivas, fazendo com que o seu número de observações seja do tipo n-1. A sua vantagem é a de basear-se nos resíduos estimados, do cálculo de análise de regressão. Tornando comum informar o d de Durbin-Watson com distintas medidas (R^2 ; R^2 ajustado; t e F). É importante estar atento às hipóteses que fundamentam estatística d:

1-o modelo de regressão inclui o termo de intercepto, estando ausente, refaz-se a regressão, incluindo o intercepto para obter SQR;

2- As variáveis explanatórias X, não são estocásticas (fixas) em amostras repetidas;

3- Os termos de erro u_t são gerados pelo processo autorregressivo de primeira ordem: $u_t = \rho u_{t-1} + \varepsilon_t$. Não é usado em casos de detenção de processos autorregressivos de ordem superior a 1;

3- O termo de erro u_t , deve ser distribuído normalmente;

5- os valores dos termos desfasados da variável dependente do modelo de regressão não são inclusos como variáveis explanatórias. Este teste não se aplica a modelos autorregressivos;

6- Não devem faltar observações nos dados, caso faltem à estatística d não faz concessão às observações em falta.

É difícil derivar a probabilidade (amostragem) de estatística d , porque depende de uma maneira complicada dos valores de X da amostra, o que se torna compreensível por d ser calculado dos \hat{u}_t , dependendo de determinados X . Diferentemente dos testes t , F ou X^2 , não há um valor crítico que leve a rejeição ou aceitação nos termos de erro u_i . Durbin e Watson, determinaram um limite inferior d_L e um limite superior d_U , que, se o d calculado estiver fora destes valores críticos, toma-se uma decisão a respeito da presença de correlação serial positiva ou negativa. Esses limites dependem do número de observações n e do número das variáveis explanatórias, e não dos valores assumidos por essas variáveis.

Os limites de d são de 0 a 4, estabelecidos da seguinte forma:

$$d = \frac{\sum_{t=1}^{t=n} (u_t - u_{t-1})^2}{\sum_{t=1}^{t=n} u_t^2} \quad (\text{XXXX})$$

Expandindo vem:

$$d = \frac{\sum u_t^2 + \sum u_{t-1}^2 - 2 \sum u_t u_{t-1}}{\sum u_t^2} \quad (\text{XXXXI})$$

$\sum u_t^2 \approx \sum u_{t-1}^2$, diferindo em apenas uma observação. Logo:

$$d \approx \frac{2(1 - \sum u_t u_{t-1})}{\sum u_t^2} \quad (\text{XXXXII})$$

Podemos definir:

$$\rho = \frac{\sum u_t u_{t-1}}{\sum u_t^2} \quad (\text{XXXXIII})$$

Como coeficiente de autocorrelação de primeira ordem amostral, um estimador de ρ . Daí temos $d \approx 2(1-\rho)$, como $-1 \leq \rho \leq 1$, de $d \approx 2(1-\rho)$ implica $0 \leq d \leq 4$, qualquer valor estimado de d encontra-se nesses limites. Daí é evidente que, se $\rho=0$, $d=2$; não havendo correlação serial então d estará em torno de 2. Se d for igual a 2 então não há correlação de primeira ordem seja ela positiva ou negativa.

Se $\rho=+1$, então existe correlação positiva perfeita nos resíduos, $d \approx 0$. Se $\rho=-1$, há correlação negativa perfeita entre os resíduos sucessivos e $d \approx 4$.

As etapas envolvidas no teste de Durbin-Watson, supondo que as hipóteses que o fundamentam são respeitadas, são:

- 1- Os resíduos da regressão obtêm-se dos MQO;
- 2- Calcula-se a estatística d de Durbin-Watson;
- 3- Determina-se d_L e d_U críticos, para dado tamanho da amostra e número de variáveis explanatórias;
- 4- Seguir as regras de decisão da estatística d e tirar as devidas conclusões.

O teste d apresenta uma desvantagem que é a de cair na zona de indecisão, impossibilitando a conclusão de existência ou não de autocorrelação de primeira ordem. Para resolver este problema, em muitas situações constatou-se que o limite d_U é aproximadamente o verdadeiro limite de significância e, no caso em que d fica na zona de indecisão usa-se o seguinte teste **d modificado**, dado o nível de significância α :

- 1- $H_0: \rho=0$ versus $H_1: \rho>0$. Rejeitar H_0 ao nível α se $d < d_U$. Há autocorrelação positiva estatisticamente significativa;
- 2- $H_0: \rho=0$ versus $H_1: \rho<0$. Rejeitar H_0 ao nível α se o nível estimado $(4-d) < d_U$. Então há evidência estatística significativa de autocorrelação negativa;
- 3- $H_0: \rho=0$ versus $H_1: \rho \neq 0$. Rejeitar H_0 ao nível 2α se $d < d_U$ ou $(4-d) < d_U$. Então há evidência estatística significativa de autocorrelação positiva ou negativa.

A zona de indecisão estreita-se à medida que o tamanho da amostra aumenta. Segundo as tabelas de Durbin-Watson. Por sua vez, o programa Shazam dá o *p-valor*, ou seja, a probabilidade exata do valor d calculado. Os cálculos computadorizados vieram facilitar o cálculo do valor p calculado da

estatística d . O teste d de Durbin-Watson é tão respeitado que muitas vezes os seus utilizadores esquecem-se de aplicar as hipóteses que o fundamentam que são:

- 1- as variáveis explanatórias, ou regressores, são não estocásticas;
- 2- o termo de erro segue a distribuição normal;
- 3- os modelos de regressão não incluem os valores desfasados do regressando; e
- 4- apenas a correlação serial de primeira ordem é levada em conta. Uma estatística d significativa, nem sempre indica autocorrelação, mas pode ser indicação de omissão de variáveis relevantes no modelo.

Se um modelo de regressão contém valores desfasados do regressando, o valor d , é frequentemente em torno de dois, sugerindo ausência de autocorrelação em tais modelos. Há um viés contra o descobrir autocorrelação em tais modelos, o que não significa que os modelos autorregressivos não sofram do problema de autocorrelação. Durbin desenvolveu o teste h para correlação serial em tais modelos, mas este teste não é tão poderoso estatisticamente como o teste de Breusch-Godfrey, de modo a dispensar o uso do teste h . Se o termo de erro u_i não for NIID, o teste d , pode não ser confiável. Então o teste de carreira terá a vantagem de não fazer nenhum pressuposto quanto à distribuição do termo de erro. Sendo grande a amostra, pode-se empregar o teste d , já que demonstra-se:

$$\sqrt{n \left(\frac{1-d}{2d} \right)} \approx N(0;1)$$

Que em amostras grandes a sua estatística d transformada em amostras grandes segue a distribuição padrão normal. Tendo em vista a relação entre d e, ρ o coeficiente de autocorrelação de primeira ordem estimado

$d = 2(1 - \rho)$ segue-se: $\sqrt{n\rho} \approx N(0;1)$, isto é, em amostras grandes, a raiz quadrada do tamanho da amostra vezes o coeficiente de autocorrelação de primeira ordem estimado também segue a distribuição normal padrão. Assintoticamente, se a hipótese nula de correlação de primeira ordem igual a zero for verdadeira, a probabilidade de se obter um valor Z (uma variável

normal padronizada), será muito pequena (para uma distribuição normal padrão o valor Z crítico bilateral a 5% é de apenas 1,96 e a 1% é de cerca de 2,58). O problema mais sério com o teste d é o pressuposto de os regressores serem não estocásticos, isto é, seus valores serem fixos em amostras repetidas. Não sendo este o caso, então o teste d não é válido em amostras pequenas, finitas ou grandes. Em geral é difícil sustentar esse pressuposto em modelos económicos com dados de séries temporais. Segundo Hayashi (2000), a estatística de Durbin-Watson pode não ser útil em econometria de séries temporais, por existirem testes mais úteis de autocorrelação, baseados em amostras grandes.

Teste de Breusch-Godfrey (BG) << um teste geral de autocorrelação>>, este teste de autocorrelação que é mais genérico, foi criado para evitar algumas das armadilhas do teste d de Durbin-Watson, por não permitir regressores não estocásticos, como os valores desfasados do regressando; esquemas autorregressivos de ordem superior, como AR(1), AR(2)...; e **medidas móveis** simples ou de ordem mais elevadas de termos de erro de ruído branco, como ε_t na equação $u_t = \rho u_{t-1} + \varepsilon_t$, $-1 < \rho < 1$, como por exemplo na regressão:

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

O termo de erro pode ser representado como:

$$u_t = \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \lambda_2 \varepsilon_{t-2},$$

Indicando uma média móvel de três períodos do termo de ruído branco ε_t .

O **teste BG**, conhecido como **teste LM**, isto é baseado no princípio do Multiplicador de Lagrange é feito a partir do modelo de regressão de duas variáveis para ilustrá-lo, embora muitos regressores lhe possam ser acrescentados. Contudo valores desfasados podem ser acrescentados ao modelo. Sendo;

$$Y_t = \beta_1 + \beta_2 X_t + u_t,$$

Se o termo de erro u_t seguir um esquema autorregressivo de ordem p , AR(p) será:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + \varepsilon_t$$

Em que \mathcal{E}_t é um termo de erro de ruído branco. Como se vê, é uma simples extensão de AR(1), e a hipótese nula a ser testada é: $H_0: p_1=p_2=\dots=p_p=0$, logo não há autocorrelação serial de qualquer ordem.

O teste BG apresenta as seguintes etapas:

- 1- Estimação da equação $Y_t = \beta_1 + \beta_2 X_t + u_t$, pelo MQO obtendo os resíduos, u_t ;
- 2- Faz-se a regressão \hat{u}_t contra X_t original, havendo mais de uma variável X no modelo original também serão incluídas e $u_{t-1}, u_{t-2}, \dots, u_{t-p}$, em que os últimos são os valores desfasados dos resíduos estimados na primeira etapa. Se $p=4$, introduz-se os quatro valores desfasados dos resíduos como regressores adicionais do modelo. Para fazer esta regressão teremos apenas $(n-p)$ observações. Em suma:

$$U_t = \alpha_1 + \alpha_2 X_t + p_1 u_{t-1} + p_2 u_{t-2} + \dots + p_p u_{t-p} + \mathcal{E}_t$$

e obtém-se R^2 dessa regressão;

- 3- Sendo grande o tamanho da amostra, Breuch e Godfrey mostraram que $(n-p)R^2 \sim \chi^2_p$, assintoticamente, $n-p$ vezes o valor R^2 obtido da regressão auxiliar u_t segue a distribuição do qui-quadrado com p graus de liberdade. Se $(n-p)R^2$ excede o valor crítico do qui-quadrado numa aplicação no nível de significância escolhido, a hipótese nula é rejeitada e p na equação $u_t = p_1 u_{t-1} + p_2 u_{t-2} + \dots + p_p u_{t-p} + \mathcal{E}_t$ é estático e diferente de zero.

O teste BG apresenta alguns pontos críticos que são:

- 1- Os regressores do modelo de regressão podem conter valores desfasados do regressando Y, ou seja, Y_{t-1}, Y_{t-2}, \dots , que podem parecer variáveis explanatórias. Comparando o modelo à restrição do teste de Durbin-Watson, pode não haver valores desfasados do regressando entre os regressores;
- 2- O teste BG é aplicável embora os termos de erro sigam um processo de **média móvel (MA)** de ordem p , isto é;

$$u_t = \mathcal{E}_t + \lambda_1 \mathcal{E}_{t-1} + \lambda_2 \mathcal{E}_{t-2} + \dots + \lambda_p \mathcal{E}_{t-p},$$

com \mathcal{E}_t termo de erro branco (que satisfaz as hipóteses clássicas);

- 3- se em $u_t = p_1 u_{t-1} + p_2 u_{t-2} + \dots + p_p u_{t-p} + \mathcal{E}_t$, $p=1$, correspondente a autorregressão de primeira ordem, o teste BG é conhecido por **teste M de Durbin**;

- 4- O teste BG apresenta a desvantagem de o valor de p (*duração da defasagem*) não é especificado com antecedência. O valor de p deve ser experimentado. Para tal pode-se usar os critérios de informação de **Akaike** e **Schwarz** para saber o número de defasagens;
- 5- Dados os valores das variáveis X e os defasados de u , supõe-se que a variância de:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \dots + \rho_p u_{t-p} + \varepsilon_t$$

Seja homocedástica. Utilizam-se diversos testes de autocorrelação por não haver um teste que venha a ser o melhor e mais eficiente, por isso o analista deve considerar diversos procedimentos de teste para detetar a presença ou estrutura, ou ambas, de autocorrelação (a eficiência de um teste estatístico é 1 menos a probabilidade de aceitar uma hipótese falsa, a eficiência máxima de um teste é 1 e a mínima é 0). Quanto mais próxima de zero estiver à eficiência de um teste, pior será, e quanto mais próxima de 1, mais eficiente se torna.

Para a correção da autocorrelação toma-se algumas medidas que são:

- 1- Verificar se trata-se de uma autocorrelação pura e não o resultado equivocado do modelo. Visto que por vezes observa-se padrões nos resíduos, porque o modelo foi mal especificado (exclusão de variáveis importantes ou porque a forma funcional é incorreta);
- 2- Sendo autocorrelação pura, pode-se transformar o modelo original de modo a que o modelo transformado não tenha problemas de autocorrelação pura, recorrendo a alguns tipos de métodos de mínimos quadrados generalizados (MQG);
- 3- Em grandes amostras, pode-se usar o método de **Newey-West** na obtenção de erros padrão dos estimadores de MQO corrigidos para a autocorrelação, que é uma extensão do método de erro padrão para heterocedasticidade de White;
- 4- Em algumas situações insiste-se no uso dos MQO. Como em pequenas amostras em que ρ seja menor que 0,3. No entanto, amostras grandes ou pequenas são uma questão relativa e deve-se recorrer a um julgamento baseado na prática.

Podemos aqui ver ainda a relação entre as variáveis dummy e a autocorrelação. Dos modelos de regressão da dummy binária, tem-se:

$$Y_i = \alpha_1 + \alpha_2 D_t + \beta_1 X_t + \beta_2 (D_t X_t) + u_t$$

Em que Y é a poupança, X a renda e D as observações por período, modelo este estimado com pressupostos usuais dos MQO. Se u_t seguir um processo autorregressivo de primeira ordem, AR(1) e se ρ conhecido ou estimado, pode-se usar o método das diferenças generalizadas para estimar os parâmetros do modelo livre de autocorrelação de primeira ordem. A presença de variáveis binárias D impõe um problema especial (a variável binária apenas classifica uma observação como pertencente ao primeiro ou segundo período). Para transformá-las seguimos os procedimentos abaixo:

- 1- Da equação acima vemos que os valores de D são zero em todas as observações do primeiro período, no segundo período para as primeiras observações o valor de D é de $1/(1-\rho)$ e 1 para as restantes observações;
- 2- A variável X_t é transformada em $(X_t - \rho X_{t-1})$. Nesta observação perde-se uma observação, a menos que se recorra à **transformação de Prais-Winsten** para primeira observação;
- 3- O valor de $D_t X_t$ é zero pra todas observações no primeiro período ($D_t = 0$ no primeiro período), no segundo período a primeira observação assume o valor de $D_t X_t = X_t$ e as demais observações passam a ser $(D_t X_t - \rho D_{t-1} X_{t-1}) = (X_t - \rho X_{t-1})$. (no 2º período $D=1$). A primeira observação do segundo período é fundamental. Tratada como sugerido acima, não haverá problemas na estimação das regressões sujeitas a autocorrelação AR(1), como a da equação Y_i .

No caso dos modelos ARCH e GARCH, vemos que se o termo de erro u_t estiver correlacionado com u_{t-1} num processo AR(1) ou com vários termos de erro desfasados num processo geral AR(ρ), a autocorrelação observada na variância σ_t^2 e seus valores desfasados em um ou mais períodos ocorre na previsão de séries temporais como preço de ações, taxa de inflação e de câmbio, que recebe os nomes de **heterocedasticidade condicional auto regressiva (ARCH)**, se a variância do erro relaciona-se com u_t^2 no período anterior e de **heterocedasticidade condicional auto regressiva generalizada (GARCH)**, se a variância de erro estiver relacionada com u_t^2 em vários períodos anteriores. O que mostra que a autocorrelação não se

limita na relação entre os termos de erro antigos e atuais, mais também na relação entre erros de variações antigas e atuais.

A auto correlação e a heterocedasticidade podem coabitar se elaborar-se um modelo que permita usar o erro padrão ECHA que considera tanto a auto correlação como a heterocedasticidade, em amostras grandes, já que há autores que consideram que a autocorrelação só é detetada após o controlo da heterocedasticidade.

2.6 Resumo

A econometria é um ramo novo das ciências económicas, que veio simplificar ainda mais a vida dos economistas tornando os seus resultados mais fiáveis, reduzindo consideravelmente os riscos de se obter resultados duvidosos na análise de dados nos diversos estudos realizados, já que o seu estudo é mais precioso no que tange à vida económica-financeira de uma empresa ou negócio.

Os modelos econometricos são as vias usadas pelos econometristas para analisarem os resultados das suas atividades sem contudo violar as regras económicas, através de cálculos de algumas funções matemáticas e do suporte estatístico.

Os modelos econometricos podem ser do tipo determinístico ou exato que é mais geral e muito pouco utilizado em econometria por retirar a hipótese do erro; já o modelo do tipo estocástico é mais específico, por quantificar as variáveis de dada experiência económica tendo em conta os conceitos da teoria económica, buscando como base conceitos estatísticos e matemáticos como o da função linear, permitindo uma comparação com a ação das variáveis intervenientes na experiência a realizar para garantir a veracidade dos resultados.

Para a elaboração de um modelo econometrico até à fase da sua implementação necessita-se passar por 5 etapas que são: etapa de especificação do modelo em que, selecionam-se as variáveis explicativas da variável explicada que é a que se vai estudar, definindo a forma funcional da relação entre a variável explicada e as explicativas de forma correta. Na etapa de seleção da amostra, recolhem-se os dados da amostra que podem ser seccionais ou transversais, temporais ou cronológicos e ainda logitudinais ou de painel. As variáveis referentes a stock refletem um dado momento,

variáveis representativas de fluxos tem períodos anuais, semestrais, trimestral, semanal, etc. Esta é segunda etapa de estimação ou ajustamento do modelo, em que, o econometrista produz valores numéricos constantes e concretos que aproximem aos valores ignorados, recorrendo a métodos com propriedades desejáveis. A inclusão da variável u apresenta efeitos aditivos de natureza aleatória para acomodar no modelo uma multiplicidade de fatores que influenciam a variável explicada e não refletido convenientemente pelas variáveis explicativas. Já na etapa de avaliação dos resultados, o coeficiente de determinação nos indica a qualidade de ajustamento possibilitando a estimação do valor da variável explicada para cada observação, valores estes que ajustados para a variável explicada coincidem com os observados. Por último, temos a etapa de utilização dos resultados, em que averigua-se a confirmação dos dados, sua refutação ou modulação da teoria e a quantificação dos impactos de diferentes fatores sobre a variável em estudo.

Os modelos económicos servem para análise de estrutura, simulações, previsões e constatação da teoria económica. Ao serem aplicados em empresas para um ajuste de preços melhora os investimentos e na previsão das vendas futuras, resulta uma optimização das vendas. Por outro lado, tem havido certa confusão entre a economia matemática e a econometria, visto que o estudo empírico e a análise teórica complementam-se reforçando uma a outra, testando a validade das teorias em relação aos dados empíricos antes da sua aplicação. Contudo, não existe estatística em econometria que não se apoie na teoria económica como guia e a economia matemática é a teoria básica para tal estudo econométrico, por conter ferramentas matemáticas necessárias para desenvolver um estudo estatístico e econométrico.

No estudo de modelos econométricos surgem problemas como a multicolinearidade, a heteroscedasticidade e a autocorrelação, que comprometem a confiabilidade dos valores dos coeficientes estimados e a inferência estatística. No entanto, estes problemas podem ser detetados e corrigidos.

3. APLICAÇÃO DOS MODELOS ECONOMETRICOS NO ENSINO SUPERIOR ANGOLANO

No contexto angolano o ensino da econometria a nível superior, começa a ganhar corpo em 2002 com o fim da guerra civil no país e consequente conquista da paz efetiva. O que fez com que houvesse uma grande abertura para o surgimento de novas Universidades e Institutos Superiores tais como: Universidade Metodista de Angola, Universidade Metropolitana Privada, Universidade 11 de Novembro, Universidade Privada de Angola, Universidade Católica de Angola, o Instituto Superior Lusíada, Instituto Superior Jean Piaget, o Instituto Superior Politécnico do Kazenga, Instituto Superior Politécnico do Huambo, Instituto Superior Dom Bosco, Instituto Superior de Ciências Sociais e Relações Humanas, entre outros. Desde a independência existiu no país uma única Universidade Pública (Universidade Agostinho Neto), com pólos em algumas das 18 (dezoito) províncias do país. Com a paz efetiva surgiram no país muitas outras Universidades e Institutos do Ensino Superior, que se espalharam nas diversas províncias como a província do Lubango, a do Huambo, de Benguela, do Uíge, de Malange e em muitas outras. No entanto, todos os pólos da Universidade Agostinho Neto foram transformados em novas Universidades, assim o pólo de Benguela transformou-se na Universidade Katiavala Bwila, o pólo do Lubango é agora a Universidade Mandume ya Ndemufayo, no Uíge surgiu a Universidade Kimpa Vita, e assim por diante. Estas fizeram com que se criassem ou lecionassem cursos de acordo com a necessidade político-social do país e para tal se deveria formular os programas com base naquilo que se ensina a nível mundial, tendo em conta o desenvolvimento tecnológico e científico atual.

Em Angola existem 29 Universidades e Institutos Superiores já reconhecidos ou legalizados pelo Ministério da Educação até Abril de 2009, segundo o Diário da República (Terça feira, 7 de Abril de 2009).

Para que se controlasse este crescente fluxo de Universidades e Institutos Superiores no nosso país, tornando possível uma expansão ordenada da rede de Instituições de Ensino Superior, o governo achou por bem criar um Ministério que pudesse controlar com maior rigor o surgimento destas novas

instituições de ensino bem como a sua componente pedagógica. O que fez com que se criasse o Ministério do Ensino Superior em Angola. Este Ministério tem trabalhado com as instituições até então já reconhecidas e tem controlado tantas outras que têm surgindo com o tempo, o que faz com que, ainda haja algumas instituições que se encontram em vias de reconhecimento.

As instituições de Ensino Superior já reconhecidas pelo governo angolano localizam-se em apenas 7 das 18 províncias do país, onde a região norte possui a maior parte destas instituições distribuídas em apenas 3 (três) províncias que são: Luanda, Uíge e Malange; destas Luanda possui 17 Instituições e as outras duas tem apenas uma instituição pública cada. Nas regiões centro e sul do país encontram-se 10 destas instituições de ensino distribuídas em 4 províncias; das quais Benguela tem 4 instituições, Huambo tem 3, Lubango tem 2 e Kwanza-Sul tem 1. Para além destas, em todas as localidades citadas, existem tantas outras instituições em vias de reconhecimento.

Visto que o Ministério da Educação, Ciência e Tecnologia distribuiu o país em regiões académicas com base nas prioridades de desenvolvimento económico e social das regiões em que estas instituições estão inseridas, estas delimitam o âmbito territorial de atuação e expansão das instituições de Ensino Superior. O que faz com que, a autora desta pesquisa limite o seu estudo à região académica nº 2, que envolve as províncias de Kwanza-Sul e Benguela, o que facilita o seu estudo por se encontrar nela localizada e ter maior facilidade de se deslocar nestas localidades e ainda porque a maior parte das instituições superiores nela localizadas são extensões de outras já existentes na província de Luanda que possui a maior parte das instituições de ensino superior do país e que pertence à região académica nº 1; podendo os resultados deste estudo serem também aplicáveis a esta região e outras regiões académicas do país.

A região académica nº 2, tem 5 instituições do Ensino Superior já reconhecidas das quais 4 lecionam cursos de gestão em que se tem a econometria como disciplina. Todas elas encontram-se na província de Benguela, mais propriamente nos municípios de Benguela e Lobito. Para além destas a região possui mais uma instituição em vias de reconhecimento

que apresenta no seu repertório cursos de gestão que envolve o ensino da econometria como disciplina. A autora trabalhou com todas as universidades desta região que têm no seu repertório cursos em que se leciona a cadeira de econometria como disciplina e que já estão reconhecidas pelo Ministério do Ensino Superior, o que faz com que a amostra corresponda a 100% da população existente na região. Verificou-se que em Benguela, das instituições superiores existentes, apenas dois Institutos Superiores não lecionam cursos em que se enfoque a econometria como disciplina que são: O Instituto Superior Politécnico de Benguela (em vias de reconhecimento) e o Instituto Superior Politécnico Católico de Benguela (em vias de reconhecimento), esta última já tem um programa para o arranque em 2014 do curso de economia que faz referencia a disciplina de econometria e se encontra em fase de elaboração dos programas das diferentes cadeiras para o referido curso e de aquisição de professores (segundo informação da secretaria pedagógica desta instituição, 2012).

As técnicas de gestão surgem na região com a necessidade que o mercado financeiro tem apresentado nos últimos anos. Já que, com a paz efetiva no país, aumentam as oportunidades de investimentos no país e na região, nas diversas empresas nacionais e internacionais que procuram se fixar no país e principalmente na região académica nº2, aumentando assim o mercado de ação de um gestor económico-financeiro, bem como a necessidade de desenvolvimento de todos os ramos de gestão. Vê-se aqui que a formação académica segue as necessidades da sociedade, não se pode falar de desenvolvimento económico-financeiro sem fazer referência ao ramo da gestão financeira.

3.1 Análise da entrevista aos coordenadores de disciplina

Da entrevista feita aos coordenadores de cursos constatou-se que os cursos lecionados na região académica nº 2, são uma extensão daquilo que é a estrutura de cada Instituição Superior no país. No ensino da econometria uma das dificuldades encontradas é o facto de que a assimilação dos conteúdos pelos estudantes é muito mais limitada ao que o professor apresenta na sala de aulas por causa da dificuldade financeira de muitos

estudantes na aquisição de material de apoio (refere-se à bibliografia sugerida nos programas e pelo professor). Mesmo assim nota-se um grande interesse por parte dos estudantes em querer aproveitar o máximo possível do que se lhes apresenta. Os professores, apesar de na sua maioria serem formados no ramo da gestão, existem outros tantos formados em Matemática e Estatística que já veem lecionando esta cadeira há bastante tempo. No entanto não tem apresentado grandes dificuldades na transmissão do conteúdo econométrico. uma vez que têm recebido algumas orientações dos demais quanto aos assuntos mais complexos de se transmitir. Tal conteúdo foi cuidadosamente selecionado procurando dar ao estudante uma formação completa em muito pouco tempo. Para os coordenadores de cursos o estudante não deve limitar-se ao que o professor apresenta, porque o tempo em que este fica em contacto com o conteúdo na instituição não é suficiente para que adquira as qualidades necessárias para exercer a sua atividade profissional com eficácia.

As instituições têm distribuído às aulas de econometria em teóricas e práticas; nas aulas teóricas tem-se procurado apresentar ao aluno todo o conteúdo necessário para uma boa assimilação e que muito o ajudará na realização da sua atividade diária, facilitando o estudo de casos e a retirada de conclusões, bem como na tomada de decisão. Nas aulas práticas tem-se procurado trabalhar mais com dados reais de empresas já existentes na região, embora não tenha sido fácil obter tais dados visto que muitas empresas ainda tem a parte administrativa e financeira muito desorganizada, e se furtam de fornecer dados de sua empresa. Outra dificuldade no campo da transmissão do conteúdo é o fato de não existirem em todas as instituições laboratórios equipados com softwares específicos para aplicação de modelos econométricos. No entanto os conteúdos estão organizados por capítulos e cada capítulo possui os seus subtemas, no entanto os capítulos não tratam dos conteúdos na mesma intensidade nas distintas instituições. Verifica-se que conteúdos como o dos problemas de uma análise de regressão são mais bem proveitosos se estudados como um capítulo cada um deles. O mesmo acontece com o conteúdo sobre análise de regressão linear simples e múltiplo.

3.2 Análise dos programas de Econometria

Contudo do estudo feito aos programas das universidades da região académica nº 2, constatou-se que a econometria como disciplina é lecionada nos cursos de Economia, Gestão de Empresas, Economia e Gestão e no curso de Engenharia Informática de Gestão. Dos quais cada instituição leciona pelo menos dois cursos distintos e todas lecionam o curso de Economia e de Gestão de Empresas como se vê na tabela I. No entanto as universidades Katiavala Bwila e Jean Piaget de Angola apresentam -nos numa primeira instância os cursos de Economia e de Economia e Gestão de onde surgem respectivamente no 3º e 4º anos curriculares os cursos de Gestão de Empresas e de Organização e Gestão Empresarial como especialidade.

Tabela I- Cursos de gestão que lecionam a econometria como disciplina

Universidades/Cursos	Katiavala Bwila	Lusíadas de Angola	Jean Piaget de Angola
Economia	x	x	
Economia e Gestão			x
Gestão de Empresas	x	x	x
Engenharia Informática de Gestão			x
Fonte: Própria			

Nestes cursos no 1º e 2º anos curriculares lecionam-se as disciplinas de Matemática e Estatística, uma vez que estas tem uma grande influência naquilo que é o ensino da econometria como disciplina. A Matemática possibilita ao estudante um bom desempenho no estudo dos conteúdos econométricos no que diz respeito à elaboração e leitura de gráficos, bem como o aprimora para a realização de cálculos de uma forma mais sintetizada e prática. Já a Estatística ajuda-o a esclarecer questões de cálculo do erro e na compreensão da questão dos desvios, variâncias, rácios e muito mais. O que nos mostra que é fundamental que os gestores tenham um conhecimento mais profundo neste campo.

Tabela II- Ano curricular em que se leciona econometria

Cursos	1º Ano		2º Ano		3º Ano	
	1ºSem	2ºSem	1ºSem	2ºSem	1ºSem	2ºSem
Economia					X	
Economia e Gestão					X	X
Gestão de Empresas					X	
Informática de Gestão					X	

Fonte: Própria

Para o ensino da econometria, existe uma gritante necessidade de se organizar os conhecimentos, o que faz com que em todas as instituições de Ensino Superior esta venha a ser lecionada apenas no primeiro semestre do 3º ano curricular. Antes é preciso que o estudante tenha adquirido uma grande quantidade de conhecimentos Matemáticos e Estatísticos, que lhe são ministrados nos 1º e 2º anos curriculares da seguinte forma: no 1º ano em todas as instituições do Ensino Superior dos cursos de Gestão, ministra-se a disciplina de Matemática como sendo uma disciplina anual ou semestral, há instituições em que esta disciplina é subdividida em vários ramos das matemáticas, vindo a ser lecionada também no 2º ano, como ocorre no curso de Engenharia Informática de Gestão. A Matemática fornece à econometria as bases para que o estudante consiga fazer uma análise e leitura corretas, de gráficos, bem como a facilidade de cálculos Matemáticos ligados à economia. Já a Estatística é lecionada no 2º ano curricular distribuída nos dois semestres deste ano na maioria destas instituições, com exceção do curso de Engenharia Informática de Gestão que já no 1º ano e 1º semestre fala de Probabilidades e Estatística. Esta ajuda o estudante a assimilar informações necessárias como o cálculo do desvio padrão, variância e covariância, importantes para o estudo do erro, heterocedasticidade entre outros aspectos econométricos. Para resumo desta interligação entre as disciplinas de Matemática, Estatística e Econometria elaborou-se a tabela III.

Tabela III-`Disciplinas interrelacionadas ao ensino da Econometria

Universidades	Cursos	1º Ano	2º Ano	3º Ano
Katiavala bwila	Gestão de Empresas	Análise Matemática	Estatística	
	Economia	Análise Matemática	Estatística	Econometria
Lusíada de Angola	Gestão de Empresas	Análise Matemática	Estatística	Econometria
	Economia	Matemática	Estatística	Econometria I e II
Jean Piaget de Angola	Engenharia Informática de Gestão	Álgebra Linear e Geometria Analítica; Análise Matemática I; Probabilidades e Estatística.	Análise numérica	Econometria
	Economia e Gestão	Análise Matemática e Estatística I	Estatística II	Econometria

Fonte: Própria

Tabela IV- Objectivos a alcançar no ensino da econometria por cursos.

Objectivos	Universidades / Institutos					
	Katiavala Bwila		I.S.P. Lusíadas		I.S.P. Jean Piaget	
	Economia	Gestão de Empresas	Economia	Gestão de Empresas	Economia de Gestão	Informática de Gestão
Avaliar e aplicar os testes e tratamento para as correções desses problemas em modelos econométricos			X			
Capacitar para diagnosticar a presença de multicolinearidade, autocorrelação e heterocedasticidade.			x			
Compreender o conceito e objetivos da econometria			x	x		
Compreender o processo econométrico fazendo distinções entre os vários tipos de modelos				x		
Compreender o processo econométrico linear simples e geral			x			
Compreender os pressupostos básicos sobre os quais se fundamenta a validade das estimativas obtidas			X	x		

Descrever o funcionamento de uma parte ou o todo de um sistema	X					
Desenvolver a teoria e prática da análise de regressão linear					x	
Dotar os estudantes dos conhecimentos usados na resolução de problemas dos modelos econométricos						X
Especificar adequadamente algumas das formas funcionais econométricas mais usuais			x			
Especificar as várias variáveis em cada relacionamento	X					
Estimar e avaliar um modelo econométrico linear simples e geral				x		
Estimar e interpretá-los				x		
Estimar estatisticamente este relacionamento	X					
Habilitar o estudante à aplicação dos testes de hipóteses					x	
Habilitar o estudante a valorizar, compreender, criticar e analisar, a pesquisar, organizar e a emitir					X	

conclusões.						
Incorporar variáveis qualitativas na especificação de modelos econométricos				x		
Inserir variáveis especiais aos modelos			X			
Introduzir o estudante na especialidade dos conceitos e técnicas da econometria e suas aplicações (análise de regressão linear)					x	
Introduzir o relacionamento entre variáveis económicas no caso de duas variáveis	X					
Permite medir o grau de confiabilidade do modelo estimado em termos probabilísticos			X	x		
Testar este relacionamento empiricamente	X					

Fonte: Própria

Os objetivos nos distintos cursos estão mais orientados para a formação de um econometrista em todas as vertentes para compreender, avaliar, incorporar, estimar e especificar o que são os elementos fundamentais do processo econométrico de uma forma mais profunda. Apesar de os objetivos parecerem ser diferentes, todos estão interligados de uma ou outra forma. Assim vemos que o objectivo que diz “*Compreender o conceito e objetivos da econometria*” do I.S.P. Lusíadas tem o mesmo teor que o objectivo do I.S.P. Jean Piaget que diz “*Introduzir o estudante na especialidade dos conceitos e técnicas da econometria e suas aplicações*” e ainda assemelha-se ao objetivo da Katiavala Bwila que diz “*Introduzir o relacionamento entre variáveis económicas no caso de duas variáveis*”. Todos estes objetivos falam da necessidade que existe de se conhecer os elementos principais que regem a econometria. Com isto podemos dizer que, de uma forma geral, os objetivos foram traçados de tal maneira que ao elaborar a sua aula o docente tenha que apresentar ao estudante as ferramentas necessárias para que este tenha em mãos as ferramentas fundamentais para o tratamento de uma experiência econométrica.

Tabela V- Bibliografia indicada no ensino da econometria

Bibliografia indicada ao estudante	Universidades / Institutos					
	Katiavala Bwila		I.S.P. Lusíadas		I.S.P. Jean Piaget	
	Economia	Gestão de Empresas	Economia	Gestão de Empresas	Economia de Gestão	Informática de Gestão
Basic Econometrics Models e Economic Fore Casts (Gujarati D.)						x
Econometria (traduzido por Alfredo Alves de Farias)			x	x		
Econometria Básica (Gujarati)			x	X		
Econometria Básica: Teoria e Aplicações (Matos, Orlando Carneiro1997).			x	x		
Econometria schaums Exercícios e teoria (Oliveira M.M. e outros)	X					
Econometric Analysis (Greene)						x
Econometric Method's (J.Johnston)	X					
Elementos de econometria (São Paulo, Kmenta, J. 1988)	X					
Essentials of Econometrics (Gujarati, D. 1999)	X					
Estatística Aplicada a Administração e Gestão					X	

Estatística Aplicada a Administração Económica					X	
Estatística Aplicada à Economia e Administração (tradução de Carlos A. Crusius e Jandyra M. Fachel)	X					
Estatística Descritiva					X	
Estatística e Introdução a Economia (Sartotis, Alexandre)			X	x		
Lições de Econometria (Lange, O. Res-Editora limitada)	X					
Matemática e Estatística Aplicada					x	
Matemática para Economia e Administração [tradução de Mathematical analysis, Business and Economic Applications (Weber, Jean e.)]			X	x		
Mathematics for Economists (Yamane, Scond. Edicion)	X					
Modelos Económicos [cap.IV(Guillaume, M.)]	X					
Theory of Econometrics: An introductory Exposition of Econometrics Methods (Koutoyiannis A. 1988)	X					

Fonte: Própria

Quanto à bibliografia, existe muita diversidade. Algumas faculdades como a Katiavala têm 50% da bibliografia sugerida ao estudante em Inglês, o que dificulta a pesquisa por parte dos estudantes, visto que muitos deles não dominam a língua inglesa, para além de existirem estudantes que no ensino médio optaram pela língua francesa e que agora se deparam com a bibliografia inglesa. No entanto, toda bibliografia sugerida está diretamente ligada ao conteúdo econométrico a lecionar. Por sua vez o Instituto Superior Politécnico Lusíadas assim como a Universidade Katiavala Bwila tem todo conteúdo virado mais à econometria como tal e, ao contrário da Katiavala, que está em fase de organização nas suas novas instalações e não possui biblioteca nem uma livraria para que os estudantes de alguma forma tenham acesso à bibliografia sugerida. O Instituto Superior Politécnico Lusíadas tem toda a bibliografia em português e uma livraria onde é possível encontrar ou encomendar todos os livros sugeridos, para além da biblioteca da instituição, o que facilita os estudantes nas pesquisas que têm a fazer. Por sua vez o Instituto Superior Politécnico Jean Piaget, tem uma bibliografia mais virada à estatística, com apenas uma pequena parte voltada diretamente ao ensino da econometria e como tal, uma pequeníssima parte voltada ao ensino da Matemática, como ocorre no curso de Economia e Gestão. Aqui encontramos um pouco de tudo, ou seja, encontra-se uma bibliografia que mostra a interligação entre as ciências Matemáticas, Estatística e Econometria. Já a diversificação da bibliografia é reduzida no curso de Engenharia Informática de Gestão que nos apresenta apenas duas sugestões para tal, limitando a possibilidade de mais opiniões em relação aos conhecimentos econométricos. No entanto também possui uma livraria onde se podem encontrar tais livros e o estudante pode adquirir a bibliografia recomendada, bem como biblioteca bem apetrechada com literatura diversificada.

Tabela VI- Programas da disciplina de Econometria das distintas Universidades/Instituições

Conteúdo programático por capítulos	Universidades / Institutos					
	Katiavala Bwila		I.S.P. Lusíadas		I.S.P. Jean Piaget	
	Economia	Gestão de Empresas	Economia	Gestão de Empresas	Economia de Gestão	Informática de Gestão
Análise de regressão de duas variáveis					X	
Análise de regressão múltipla					X	
Apresentação dos resultados de regressão	X					
Autocorrelação			X			
Coeficiente de determinação quadrado (R^2)	X					
Equações simultâneas	X					
Estudo formal do processo econométrico de aplicações de modelos			X	X		
Formas Funcionais de modelos econométricos	X					
Heterocedasticidade			X			
Identificação	X					
Intervalos de confiança para parâmetros	X					
Introdução	X					
Introdução a teoria de séries						X

Modelo linear geral			X			
Modelo linear geral e variável binária				X		
Modelos explicativos dinâmicos						X
Multicolinearidade			X			
Natureza da análise de regressão					X	
O modelo de regressão simples	X					
Problemas de estimação					X	
Procedimentos econométricos						X
Reflexões metodológicas						X
Regressão múltipla-extensões	X					
Teste de hipóteses	X					
Uso de determinantes para resolver sistemas de equações	X					
Variáveis especiais			X	X		
Violação dos pressupostos básicos e soluções	X					

Fonte: Própria

Embora os cursos tenham nomes diferentes, em termos de conteúdo estes não diferem muito, naquilo que se apresenta ao estudante nas distintas disciplinas, apenas se vê uma diferença na organização temática dos assuntos a tratar de uma instituição para outra. Verifica-se, no entanto, que para o ensino da econometria existem unidades temáticas que numa universidade são analisados como um assunto dentro de uma unidade temática enquanto que na outra representa já uma unidade temática como é, por exemplo, o que acontece com o tema: “*Violação de pressupostos básicos e soluções*”, do curso de Economia da Katiavala Bwila onde se fala dos testes mais usuais no estudo da heteroscedasticidade, multicolinearidade e autocorrelação; enquanto que a Lusíada trata cada um destes elementos (heteroscedasticidade, multicolinearidade e autocorrelação) como uma unidade temática, onde estuda todos os fatores que intervêm na sua existência como tal. Também existem temas com nomes diferentes à primeira vista e que tratam do mesmo assunto a partir de diferentes pontos de vista como é o caso dos temas; “*Introdução*” da universidade Katiavala e da Piaget dos cursos de Economia e Economia e Gestão, respectivamente, e “*Estudo formal do processo económico de aplicação de modelos*” da Lusíada nos cursos de Economia e Gestão de Empresas, que tratam de assuntos como a definição de econometria, sua importância, sua classificação, seus objetivos entre outros aspetos básicos. Ainda para o ensino da econometria torna-se importante conhecer o conteúdo Matemático programado para estes cursos. Desta forma vemos que todas as universidades começam por rever aquilo que foi o conteúdo lecionado no ensino médio, aprofundando um pouco mais este ou aquele assunto. Vemos também que o conteúdo programático das diversas faculdades em relação à disciplina de Análise Matemática é o mesmo, diferindo apenas na questão organizacional dos conteúdos em alguns casos. A seu modo vemos que cada faculdade retira da Matemática conteúdos de álgebra e geometria que têm aplicação direta nas disciplinas económicas, como é o caso do estudo e análise da função linear, tendo em conta as suas representações gráficas e o estudo completo de uma função, retratando a vida económica no uso de matrizes, buscando os determinantes que ajudam de certa forma na compreensão e aplicação dos modelos económicos, facilitando a análise de resultados. O integral de Riemann

necessário na análise de resultados das experiências económicas. No curso de Engenharia Informática de Gestão faz-se um estudo detalhado daquilo que é a disciplina de Matemática. Aqui se estuda a disciplina nas várias vertentes do saber matemático, tornando claro ao estudante as áreas deste saber que estão diretamente ligadas à estatística e econometria, bem como a interligação desta com as várias ciências assim como a física e outras. O estudante do curso de Informática de Gestão termina com um amplo conhecimento daquilo que é a Matemática, como dividi-la em secções e como ligá-la a si mesma. Ainda consegue ver claramente a ampla aplicação da Matemática na vida prática do dia a dia. Por sua vez, o conteúdo estatístico aqui apresentado tem uma profunda aplicação à disciplina de econometria, visto que se fala de regressão linear e correlação de variáveis. Apesar de que o Instituto Superior Politécnico Lusíada de Angola foque muito ainda a descrição das variáveis aleatórias e as probabilidades que não são menos importantes para a compreensão rápida e clara daquilo que são os modelos econométricos e suas aplicações nas atividades económicas. O conteúdo estatístico transmitido é aquele que serve de base ou que nos apresenta as ferramentas necessárias para o tratamento da econometria como tal. Temos aqui elementos como, conceito de variável aleatória e outros para a compreensão da aleatoriedade em econometria, o conceito de teste de hipóteses que ajuda na compreensão da aplicação de hipóteses, Modelo de Regressão Linear que nos dá uma ideia de como é e será trabalhar com modelos deste tipo, o Método dos Mínimos Quadrados muito utilizado em econometria para análise de erros nos modelos econométricos, os Erros de primeira e segunda espécie que nos ajudam a diferenciar os tipos de erros encontrados nos modelos econométricos usados, bem como os mecanismos de superá-los. Com isto vemos que a probabilidade e a estatística são fundamentais para o ensino da econometria por introduzirem os elementos primordiais da econometria como as variáveis aleatórias, as funções de distribuição, modelos lineares simples, covariância e correlação, intervalos de confiança e testes de hipóteses, que compõem as bases econométricas para aplicação de modelos econométricos, proporcionando ao econometrista as ferramentas de trabalho seguro para as questões que se lhe apresentam, a fim de obter na sua análise econométrica resultados fiáveis.

Tabela VII – Conteúdo Matemático, Estatístico e Econométrico programado para os cursos de Economia e Gestão de Empresas.

DISCIPLINA		UNIVERSIDADES		
		KATIAVALA BWILA	I.S.P. LUSIADAS	I.S.P. JEAN PIAGET
MATEMATICA	ANALISE MATEMATICA I	<p>I. Sucessões: noções topológicas; definição de sucessão; limite de uma sucessão; critérios de convergência e de divergência.</p> <p>II. Limite de uma função real de variável real: definição; limites laterais; propriedades e cálculo dos limites; número de Néper.</p> <p>III. Funções contínuas e descontínuas em \mathbb{R}: definição de continuidade; continuidade lateral; descontinuidade.</p>	<p>I. Álgebra linear: matrizes ou cálculo matricial (generalidades e definições; álgebra de matrizes; inversa de uma matriz; matriz transposta); determinantes(definições; regras de cálculo; teorema de Laplace; regra de Cramer de Sarrus; propriedades; matriz adjunta; menor complementar, complement algébrico; valores e vectores próprios de uma matriz, diagonalização; equações matriciais); sistemas de equações lineares(representação matricial de sistemas de equações; método de matriz inversa da resolução de sistemas das equações; regra de Cramer para resolução de sistemas de equações lineares; método de Gauss; sistemas homogénios.</p> <p>II. Cálculo diferencial em \mathbb{R} e em \mathbb{R}^n: Funções (definições, apresentação analítica, de tabela e gráfica, domínios e contradomínios;funções reais de variável real, gráficos e análise; funções trigonométricas e inversas trigonométricas, gráficos e análise; noções de limites e continuidade de uma função); derivadas (derivada de uma função real de variável real, definição e interpretação; principais fórmulas e regras de derivação; derivação de funções compostas e implícitas; derivação logarítmica, de funções apresentadas na forma paramétrica, e de funções inversas; derivadas de diferentes ordens); aplicações das derivadas (teorema de Lagrange. Aplicações; breves noções sobre séries de potências; breve referência as noções de limite e continuidade num ponto; problemas de máximos e mínimos; aplicação em economia)</p>	<p>I. Axiomática dos números reais</p> <p>II. Sucessões geométricas e aritméticas.</p> <p>III. Noção de limite, teorema de Bolzano-Weierstrass.</p> <p>IV. Séries numéricas: critérios de comparação de D'Alembert e de Cauchy; séries alternadas e de potências.</p> <p>V. Funções reais de variável real: continuidade de limite; continuidade global; teorema do valor médio e de Weierstrass; levantamento de indeterminações.</p> <p>VI. Diferenciabilidade: definição; teorema de diferenciabilidade; derivada da função composta; derivadas parciais de ordem superiorà primeira; teorema de Schwarz.</p> <p>VII. Aplicações: estudo completo de uma função; levantamento de indeterminações pelo teorema de L'Hospital.</p> <p>VIII. Primitivação: métodos gerais; primitivação de funções racionais; condições de integrabilidade; integração por substituição, por partes; integração de funções irracionais, trigonométricas e racionais.</p> <p>IX. Integrais definidos</p> <p>X. Noção de matriz: tipos de matrizes; propriedades; cálculo da matriz inversa; soma e multiplicação de marizes; sistemas de equações lineares; resolução de sistemas de equações; método de Cramer e de Gauss.</p>

	ANALISE MATEMATICA II	<p>I. Derivadas: definição; derivadas laterais; regras de derivação; derivadas das funções compostas; derivadas das funções logarítmicas e exponenciais; derivadas de ordem superior; derivadas parciais; extremos de uma função contínua; aplicações económicas das derivadas.</p> <p>II. Primitivas: Integração indefinida (noção e definição; Integração imediata; Integração por mudança de variável; Integração por partes); Integração definida (definição; fórmula de Newton-leibniz; propriedades; cálculo das coordenadas de figuras planas); Integração imprópria (definição; propriedades).</p> <p>III. Séries: definição; séries numéricas não alternativas (definição; condições de convergência e divergência; critérios de convergência); séries numéricas alternadas (definição; critério de Leibniz).</p>	<p>I. Cálculo integral: primitivação de uma função real de variável real; método imediato de primitivação; integral indefinido (método imediato de integração, de integração por partes, por substituição e da decomposição; integrais de funções irracionais, trigonométricas e exponenciais); integral de rieman para funções reais de variável real. Definição. Propriedades; cálculo de áreas de regiões limitadas do plano; integrais impróprios. Cálculo de áreas de regiões não limitadas do plano.</p> <p>II. Funções de várias variáveis e integrais duplos: funções de várias variáveis; derivadas parciais; integrais múltiplos; integrais duplos (definição, propriedades e cálculos nos casos simples); mudança de variável. Cálculo; aplicação de integrais duplos.</p> <p>III. Equações diferenciáveis: equações diferenciáveis não lineares de 1ª ordem (variáveis separadas, homogéneas, exatas, fatores integrantes); equações diferenciais lineares (1ª ordem, 2ª ordem com coeficientes constantes, estabilidade)</p> <p>IV. Breve introdução à análise numérica: interpolação polinomial; reta dos quadrados mínimos. Aplicações.</p>	
--	-----------------------	--	--	--

ESTATÍSTICA	ESTATÍSTICA I	<p>I. Estatística descritiva: conceitos fundamentais (definição de estatística; população, amostra, atributo, atributo qualitativo e quantitativo, variável, variável discreta e contínua, elemento e dado estatístico, nível de medida dos dados e variáveis); distribuição de frequências (definição. Conceitos de modalidade, frequência absoluta, frequência relativa, frequência acumulada absoluta e relativa; apresentação dos dados. Distribuição de frequências de variáveis discretas. Diagrama de barras; apresentação dos dados. Distribuição de frequências de variáveis contínuas; histograma de frequências. Polígono de frequências); medidas de estatística descritiva (medidas de tendência central: média aritmética, mediana e moda; medidas de tendência não central: quartis, decis e percentis; medidas de dispersão: desvio absoluto médio, variância, desvio padrão e coeficiente de variação; medidas de concentração: curva de Lorenz e índice de Gini; medidas de assimetria; medidas de curtose).</p> <p>II. Probabilidades: revisões (álgebra dos conjuntos; análise combinatória. Técnicas de contagem. Princípio de multiplicação. Combinações, permutações e arranjos); experiências aleatórias. Resultados. Acontecimentos; conceitos de probabilidades [conceito clássico de probabilidades; conceito frequentista (estatística); conceito subjetivo; axiomas e principais teoremas de probabilidades]; probabilidade condicionada (definição; fórmula de probabilidade total; fórmula de Bayes); independência.</p> <p>III. Variáveis aleatórias: noção de variável aleatória. Definição. Função de distribuição. Propriedades; variáveis aleatórias unidimensionais (variáveis aleatórias discretas. Função de probabilidade; variáveis aleatórias contínuas. Função de densidade); variáveis aleatórias bidimensionais (definição. Função de distribuição conjunta. Propriedades; variáveis aleatórias discretas. Função de probabilidade conjunta. Distribuição marginal. Distribuições condicionais; variáveis aleatórias contínuas. Função de densidade conjunta. Distribuições marginais. Distribuições condicionais; Independência de variáveis).</p> <p>IV. Parâmetros das variáveis aleatórias: parâmetros de localização (de tendência central. Valor esperado. Mediana. Moda; quantis: quartis, decis, percentis); momentos. Momentos ordinários iniciais e centrados; parâmetros de dispersão. Variância, desvio padrão e coeficiente de variação; parâmetros de associação. Covariância e coeficiente de correlação.</p> <p>V. principais distribuições teóricas discretas: distribuição uniforme; distribuição binomial (distribuição de Bernoulli); distribuição de Poisson; distribuição Hipergeométrica.</p> <p>VI. Principais distribuições teóricas contínuas: distribuição uniforme; distribuição exponencial; distribuição normal (ou de Gauss).</p> <p>VII. Resultados assintóticos: lei dos grandes números (desigualdade de Tchebichev; lei fraca dos grandes números; lei forte dos grandes números); teorema do limite central. Definição. Aproximação normal das distribuições binomial e de Poisson.</p>	<p>I. Estatística descritiva: conceitos; distribuições de frequências; representações gráficas; medidas de localização, dispersão e de assimetria; Box-Plot.</p> <p>II. Teoria das probabilidades: conceitos (experiência aleatória, espaço de resultados, acontecimentos); definição de probabilidade (clássica, frequentista e axiomática); probabilidade condicional; acontecimentos independentes; teorema da probabilidade total; teorema de Bayes.</p> <p>III. Variáveis aleatórias discretas: conceito de variável aleatória; função massa de probabilidade; função de distribuição; parâmetros de variáveis aleatórias; distribuições de probabilidade discreta (prova de Bernoulli; distribuição binomial e de Poisson).</p> <p>IV. Variáveis aleatórias contínuas: conceito; função densidade de probabilidade; função de distribuição; parâmetros de v.a.s. contínuas; distribuições de probabilidade contínuas (distribuição uniforme e normal).</p>	<p>I. Introdução: nota histórica; conceitos gerais (definição de estatística, bioestatística, população, amostra, dados e variáveis); objetivos do estudo da bioestatística.</p> <p>II. Tipos de estatística: descritiva e inferencial.</p> <p>III. Variável: tipos de variáveis; níveis de medida de uma variável.</p> <p>IV. Estatística descritiva: apresentações estatísticas; apresentação tabular dos dados (frequências simples e agrupadas; apresentação gráfica dos dados; histogramas e polígonos de frequências, curvas de frequências, diagrama de barras, gráficos de linhas, gráficos circulares (tartes)); descrição de dados “características amostrais”, medidas de tendência central “localização”. Média, moda e mediana (medidas de variabilidade “dispersão”, amplitude, desvio padrão, variância); teorema elementar da probabilidade (noções básicas de probabilidades; probabilidades condicionais; teorema de probabilidade total (teorema de Bayes)); amostragem, etapas de um levantamento por amostragem, tipos de amostragem (amostragem não probabilística e probabilística); distribuições estatísticas [distribuições de variável discreta (Binomial e Poisson); distribuições de variável contínua (normal)].</p> <p>V. Estatística inferencial (indutiva): distribuições amostrais e teorema do limite central; estimação de parâmetros populacionais (estimação pontual; estimação por intervalos de confiança; uso da distribuição t de Student); teste de hipóteses [contraste de hipóteses em distribuições normais; comparação de duas médias (teste t), homocedasticidade, heteroscedasticidade, observações emparelhadas]; números índices.</p>
-------------	---------------	--	--	--

	<p>ESTATÍSTICA II</p>	<p>I. Teoria de estimação: noção de inferência estatística (tipos de amostragem; distribuições amostrais. Distribuição normal, qui-quadrado, student e Fischer. Distribuições amostrais de médias, variâncias, proporções, diferenças de médias, diferenças de proporções e razão de variâncias); estimação pontual de parâmetros (definições. Estimação, estimador e estimativa; propriedades de estimadores); estimação de parâmetros por intervalos (definições. Intervalos de confiança. Princípios; intervalos de confiança para médias, variâncias e proporções; intervalos de confiança para diferença de duas médias; intervalos de confiança para diferença de duas proporções; intervalos de confiança para diferença de duas variâncias).</p> <p>II. Testes de hipóteses: conceitos básicos [hipótese estatística. Hipótese nula e alternada; nível de significância; erros do tipo I e do tipo II; critério estatístico; região de aceitação e região crítica (rejeição)]; princípio. Passos de um teste de hipótese (formulação das hipóteses e estabelecimento do nível de significância; definição do teste. Teste bilateral e teste unilateral; cálculo dos pontos críticos do critério estatístico, estabelecimento das regiões críticas e de aceitação e definição das regras de decisão; cálculo do valor do teste; comparação e tomada de decisão); testes paramétricos (testes para a média de uma população (com variância populacional conhecida e desconhecida); teste para a proporção; teste para variância uma população normal; teste para a diferença de duas médias populacionais (com variâncias populacionais conhecidas e desconhecidas); teste para a diferença de duas proporções; teste para a razão de duas variâncias); testes não paramétricos (teste qui-quadrado de ajustamento e de independência).</p> <p>III. Análise de variância (ANOVA): conceitos básicos. Princípios; análise de variância simples [soma de quadrados (entre e dentro dos grupos); graus de liberdade (entre e dentro dos grupos); variâncias entre e dentro dos grupos; valor do teste; comparação e tomada de decisão; quadro de análise de variância].</p> <p>IV. Correlação e regressão: diagramas de dispersão; regressão linear simples. O método dos mínimos quadrados; correlação linear; coeficiente de determinação e análise de variância.</p> <p>V. Análise das séries temporais: introdução (definição; componentes de uma série temporal; métodos de decomposição de uma série temporal); análise da tendência (método dos mínimos quadrados e das médias móveis); análise da componente sazonal [método da percentagem média; cálculo dos índices sazonais (num modelo de decomposição aditivo e num modelo de decomposição multiplicativo)]; análise da componente cíclica.</p> <p>VI. Números índices: introdução (definição; tipos de números índices); índices simples. Propriedades; índices compostos ou agregados; índices compósitos ou sintéticos; manipulação dos números índices; deflação de séries temporais [o índice de preços no consumidor (IPC); alguns problemas na determinação do UPC].</p>	<p>I. Estimação pontual: estatísticas; estimador e estimativa; propriedades dos estimadores.</p> <p>II. Estimação por intervalos de confiança: noção de intervalos de confiança; cálculo de intervalos de confiança; intervalos de confiança para médias de normais; intervalos de confiança para proporções.</p> <p>III. Testes de hipóteses: conceito de teste de hipóteses; conceito de estatística de teste e regra de decisão; erros de primeira e segunda espécie; conceito de potência do teste.</p> <p>IV. Regressão linear e correlação simples: introdução; modelo de regressão linear; método dos mínimos quadrados; coeficiente de determinação e correlação.</p>	<p>I. Análise de variância (ANOVA): anova a um critério de classificação; anova a dois critérios de classificação; anova multifactorial; comparações múltiplas (Tukey, Bonferony, Newman Keuls).</p> <p>II. Análise de correlação e de regressão: correlação e regressão simples e múltipla.</p> <p>III. Teste de qui-quadrado (tabelas de contingência)</p>
--	-----------------------	--	---	--

ECONOMETRIA	ECONOMETRIA I	<p>I. Introdução: regressão simples (relação determinista; relação estocástica); regressão múltipla (pressupostos básicos; derivação de equações normais necessárias para calcular os coeficientes de uma função); uso de determinantes para resolver sistemas de equações (a regra de Cramer; o método do inverso; o método pivotal ou de eliminação de Gauss; outros métodos).</p> <p>II. Casos aplicados: testes de hipóteses (T-teste; F-teste); coeficiente de determinação quadrado (R^2); intervalos de confiança para os parâmetros; apresentação dos resultados da regressão; violação dos pressupostos básicos e soluções (-heterocedasticidade: o uso do teste de Breusch Pagan, de Goldfeld-Quandt, e do teste de White, uso de gráficos, de resíduos e dummies, etc; - teste de ausência de auto-regressão: a estatística de Durbin-Watson e outros métodos “Durbinh, etc”; soluções possíveis: transformação da equação em causa, ou transformação através de logaritmos; - perturbações auto: regressivas, estimação interativa de duas etapas “biépatica”, método de Durbin, o uso das primeiras diferenças; -testes de multicolinearidade: o uso da matriz de correlação; o uso do determinante. Soluções possíveis: identificação da causalidade e sua correção, etc.</p> <p>III. Formas funcionais de modelos econométricos: regressão múltipla-extensões (o modelo de Logit e a sua estimação; variáveis independentes não quantificáveis “dummies de 0,1; dummies de 0, 1, 2, etc”); equações simultâneas [caso de correlação entre variáveis independentes e o erro; casos de erros nas ambas variáveis (dependente e independente)]; identificação (condições necessárias e condições suficientes); outros e diversos [métodos avançados de estimação (e.g. MLE); omissão de uma variável independente importante; etc].</p>	<p>I. Estudo formal do processo econométrico de aplicação de modelos: conceitos e objetivos da econometria; conceito de modelo (classificação de modelos em economia; modelos econométricos; pesquisa econômica e uso de modelos econométricos); especificação de modelos (conceitos de especificação; fontes de informações para a especificação de modelos; restrições à construção de modelos sem teoria; formas funcionais linealizáveis; critérios de escolha da forma funcional).</p> <p>II. O modelo de regressão simples: o modelo linear simples (pressupostos do modelo; método de estimação de modelos; método dos mínimos quadrados; exemplos numéricos da aplicação do modelo linear simples “avaliação de modelos estimados; caracterização de uma variável aleatória; qualidade desejáveis dos estimadores; critérios de avaliação de estimativas de modelos; média, variância dos parâmetros estimados; decomposição da soma de quadrados; variância amostral ou residual; coeficiente de determinação; estatística F; estatística t; testes de hipóteses; intervalos de confiança; correlação parcial”).</p> <p>III. Modelo linear geral: modelo linear geral [pressupostos básicos; estimação dos parâmetros do modelo; estatística de avaliação; análise de regressão e correlação parcial; exemplo numérico da aplicação do modelo linear geral e de correlação parcial (variância amostral ou residual; coeficiente de determinação; estatística F; estatística t); teste de hipóteses; intervalos de confiança; correlação].</p>	<p>I. Introdução: econometria; metodologia da econometria; tipos de econometria; o papel do computador.</p> <p>II. Natureza da análise de regressão: origem histórica do termo regressão; interpretação moderna da regressão; regressão versus causalidade; regressão versus correlação.</p> <p>III. Análise de regressão de duas variáveis: conceito de função de regressão populacional (FRP); significado do termo linear; significado do termo de perturbações; conceito de função de regressão amostral (FRA).</p> <p>IV. Problema da estimativa: método dos mínimos quadrados ordinários; o modelo clássico de regressão linear; as hipóteses subjacentes ao método dos mínimos quadrados; precisão ou erros-padrão das estimativas por mínimos quadrados; propriedades dos estimadores de mínimos quadrados (teorema de Gauss-Markov).</p> <p>V. Análise de regressão múltipla: o modelo de três variáveis (noção e hipóteses); interpretação da equação da regressão múltipla; significado dos coeficientes de regressão parcial; estimativa por MQO dos coeficientes de regressão parcial.</p>
	ECONOMETRIA II		<p>I. Modelo linear geral: aplicação do modelo linear geral (forma linear; forma logarítmica ou potencial; forma exponencial ou semilogarítmica I; forma hiperbólica ou recíproca I; forma quadrática).</p> <p>II. Variáveis especiais: variável tempo (o tempo como variável explicativa); variáveis binárias ou dummies (conceito e objetivos; como incorporar variáveis binárias ao modelo; aplicação do modelo linear geral com utilização de variáveis dummies).</p> <p>III. Multicolinearidade: conceito e consequência; casos de multicolinearidade; diagnóstico da multicolinearidade; tratamento da multicolinearidade; aplicação de testes de multicolinearidade.</p> <p>IV. Autocorrelação serial: conceito e consequências; fontes de autocorrelação; diagnóstico de autocorrelação; testes e correção do problema da autocorrelação serial; aplicação do teste de Durbin-Watson; aplicação do teste de Godfrey; correção pelo método iterativo de Cochrane-Orcutt.</p> <p>V. Heterocedasticidade: conceitos e consequências; pressupostos sobre a natureza da variância dos resíduos e correção da heteroscedasticidade; testes e correção da heteroscedasticidade; aplicação do teste de Quandt-Goldfeld; aplicação do teste de Glejser; aplicação do teste de Park.</p>	

Do exposto na tabela VI (acima) podemos ver que não podemos falar de Econometria sem antes buscarmos bases Matemáticas e Estatísticas para a fundamentarmos. Da Matemática, a Econometria busca noções de condição, variável independente e dependente, de representação gráfica bem como sua interpretação, retira o conceito de função entre outros; já da Estatística busca as noções de variância, desvio padrão, covariância, erro, hipótese e muito mais. No entanto, estes conteúdos lecionados nos cursos de Economia e de Gestão de Empresas apresentam várias diferenças de uma instituição para outra no que diz respeito à sua organização por unidades temáticas (capítulo). O teor apresentado aos estudantes é até certo ponto semelhante, mas há uma necessidade de unificá-los tanto em unidades temáticas ou capítulo como nos assuntos tratados em cada unidade temática, visto que os conteúdos de cada disciplina são vastos, os docentes têm encontrado dificuldades no cumprimento dos programas e a unificação dos conteúdos requereria que se fizesse um corte vertical no conteúdo até então programado. Quanto à Econometria que é lecionada apenas no primeiro semestre do terceiro ano curricular dos cursos de economia em todas as instituições, tem sido difícil para os docentes cumprirem com todo o programa o que mostra que devido à sua importância na formação de um gestor deveria se aumentar a carga horária desta disciplina para que se cumpra com o programado, ocupando os dois semestres do terceiro ano curricular e preparando assim da melhor forma os seus gestores. Com isto é necessário que os conteúdos econométricos lecionados nas distintas instituições superiores sejam os mesmos, isto é, tenham a mesma organização e que comecem todos com uma unidade introdutória onde se apresentem a importância, o historial, os objetivos da disciplina, entre outros aspectos. Também é importante que as unidades temáticas estejam organizadas da melhor forma possível de tal modo que o conteúdo seja apresentado ao estudante de forma gradual, partindo do simples para o complexo, focando os elementos fundamentais na aplicação laboral dos conhecimentos econométricos. Para uma melhor formação de um gestor visando excelentes resultados no local de trabalho é necessário que haja mais debates e colóquios dirigidos aos estudantes com vista a conhecerem os demais métodos econométricos aplicáveis em diferentes situações, bem como as

instituições investirem mais na aquisição de pacotes informáticos que facilitem a aplicação de modelos econométricos mais complexos uma vez que os modelos ensinados aos estudantes são os modelos de regressão linear simples e o múltiplo (com três variáveis) devido à falta de laboratórios informáticos equipados com softwares que facilitem o estudo de casos mais complexos, uma vez que no estudo de modelos econométricos de uma forma geral utiliza-se apenas o software Excel. Deve-se criar nos estudantes o hábito de leitura, visto que os estudantes apresentam uma limitação na interpretação económica dos problemas econométricos que se lhes apresentam, e no ensino da Matemática e Estatística enfatizar a importância de se interpretar todos os gráficos estudados bem como a aplicabilidade dos casos estudados.

Sendo o gestor uma alavanca para o crescimento de uma empresa e visando o crescimento económico da região académica em estudo, é urgente que as instituições superiores da região reestruturem os seus programas de tal forma que o gestor vindo de qualquer uma das instituições esteja à altura de responder às ansiedades do mercado de trabalho que lhe espera.

O Ministério do Ensino Superior deve controlar não só o surgimento anárquico de novas instituições, como deve criar um departamento que analise os programas sugeridos tendo em vista uma formação que leve o país a um desenvolvimento mais amplo olhando para a qualificação profissional dos cidadãos.

4. Resultados

4.1 Resultados da revisão bibliográfica

A econometria tem um papel importante na obtenção de dados fiáveis para as operações económicas atuais, devido ao facto das várias ações das suas atividades serem analisadas em um curto espaço de tempo.

Os modelos econométricos são as vias usadas pelos econometristas para analisar os resultados das suas atividades sem, contudo, violar as regras económicas, através de cálculos de algumas funções matemáticas com suporte estatístico, para que se alcancem melhores resultados na análise das atividades económicas em estudo.

O modelo econométrico do tipo determinístico ou exato é mais geral e muito pouco utilizado em econometria por retirar a hipótese do erro, visto que a Economia é uma ciência social, os seus resultados variam de acordo a situação em estudo. O modelo exato não é o mais indicado para o estudo das medições económicas aplicadas à Econometria, por analisar as variáveis em estudo em grupos tendo em conta apenas o momento de sua distribuição.

Nas fases de elaboração de um modelo até o momento de sua aplicação vemos que existe um conjunto de etapas, a percorrer como a de especificação do modelo (selecionam-se as variáveis explicativas da variável explicada definindo a forma funcional da relação entre a variável explicada e as explicativas de forma correta); a etapa de seleção da amostra (recolhem-se os dados da amostra que podem ser seccionais ou transversais, temporais ou cronológicas e ainda longitudinais ou de painel); a etapa de estimação ou ajustamento do modelo (o econometrista produz valores numéricos constantes e concretos que se aproximem aos valores ignorados, recorrendo a métodos com propriedades desejáveis); a etapa de avaliação dos resultados (o coeficiente de determinação indica-nos a qualidade de ajustamento, possibilitando a estimação do valor da variável explicada para cada observação, valores estes que ajustados para a variável explicada coincidem com os observados) e a etapa de utilização dos resultados (averigua-se a confirmação dos dados, especificando o modelo estimado e

validado estatisticamente, pode servir de base a antecipações de conhecimentos futuros ou diferentes).

Um modelo económico é econométrico e aplicável, se as fases citadas estiverem bem identificadas, os resultados obtidos estejam corretos e seja aplicável a situações pesquisadas.

Os modelos económicos servem para análise de estrutura, simulações, previsões e constatação da teoria económica. Ao serem aplicados em empresas para um ajuste de preços melhora os investimentos e, na previsão das vendas futuras, resulta uma optimização das vendas.

Por outro lado não existe estatística em econometria que não se apoie na teoria económica como guia. A Economia Matemática é a teoria mais básica para tal estudo econométrico, por conter as ferramentas matemáticas necessárias para desenvolver um estudo estatístico e econométrico.

4.2 Resultados da entrevista aos coordenadores e aos programas

Da análise feita aos programas e da entrevista aos coordenadores de disciplina, vê-se que: a Econometria é uma disciplina fundamental para a formação legal de um gestor. No entanto, para que se leccione a econometria é fundamental que os estudantes tenham bases matemáticas e estatísticas, o que faz com que esta seja apresentada ao estudante no terceiro ano curricular, já que no 1º e 2º anos estes adquirem conhecimentos matemáticos e estatísticos que servem de base sólida aos conhecimentos econométricos por adquirir. Todas as instituições superiores lecionam a disciplina de Econometria no primeiro semestre do 3º ano curricular e apenas uma subdivide-a nos dois semestres deste ano curricular como econometria I e econometria II. Querendo ter um gestor dotado de conhecimentos econométricos mais amplos e profundos no que diz respeito aos modelos econométricos tendo em conta a sua importante influência na análise e tomada de decisão do gestor nas diversas experiências económicas que se lhe apresentam, todas as instituições deveriam ter a econometria como uma disciplina anual e não semestral. Vemos que o conteúdo sobre *modelos de regressão simples* também conhecidos como *modelo de regressão linear*

simples é um exemplo de que existem conteúdos que para uma instituição é uma unidade temática enquanto que para outra instituição parece como um subtema de uma unidade temática, mostrando que de uma forma geral o mesmo conteúdo pode ser tratado de diferentes maneiras sem alterar os seus aspectos primordiais. A vantagem é de que se o assunto é tratado como unidade temática, há mais possibilidades de se avaliar todos os detalhes e pormenores importantes no estudo deste tema. Já na questão dos componentes que indicam a existência de problemas na estimação de dado modelo que são a heterocedasticidade, colinearidade e autocorrelação algumas instituições estudam cada um destes aspectos em uma unidade temática enquanto outras os estudam como subtemas de uma unidade temática. Uma vez que os modelos econométricos mais apresentados aos estudantes são os de regressão linear simples e múltipla, os problemas que surgem na análise de uma regressão devem ser bem estudados, passando por todos os seus aspetos com muita paciência e cautela para que o estudante os possa melhor distinguir e corrigir.

Verificou-se que o curso de Gestão Empresarial é ministrado como uma especialidade do curso de Economia, e que estes se dissociam em alguns casos no terceiro ano curricular e os estudantes neste caso acabam ficando sem uma formação em econometria e em outros casos se dissociam no quarto ano curricular e os estudantes são formados com conhecimentos econométricos. Este facto faz-me sugerir que as instituições superiores criem mecanismos que permitam esta dissociação apenas no quarto ano curricular, porque o gestor torna-se mais completo se possuir conhecimentos econométricos. Os objectivos traçados para esta cadeira de uma forma geral visam transmitir da melhor forma os conhecimentos ligados aos modelos econométricos principalmente os de regressão linear simples que é o mais aplicado durante o processo curricular e que pode ser aplicado a partir do programa informático do Excel. Não existe falta de bibliográfica para análise e pesquisa dos estudantes. O que se tem verificado é que os estudantes não tem uma cultura de investigação, limitam-se às informações oferecidas pelo seu docente o que não é suficiente, para uma boa formação. O estudante deve procurar comprovar os conhecimentos a ele transmitidos a partir de diferentes pontos de vista. A maioria destas instituições superiores oferece

vasta bibliografia em português de diferentes autores e origens, facilitando de certa forma a diversificação dos conhecimentos científicos neste campo do saber. No entanto, de uma forma geral, o aproveitamento dos estudantes tem estado ao mesmo nível nas distintas instituições do ensino superior. As dificuldades na aprendizagem assemelham-se muito, pelo facto de na sua maioria os estudantes não terem uma base sólida no campo da Matemática e Estatística, disciplinas que servem de alicerce para o ensino dos conteúdos econométricos; outro indicativo de dificuldades no aprendizado é a falta dos programas de softwares próprios para o ensino dos modelos econométricos limitando o campo de ação do docente e do estudante, uma vez que os conhecimentos práticos diferem consideravelmente dos conhecimentos teóricos, ou seja, ouvir falar de regressão linear múltipla difere muito de ter meios para aplicar às fórmulas estas regressões, vendo na prática como identificar as variáveis. Dificultando assim a compreensão do que vem a ser, na prática, a utilização destes meios na vida ativa e como estes contribuem para tornar mais simples e prático o seu trabalho.

5. Conclusões

Feito o estudo da aplicação dos modelos econométricos no ensino superior angolano, no caso da região académica nº 2, concluiu-se que:

A Econometria vem para simplificar ainda mais a vida de um gestor tornando os seus resultados mais precisos e fiáveis, reduzindo consideravelmente os riscos de se obter resultados duvidosos na análise de dados nos diversos estudos realizados.

O modelo econométrico do tipo estocástico é mais específico que o modelo determinístico, por quantificar as variáveis de dada experiência económica tendo em conta os conceitos da teoria económica, buscando como base conceitos matemáticos como o da função linear, permitindo uma comparação da ação das variáveis intervenientes na experiência a realizar para garantir a veracidade dos resultados.

Existe uma série de etapas importantes a percorrer na elaboração de um modelo econométrico até a fase de sua aplicação eficaz numa dada experiência económica.

No estudo de modelos econométricos surgem problemas como a multicolinearidade, a heterocedasticidade e a autocorrelação, que comprometem a confiabilidade dos valores dos coeficientes estimados e a inferência estatística. No entanto, existem formas específicas para contornar estes problemas.

Existe uma gritante necessidade de se unificar os programas de Econometria de tal modo que as unidades temáticas venham a retratar os conteúdos ao mesmo grau ou na mesma intensidade.

Para a realização desta pesquisa colocaram-se as seguintes questões: - Será que os estudantes têm recebido bagagem suficiente para contribuírem positivamente no desenvolvimento económico da região? Como se têm preparado para encarar as dificuldades que o mercado lhes apresenta? Da análise feita aos programas e da entrevista aos coordenadores de disciplina concluiu-se que, os estudantes têm recebido bases econometricas suficientes para o cabal cumprimento de suas tarefas laborais contribuindo assim para o desenvolvimento da região, uma vez que, os docentes procuram apresentar a estes todas as bases necessarias para a gestão de uma empresa pequena ou de médio porte, que corresponde mais a realidade da região.

Quanto aos objectivos da investigação concluiu-se que:

- 1- O conteúdo econometrico apresentado aos estudantes dos cursos superiores de gestão está bem delimitado e contém todas as componentes necessarias para a formação de um gestor capaz de responder as ansiedades do mercado económico-financeiro da região. Tal conteúdo foi cuidadosamente selecionado procurando dar ao estudante uma formação completa, facilitando o estudo de casos e a retirada de conclusões, bem como na tomada de decisões, contribuindo assim para a sua eficiência laboral.
- 2- Analisada a bibliografia e o sistema de ensino concluiu-se que, apesar de existir algumas limitações na aquisição do material bibliografico, os estudantes têm a seu dispor vasta bibliografia econometrica para a sua formação tornar-se eficiente, só que uma boa parte dela encontra-

se escrita em Inglês. Existe uma necessidade tal de se ter mais bibliografia em português. Os docentes ajudam utilizando todos os recursos disponíveis e que segundo estes são insuficientes e muito limitados. Tem-se procurado equipar os laboratórios informáticos com software específicos para aplicação prática dos modelos econométricos, o que em muitas instituições estes ainda encontram-se apenas nos projectos e não há nada de concreto como é o caso da Katiavala Bwila. O sistema de ensino apresenta algumas lacunas, uma vez que, em muitos casos o estudante chega ao ensino superior sem bases de Estatística o que dificulta o docente de Econometria na transmissão dos conteúdos econométricos perdendo tempo ensinando ainda as bases e não o conteúdo econométrico propriamente dito. O que faz com que os docentes venham a fazer alguns cortes no conteúdo econométrico a apresentar, o que contribui de certo modo para a fraca assimilação por parte dos estudantes. Contudo tem-se procurado apresentar aos estudantes os conhecimentos necessários para que com eficiência possam tomar decisões corretas no exercício da sua atividade laboral e atender as necessidades económicas-financeiras das empresas da região desenvolvendo-as economicamente.

- 3- Quanto aos docentes, na sua maioria tem alguma formação no ramo das ciências económicas, embora haja outros formados apenas em Matemática ou estatística, no entanto, estão capacitados para esta atividade devido à experiência adquirida nos vários anos de trabalho lecionando a cadeira de Econometria e não têm encontrado dificuldades significativas ao lecionarem tal cadeira.

No entanto as instituições superiores devem criar condições para que haja laboratórios de informática com software específicos para o estudo de modelos econométricos, visto que, a falta de laboratórios equipados com software específicos tem sido muitas das vezes apontada como a causa principal para a falta da realização das aulas práticas.

Bibliografia

Andrade, João Sousa, Apontamentos sobre econometria aplicada (dezembro 2001 – Maio 2004).

Bernanke, Ben S. E Blinder. The federal funds rate and the channels of monetary transmission. .american economic review , Alan (1992).

Bernanke, Ben S. e Mihov, Ilian (1998). Measuring monetary policy. Quarterly journal of economics.

Blanchard, O. J. E Quah, D. The dynamic effects of aggregate supply and demand disturbances. American economic review. Vol.79, (1989).

Boletim – Sociedade portuguesa de estatística (publicação semestral outono de 2009).

Fumio Hayashi, Princeton, NJ. Econometrics. Princeton University Press (2000),

Gujarati, D. Econometria básica. Terceira edição, markrow books (2000).

Gujarat, Damodar N. e Porter, Dawn C. Econometria básica. Quinta edição (2008). Publicado em Língua Portuguesa por AMGH editora Ltda (2011).

Harris, Ethan. <<Forecasting Automobile output>> federal reserve bank of New York quarterly review. Reeditado em Mansfield, managerial. Econometrics and operations rescarch. 5ª edição (inverno 1985-86).

Hendry, D. F. econometric Methodology: a personal perspective. In: Bewlry, T. F. advances in econometrics. Cambridge university press (1987).

John Wiley & Sons; Science ed. (1965), Econometrics.

Mahia, Ramon, Econometria I sesion De Raposo, conceitos basicos en torno a la construccion de un modelo econométrico y las características del modelo básico de regresión lineal, Octubre 2004.

Mansfield, Edwin; ciência e técnica; ECONOMIA EMPRESARIAL – teoria, aplicação e casos, instituto Piaget (1995).

Oliveira, M. Mendes; Santos, Luis Delfim e Fortuna, Natercia. ECONOMETRIA, Escolar Editora (2011).

Peixoto, Pedro da Silva, universidade de São Paulo, (2005), Tese: O uso de modelos econométricos em empresas; curso de matemática aplicada e computacional.

Programas da disciplina de econometria das universidades Jean Piaget-Benguela (2012). Lusíadas de Angola-Benguela (2012) e Katiavala Buila- de Benguela(2012).

Sims, C. "Macroeconomics and reality", Econometria (1980)

Spanos, A. Cambridge university press (1989), Statistical foundations of econometric modelling Cambridge.

Tintner, Gerhard. Econometrics. Nova York: John Wiley & sons (1965)

Science ed.

Valle, Carlos Antônio Abanto, UFRJ, Tese: Métodos de simulação estocástica em modelos dinâmicos não lineares: uma aplicação em modelos de volatilidades (Agosto 2005).

Sites pesquisados:

<http://cangue.blogspot.com/2012/01/instituicoes-de-ensino-superior-privado.html> (atualizado 11 de abril 2013).

<http://www.angola.gov.ao/portaiscidadao/conhecimento2.htm>

"pt.Wikipedia.org/wiki/Econometria.Econometria-wikipedia,aenciclopedia livre".

"bdadolfo.blogspot.com/2009/08/mitos-sobre-econometria.html"

Adolfo sachside-opiniões sobre econometria".

"Econometria,pt,scribd.com/doc/541619005/econometria"

Anexos

Entrevista aos Coordenadores de Disciplina

Venho por esta solicitar a sua cooperação para dar avanço ao meu trabalho de investigação relacionado a aplicação dos modelos econométricos no ensino superior angolano, pelo que desejo desse resposta às questões abaixo. Estou certa que os seus critérios serão de grande utilidade, dado o rol que desempenha na direção do coletivo de professores da disciplina. Desde já envio os meus agradecimentos.

1. Que nível profissional quanto ao agregado pedagógico e académico possuem os docentes ligados a disciplina de econometria?
 - a) Superior com agregado pedagógico
 - b) Superior sem agregado pedagógico
 - c) Mestre com agregado pedagógico
 - d) Mestre sem agregado pedagógico
 - e) Outra classificação_____.
2. Considera que é importante o tratamento metodológico para o seu coletivo de professores? Por quê?
3. Considera ser importante o conhecimento sobre modelos econométrico na formação de um gestor? Justifique.
4. Tem recebido dos professores alguma reclamação sobre dificuldades na transmissão do conteúdo referente aos modelos econométricos?
 - a) Se sim, pode citar algumas destas dificuldades?
5. Considera que é necessário vincular o conteúdo matemático e estatístico aos conhecimentos econométricos? Por quê?
6. Acha que o Tema sobre aplicação dos modelos econométricos tem sido bem tratado e que os estudantes conseguem assimilar o conteúdo correspondente?
7. Considera a carga horaria um fator importante na assimilação do conteúdo por parte do estudante?

8. Considera ser eficiente a metodologia usada no ensino dos modelos econométricos? Justifique.
9. No ensino dos modelos econométricos tem-se buscado auxílio nos conhecimentos informático? De que forma.