Mark B. STEWART & Kenneth F. WALLIS. *Introductory Econometrics*. 2ª ed. Basil Blackwell, 1986.

por Rodolfo Hoffmann

Os primeiros capítulos do livro, correspondendo a aproximadamente 1/3 do seu volume, são dedicados à análise dos modelos econômicos estáticos e dinâmicos, sem se envolver nos problemas estatísticos de estimação de parâmetros e teste de hipóteses. Nesses capítulos são apresentados e discutidos, através de vários exemplos, os conceitos de variável endógena, variável exógena, equações simultâneas, variáveis defasadas, as condições de estabilidade de modelos dinâmicos e o problema de identificação. Isso distingue essa obra da maioria dos textos de Econometria, nos quais a análise dessas questões está espalhada pelos capítulos finais e os capítulos iniciais são dedicados aos métodos estatísticos de estimação de parâmetros e teste de hipóteses em regressão linear simples e múltipla. A principal exceção é o livro de C.F. Christ, "Econometric Models and Methods", cuja influência é explicitamente conhecida por Stewart e Wallis. Deve-se assinalar, entretanto, que o texto desses autores é mais elementar e mais didático.

O livro é bem escrito e os exercícios que apresenta o valorizam como obra didática. Pelas suas qualidades, seria recomendável a sua tradução (uma boa tradução!), complementando a literatura já disponível em português.

Ao explicar o teste de uma restrição linear através do teste F, os autores afirmam (p.195): "Thus we expect our test statistic to take a rather small (positive) value, i.e., to be 'close' to zero". Na realidade, quando a hipótese da nulidade for verdadeira, as esperanças do numerador e do denominador do F calculado são iguais à variância do erro (σ^2). Então o valor de F tende a ficar

próximo de 1, e não próximo de zero. A afirmativa de que o valor de F fica próximo de zero não se justifica como uma simplificação didática. Há casos em que um valor de F bastante próximo de zero não deve ser interpretado, simplesmente, como um resultado não-significativo no teste da validade da restrição linear. Vejamos um caso em que o valor de F bastante próximo de zero indica que houve fraude ou manipulação dos dados. Sejam Y_{1t} e Y_{2t} áreas plantadas com soja, no ano t , nos municípios 1 e 2, respectivamente. Vamos admitir que foram fornecidos valores de Y_1 e Y_2 para os últimos T anos, mas desconfiamos que, nesse período, a pessoa encarregada de coletar os dados no município 2 não se dava a esse trabalho, tendo desenvolvido um método mais cômodo: ele aguardava a coleta dos dados no outro município, verificava qual havia sido a variação percentual da área plantada em relação ao ano anterior, a aplicava a mesma taxa de variação para "estimar" a área plantada no seu município. Vejamos como testar a hipótese de que os dados foram forjados dessa maneira. Seja Z_{kt} , com $k = 1, 2$ e $t = 1, \dots, T$, uma variável binária que assume valor zero quando se trata da área plantada no município 1 e assume valor 1 quando se trata da área plantada no município 2, isto é, $Z_{1t} = 0$ e $Z_{2t} = 1$ para todo t .

Consideramos o modelo $\ln Y_{kt} = \alpha + \gamma Z_{kt} + \beta t + \delta t Z_{kt} + u_{jt}$

A hipótese de que a taxa de crescimento da série dos Y_1 é igual à taxa de crescimento da série dos Y_2 pode, dessa maneira, ser escrita $H_0 : \delta = 0$. Calculado o valor de F para testar essa hipótese, podemos considerar 3 possibilidades: 1ª) um valor bastante grande permite afirmar que as taxas de crescimento são estatisticamente diferentes; 2ª) um valor de F "próximo" de 1 não permite rejeitar a hipótese $H_0 : \delta = 0$, mas não é indicação de que os dados foram forjados; 3ª) um valor de f muito próximo de zero indica que as taxas de crescimento são extraordinariamente semelhantes, o que pode ser considerado indicador de que os dados foram forjados. É claro que a delimitação do que é "muito próximo de zero" é feita considerando valores críticos, na extremidade inferior da distribuição de F . Esses valores críticos podem ser obtidos das tabelas usuais de F (que apresentam valores críticos que delimitam regiões de rejeição na cauda superior da distribuição) considerando que $F_{m,n} = F_{n,m}^{-1}$.