A QUESTÃO DA HETEROCEDASTICIDADE NA COMPARAÇÃO DE DUAS MÉDIAS

RONALDO SEICHI WADA

Orientador : Prof. Dr. IZAIAS RANGEL NOGUEIRA

Dissertação apresentada à Escola Superior de Agricultura "Luiz de Queiroz", da Universidade de São Paulo, para obtenção do título de Mestre em Agronomia, Área de Concentração em "Estatística e Experimentação Agronômica".

PIRACICABA
Estado de São Paulo - Brasil
Abril, 1985

*A minha esposa e filhos,
a meus pais,*

DEDICO.

AGRADECIMENTOS

Expressamos nossos agradecimentos a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho, em especial:

- Ao Prof. Dr. Izaias Rangel Nogueira, que além de orientador, foi um grande amigo, transmitindo-nos com muita dedicação seus conhecimentos e experiência;
- À Profª Drª Sonia Vieira, pelo seu apoio constante, e pelas sugestões apresentadas durante a elaboração deste trabalho;
- Ao Prof. Dr. Rodolfo Hoffmann, pelas sugestões apresentadas;
- À Universidade Metodista de Piracicaba (UNIMEP), pelo auxílio financeiro, e ao Centro de Ciências Exatas da UNIMEP pela disposição de seus equipamentos para a realização deste trabalho;
- À Bibliotecária Ivany do Carmo Guidolin Gerola, pela revisão técnica;
- À Maria Izalina Ferreira Alves, pela datilografia deste trabalho;

- Aos amigos Mário Alcides de Oliveira Scafí e Elizabete Rodrigues de Oliveira Scafí, pela revisão do texto original;
- Finalmente, nossos agradecimentos a todos professores e funcionários do Departamento de Matemática e Estatística da ESALQ.

Í N D I C E

	Pág.
RESUMO	vii
SUMMARY	x
1. INTRODUÇÃO	1
2. MATERIAL E MÉTODOS	12
2.1 - Conceitos Básicos sobre Erro Tipo I e Erro Tipo II	12
2.2 - A Comparação de Duas Médias	15
2.2.1 - Comparação de duas médias quando as variâncias populacionais são conhecidas	18
2.2.2 - Comparação de duas médias quando as variâncias populacionais são desconhecidas mas iguais por pressuposição	20
2.2.3 - Comparação de duas médias quando as variâncias populacionais são desconhecidas e não supostas iguais ...	28
2.2.3.1 - Procedimentos alternativos baseados no cálculo do número de graus de liberdade	29
2.2.3.2 - Procedimento alternativo baseado no cálculo do valor crítico	35
2.2.3.3 - Comparações de médias nos ensaios em "split-plot"..	37

2.3 - Simulação de Amostras por Geração de Variáveis	38
2.3.1 - Geração de variáveis com distribuição normal reduzida	38
2.3.2 - Geração de pares de amostras de variáveis com distribuição normal ..	39
2.3.3 - Estimativa do nível de significância	44
2.3.4 - Estimativa do poder do teste	48
3. RESULTADOS	51
4. DISCUSSÃO	54
5. CONCLUSÕES	75
6. LITERATURA CITADA	78
7. APÊNDICE	83

A QUESTÃO DA HETEROCEDASTICIA NA COMPARAÇÃO DE DUAS MÉDIAS

Autor: Ronaldo Seichi Wada
Orientador: Prof. Dr. Izaias Rangel Nogueira

R E S U M O

O problema de Behrens-Fisher é clássico na literatura. No contexto mais simples, este problema se refere à comparação de duas médias, usando o teste t de Student, quando as variâncias são desconhecidas e não são supostas iguais, ou seja, quando ocorre heterocedasticia. Nesses casos, existem diversas propostas alternativas para o teste. Comparamos essas propostas, simulando 216 mil pares de amostras, nas seguintes condições:

- 1) Amostras com médias iguais e variâncias iguais ($\sigma_1^2 = 4, 8, 16$), com total de elementos nas amostras $N = 12, 24, 36$, nas razões $n_2/n_1 = 1, 2, 3$;
- 2) Amostras com médias iguais e variâncias diferentes, nas razões $\sigma_2^2/\sigma_1^2 = 2, 3, 4$ ($\sigma_1^2 = 4, 8, 16$),

com total de elementos nas amostras $N = 12, 24, 36$, nas razões $n_2/n_1 = 1, 2, 3$;

- 3) Amostras com médias diferentes e variâncias iguais ($\sigma_1^2 = 4, 8, 16$), com total de elementos nas amostras $N = 12, 24, 36$, nas razões $n_2/n_1 = 1, 2, 3$;
- 4) Amostras com médias diferentes e variâncias diferentes, nas razões $\sigma_2^2/\sigma_1^2 = 2, 3, 4$ ($\sigma_1^2 = 4, 8, 16$), com total de elementos nas amostras $N = 12, 24, 36$, nas razões $n_2/n_1 = 1, 2, 3$.

Os resultados obtidos permitiram concluir que:

- 1) Os resultados obtidos através do teste t usual, e do teste t , nas formas propostas por Welch-Satterthwaite e Aspin-Welch são praticamente iguais, quando as amostras são de mesmo tamanho. Então, nesses casos, é recomendável proceder ao teste t , na forma usual, porque os cálculos são mais simples.
- 2) Se a razão de tamanhos de amostras for da ordem de 2 e 3, o teste t , na forma usual, não leva a bons resultados. Nesses casos não é recomendável proceder ao teste usual, a menos que a aplicação do teste seja precedida por uma transformação de variáveis que estabilize a variância.

- 3) Se a razão de tamanhos de amostra for da ordem de 2 e 3, o teste t , na forma proposta por Aspin-Welch, leva a resultados ligeiramente melhores em termos de poder de teste àqueles obtidos quando se procede ao teste t , na forma proposta por Welch-Satterthwaite, o que implica em um teste de maior poder, consequentemente, menor probabilidade de cometer erro tipo II. Essa diferença, entre as duas formas de teste, diminui com o aumento do total de elementos nas duas amostras. No entanto, se houver interesse em um teste associado a menor probabilidade de cometer erro tipo I, é razoável proceder ao teste t , na forma proposta por Welch-Satterthwaite, o que não implica em perda considerável do poder.
- 4) O teste t , na forma proposta por Cochran, tem menor poder se a razão de tamanhos de amostra for da ordem de 2 e 3. O poder cresce com o aumento do total de elementos nas duas amostras. No entanto, este teste pode ser mais seguro se, no problema dado, existe prejuízo real em cometer erro tipo I.

ON THE HETEROSCHEDASTICITY IN THE COMPARISON
BETWEEN TWO MEANS

Author: Ronaldo Seichi Wada
Adviser: Prof. Dr. Izaias Rangel Nogueira

S U M M A R Y

It is classical for statisticians the Behrens-Fisher problem. In the context of two samples, the Behrens-Fisher problem is, in short, the testing of the equality of two means by using the Student t-test when the variances are by assumption unknown and unequal, i.e., when heteroschedasticity is assumed. There are some alternative procedures for such tests, compared herein by simulating 216 thousands of two-samples, in the following conditions:

- 1) Samples with equal means and equal variances, sample sizes with ratio $n_2/n_1 = 1, 2, 3$, and $N = n_1 + n_2 = 12, 24, 36$.

- 2) Samples with unequal means and equal variances ($\sigma_1^2 = 4, 8, 16$), sample sizes with ratio $n_2/n_1 = 1, 2, 3$, and $N = n_1 + n_2 = 12, 24, 36$.
- 3) Samples with equal means and unequal variances, ratio $\sigma_2^2/\sigma_1^2 = 2, 3, 4$ ($\sigma_1^2 = 4, 8, 16$), sample sizes with ratio $n_2/n_1 = 1, 2, 3$, and $N = n_1 + n_2 = 12, 24, 36$.
- 4) Samples with unequal means and unequal variances, ratio $\sigma_2^2/\sigma_1^2 = 2, 3, 4$ ($\sigma_1^2 = 4, 8, 16$), sample sizes with ratio $n_2/n_1 = 1, 2, 3$, and $N = n_1 + n_2 = 12, 24, 36$.

It was concluded, from the findings, that:

- 1) The Student t-test, the Welch-Satterthwaite and the Aspin-Welch t's give similar results when sample sizes are equal. Thus in such cases the Student t-test should be applied, because computation is easier when this procedure is adopted.
- 2) When the sample size ratio is 2 or 3, the Student t-test do not lead to reasonable results. So this procedure is appropriate only when a transformation makes the samples variances nearly equal.
- 3) When the sample size ratio is 2 or 3, the Aspin-Welch approximate t leads to greater power, com-

pared to the Welch-Satterthwaite approximate t-test. Differences between procedures are smaller when larger sample sizes are used. However, it is reasonable to use Welch-Satterthwaite procedure when there is awareness that type I error should not be committed. In such cases, the loss in power can be disregarded.

- 4) When the sample size ratio is 2 or 3, the Cochran approximate t is less powerful, but increases with the sample size. It should be noted, however, that the Cochran solution might be safer when type I error should not be committed.

1. INTRODUÇÃO

A inferência estatística, tanto em seus aspectos teóricos como práticos, tem seu começo ligado ao trabalho desenvolvido por STUDENT (1908). Nesse trabalho foi apresentada a distribuição da razão z , definida por esse autor como sendo a diferença entre média da amostra e média da população, dividida pelo desvio padrão da amostra, isto é:

$$z = (\bar{x} - \mu) / \sqrt{\frac{n}{\sum (x_i - \bar{x})^2} / n}$$

Também foi calculada uma tabela que dá os valores da distribuição, para amostras de tamanho 4 a 10. Essa tabela é diferente das tabelas de t , utilizadas hoje, que foram ampliadas e generalizadas. O histórico da tran-

sição entre o valor de z , proposto por STUDENT (1908) e o valor de t , como conhecido hoje, é apresentado em EISENHART (1979) e BOX (1981).

Além de dar o desenvolvimento teórico, STUDENT (1908) também se preocupou em dar alguns exemplos práticos. Ele aplica o teste em três experimentos, sendo um da área médica e dois da área agrônômica. Tais dados hoje seriam submetidos à análise de variância, devido à possibilidade de comparações múltiplas.

Em suas conclusões, STUDENT (1908) afirma que a distribuição por ele estabelecida leva a bons resultados, mesmo nos casos em que a variável tenha distribuição apenas aproximadamente normal.

O trabalho de STUDENT (1908) passou despercebido durante alguns anos e, só começou a ser melhor discutido depois de descoberto por Fisher. A correspondência entre Fisher e Student, discutida por BOX (1981), filha de Fisher, mostra aspectos curiosos sobre o desenvolvimento do teste t . A autora conta que Student era uma pessoa extraordinária, com uma variedade enorme de interesses. Trabalhava na Cervejaria Guinness, mas desenvolveu pesquisas muito além das necessidades do seu trabalho. Aliás, Student é apenas um pseudônimo. O nome do autor é William Sealy Gosset (1876-1937). Durante 1907, esteve trabalhan-

do com Karl Pearson na Universidade de Londres, onde desenvolveu o teste z , depois transformado em t .

Entretanto, de acordo com BOX (1981), Gosset não acreditava que seu trabalho fosse tão importante como enfatizava Fisher, que escreveu: "O estudo correto das distribuições estatísticas começa em 1908 com o artigo de Student 'The probable error of a mean'", em seu livro *Statistical Methods for Research Workers* (1925).

Outro aspecto curioso, comentado por BOX (1981) refere-se à transformação do valor z , proposto por STUDENT (1908) para $t = z(n-1)^{1/2}$. Embora contasse com a ajuda de Fisher, Student recalculou a tabela sozinho, com dificuldade, uma vez que a máquina de calcular era propriedade da Cervejaria. Como essa máquina era muito usada no verão, para cálculos da empresa, os cálculos estatísticos tinham que ser adiados até o inverno, quando diminuíam as vendas. A transformação de z de STUDENT (1908) em $t = z(n-1)^{1/2}$ foi objeto de artigo publicado por Fisher em 1925. Student mencionou explicitamente o uso da designação " t " pela primeira vez em uma carta dirigida à Fisher, datada de 7 de novembro de 1922, segundo EISENHART (1979).

Segundo ARMITAGE (1982), Fisher propôs a casualização por volta de 1920. Entretanto, esta só foi incorporada à pesquisa agrônômica na década de 1940 (BROWN

e HOLLANDER, 1977). Em 1923, embora a idéia de casualização ainda não estivesse publicada, Student expressou descrença sobre sua validade, segundo BOX (1981). Mais tarde, no entanto, admitiu sua importância. Aliás, a principal crítica de STUDENT (1931) ao famoso experimento de Lanarkshire era a falta de casualização.

Esse experimento foi conduzido em 1930, nas escolas de Lanarkshire, Escócia. Durante quatro meses 10.000 escolares receberam leite diariamente, sendo que 5.000 receberam leite cru e 5.000 receberam leite pasteurizado. Outras 10.000 crianças foram utilizadas como controle. Então, um total de 20.000 crianças foram medidas e pesadas, tanto no início como no final do experimento. O estudo mostrou que as crianças que receberam quantidade adicional de leite ganharam mais peso, durante o período, do que as crianças do grupo controle. Seria a quantidade de leite a causa de maior ganho de peso?

A Figura 1.1, apresentada por ALTMAN (1982), mostra as mudanças de peso para as meninas. Existem duas características que chamam a atenção nesse gráfico. A primeira é o fato de que os grupos controle, de todas as idades, eram mais pesados do que os grupos que receberam leite, tanto no início como ao final do experimento. Isto foi explicado por STUDENT (1931) pelo fato de que alguns pes-

quisadores não respeitaram a casualização e deram o leite para as crianças mais pobres.

Figura 1.1 - Comparação do grupo controle e grupo experimental (quantidade adicional de leite), para as meninas do experimento de Lanarkshire..

O segundo aspecto curioso do gráfico é o fato de que a taxa de crescimento observada, em cada grupo, é muito menor do que a que seria esperada, quando se observa o grupo seguinte de idade. A explicação, segundo ALTMAN (1982), é simples: o estudo começou em fevereiro e terminou em junho e, em ambas as ocasiões, as crianças foram pesadas com roupas. O aumento de peso, relativamente

pequeno, é devido à quantidade diferente de roupas usadas no inverno e no verão. Ainda, o crescimento ponderal menor do grupo controle pode ser explicado pelo fato de que essas crianças — mais ricas — usavam roupas mais pesadas no inverno.

O teste z de STUDENT (1908) foi transformado em teste t por Fisher em 1925. Este trabalho de Fisher apresenta uma forma geral para tratar os testes de significância de uma média, da diferença de duas médias, de coeficientes de correlação simples e parciais e dos coeficientes de uma regressão.

No entanto, o problema de Behrens-Fisher foi levantado bem mais tarde. Basicamente, esse problema consiste em fazer inferência sobre a diferença de duas médias, que têm variâncias desconhecidas e essas variâncias não são supostas iguais.

Se as variâncias são iguais, a estatística

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \theta}{\sqrt{\left[\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} \right] \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

tem distribuição de t com (n_1+n_2-2) graus de liberdade.

Se as variâncias são desiguais, o teste da diferença de médias não tem solução simples. A forma de proceder ao teste, nesses casos, foi proposta por FISHER (1936), que notou a equivalência dessa solução com o teste anteriormente proposto por Behrens em 1929. A estatística é:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

cujá distribuição não é a de t de Student.

É interessante lembrar que Fisher nunca deu maior importância ao problema. Em seu livro Statistical Methods for Research Workers (1932), Fisher diz que o teste t, na forma clássica, não exige a pressuposição de que as variâncias sejam iguais. Ele diz, textualmente, que a igualdade de variâncias é parte da hipótese, uma vez que está em teste o fato de as duas amostras terem provindo da mesma população. Nessas condições, a validade do teste é, evidentemente total. Nesse livro, o autor ainda lembra que a hipótese poderia ser colocada na forma: "As amostras seriam provenientes de duas populações com a mesma média?", mas esta seria apenas uma discussão acadêmica, de pouca validade para a discussão de experimentos reais. Segundo

WALLACE (1980), Fisher não desconhecia o problema das variâncias desiguais, mas a este se referia como um problema teórico.

No entanto, o problema não é apenas teórico. Ocorre em situações reais. Por exemplo, é razoável pressupor que qualquer tratamento, que determine uma diminuição drástica de incidência leve a uma redução tanto da média como da variância, em relação ao controle (a média e a variância do grupo tratado tendem a zero).

WALLACE (1980) argumenta que a colocação de Fisher é discutível. Esse autor cita um exemplo bastante antigo apresentado na Tabela 1.1, sobre quantidade de nitrogênio, obtida do ar e através de processo químico, para ilustrar seu argumento. Existem dados negativos e no trabalho de Wallace havia um erro de datilografia, corrigido aqui.

Tabela 1.1- Quantidade de nitrogênio, obtida por dois processos.

	PROCESSOS	
	A partir do ar (1893)	Químico (1894)
	1.035	143
	1.026	-110
	1.024	-184
	1.012	182
	1.027	
$\bar{x}$	1.025	8
s^2	69	33.096

Fonte: RAYLEIGH (1893-1894).

De acordo com WALLACE (1980), os dados apresentados na Tabela 1 são um exemplo claro do problema de Behrens-Fisher. A hipótese da nulidade especifica apenas a igualdade de médias. Não existe razão para pressupor que processos diferentes de obter nitrogênio tenham a mesma variabilidade. Na realidade, as variâncias amostrais são bastante desiguais. Ainda, o processo químico está associado a um valor tão alto de variância que leva à incerteza sobre o valor da média.

WALLACE (1980) sugere que, neste exemplo, é razoável proceder a um teste para comparar a média, obtida através de processo químico, com a média 1.025, mas usando este valor como "conhecido" μ . A variância é estimada com 3 graus de liberdade, que serão os graus de liberdade do teste.

É claro que, neste exemplo, a razão de variâncias é tão alta que o torna um caso extremo. Então, o procedimento parece razoável. É fácil ver que:

$$\frac{s_2^2}{s_1^2} = \frac{33.096}{69} = 479,65 ,$$

valor que dificilmente ocorre em situações reais.

A idéia de usar um valor especificado para a média de grupo controle, obtido através de valores his-

tóricos foi enfatizada por GEHAN e FREIREICH (1974). Esses autores argumentam que, dessa forma, se obtem maior precisão na comparação. A variância que se usaria no teste seria σ^2/N , onde N é o número total de elementos no grupo tratado. Esta colocação, no entanto, é criticada por MEIER (1975) que, entre diversos argumentos, afirma que dificilmente um pesquisador usaria controles históricos para testar o efeito de drogas, dada a influência secular sobre as respostas. É claro que, na área médica, o uso de controles históricos é às vezes defensável, mas não tem sentido em experimentos da área agrônômica e outras.

O problema de Behrens-Fisher foi tratado por diversos autores. Não existe uma solução que possa ser usada com facilidade na prática. Foram propostas diversas formas alternativas para testar duas médias, quando as variâncias são desconhecidas e supostas desiguais. Assim, existe a aproximação proposta por WELCH (1937) que é equivalente à generalização proposta por SATTERTHWAITTE (1946) para o caso particular de duas amostras. Esta solução consta em diversos livros-textos e tem sido bastante utilizada na prática. Também existe a aproximação proposta por WELCH (1947) e por ASPIN (1948), na realidade bastante semelhantes. Esta aproximação, citada por alguns como de Aspin-Welch, consta também em diversos livros-textos.

tos. Ambas as aproximações são, na realidade, correções para os graus de liberdade que devem ser usados quando se procede ao teste t .

Existe ainda a proposta de COCHRAN e COX (1957), que propõem uma fórmula para recalcular o valor crítico de t .

É claro que, nestas situações, também podemos optar por testes não-paramétricos descritos, por exemplo, em HOLLANDER e WOLFE (1973), LEHMAN (1975), CAMPOS (1983). No entanto, o teste t de Student é muito mais conhecido e utilizado. Por isso, neste trabalho apresentamos uma comparação do poder do teste t , nos casos em que as variâncias são desiguais, usando o t , na forma clássica (ponderando as variâncias) e usando as aproximações citadas, para as seguintes razões de variância $\sigma_2^2/\sigma_1^2 = 2, 3, 4$.

2. MATERIAL E MÉTODOS

2.1 - Conceitos Básicos sobre Erro Tipo I e Erro Tipo II

No procedimento de testes estatísticos, por definição, erro tipo I consiste em rejeitar a hipótese de nulidade, quando essa hipótese é verdadeira. Erro tipo II consiste em não rejeitar a hipótese da nulidade, quando essa hipótese é falsa.

Estas definições são apresentadas, resumidamente, na Tabela 2.1.

Tabela 2.1 - Erro tipo I e erro tipo II em função da decisão.

DECISÃO	H_0	
	Verdadeira	Falsa
Não rejeitar H_0	Decisão correta	Erro tipo II
Rejeitar H_0	Erro tipo I	Decisão correta

No caso particular, discutido nesse trabalho, a hipótese da nulidade corresponde à hipótese de igualdade de duas médias, isto é:

$$H_0: \mu_1 = \mu_2$$

O erro tipo I ocorre com probabilidade α . Essa probabilidade é denominada nível de significância do teste. O erro tipo II está associado à probabilidade β . Então

$$\alpha = P\{\text{Rejeitar } H_0 \mid H_0 \text{ é verdadeira}\}$$

$$\beta = P\{\text{Não rejeitar } H_0 \mid H_0 \text{ é falsa}\}$$

Por definição, poder do teste é o valor $(1 - \beta)$, indicado por alguns pela letra grega π (GIBRA, 1973). Então, poder do teste é a probabilidade de rejeitar H_0 , quando H_0 é falsa. Assim:

$$1 - \beta = P\{\text{Rejeitar } H_0 | H_0 \text{ é falsa}\}$$

Neste trabalho vamos definir, como hipótese alternativa, a hipótese de que a média da primeira amostra é maior do que a média da segunda, pré-fixada com valor igual a 10:

$$H_1: \mu_1 - \mu_2 > 0$$

Neste caso, o teste é unilateral. Vamos pressupor que a variável aleatória X em análise tenha distribuição normal. Se as variâncias σ_1^2 e σ_2^2 , das duas amostras, forem conhecidas e iguais, as probabilidades α e β podem ser facilmente visualizadas num gráfico, como mostra a Figura 2.1.

Figura 2.1 - Probabilidade de erro tipo I (α) e probabilidade de erro tipo II (β).

Neste trabalho, α foi pré-fixado em 0,05, valor bastante utilizado na prática. Aliás, em vários campos da área biológica tornou-se usual mencionar a probabilidade de 0,05 ou menos como sendo aquela que mostra evidência **estatisticamente significativa** contra a hipótese da nulidade (FINNEY, 1955). Quando o valor calculado de um teste excede o valor das tabelas estatísticas para $\alpha = 0,05$, muitos autores (por exemplo: SNEDECOR e COCHRAN, 1967 e PIMENTEL GOMES, 1982) usam um asterisco, para indicar a significância do resultado.

2.2 - A Comparação de Duas Médias

Vamos supor que pretendemos comparar as médias de duas amostras aleatórias independentes. Façamos X_{ij} indicar a variável em análise, onde o índice j representa a j -ésima observação ($j=1,2,\dots,n_i$) da i -ésima amostra ($i=1$ e 2 se referem às amostras em comparação).

Vamos supor ainda, que X_{1j} tenha distribuição normal de média μ_1 e variância σ_1^2 , e que X_{2j} tenha distribuição normal de média μ_2 e variância σ_2^2 , isto é:

$$X_{1j} \sim N(\mu_1, \sigma_1^2)$$

e

$$X_{2j} \sim N(\mu_2, \sigma_2^2)$$

Para testar a hipótese $H_0: \mu_1 = \mu_2$, existem testes estatísticos paramétricos e não-paramétricos. No presente trabalho discutiremos o teste t e, para tal, usaremos importante teorema da Estatística:

Teorema 1:

Se $X_1, X_2, \dots, X_n$ são variáveis aleatórias independentes, com distribuições normais de médias $\mu_1, \mu_2, \dots, \mu_n$ e variâncias $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$, respectivamente, a variável aleatória:

$$Y = K_1 X_1 + K_2 X_2 + \dots + K_n X_n = \sum_{i=1}^n K_i X_i,$$

que é uma combinação linear de variáveis com distribuições normais, tem também distribuição normal, com média $(K_1 \mu_1 + K_2 \mu_2 + \dots + K_n \mu_n)$ e variância $(K_1^2 \sigma_1^2 + K_2^2 \sigma_2^2 + \dots + K_n^2 \sigma_n^2)$, onde $K_1, K_2, \dots, K_n$ são constantes quaisquer ^{1/}.

Como consequência, se a variável X_i tiver distribuição normal, então a média $\bar{X}$, proveniente de uma amostra constituída de n observações, também tem distribuição normal porque:

$$\bar{X} = \sum_{i=1}^n \left(\frac{1}{n}\right) X_i$$

^{1/} Ver, por exemplo; HOGG e CRAIG (1970, p. 158-9); BROWNLEE (1965, p. 82).

é uma combinação linear de variáveis normalmente distribuídas. Além disso, se a amostra for aleatória, as n observações são independentes e

$$V(\bar{X}) = \frac{\sigma^2}{n} \quad ,$$

onde σ^2 é a variância de X .

Consequentemente, a diferença $(\bar{X}_1 - \bar{X}_2)$, onde $\bar{X}_1$ e $\bar{X}_2$ são as médias das amostras aleatórias, independentes, de tamanhos n_1 e n_2 respectivamente, tem distribuição normal, de acordo com o teorema anteriormente citado. Como $\bar{X}_1$ e $\bar{X}_2$ são variáveis independentes, normalmente distribuídas e com médias $E(\bar{X}_1) = \mu_1$ e $E(\bar{X}_2) = \mu_2$, a variável $(\bar{X}_1 - \bar{X}_2)$ tem média:

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$$

e variância

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) \quad .$$

Como $V(\bar{X}_1) = \frac{\sigma_1^2}{n_1}$ e $V(\bar{X}_2) = \frac{\sigma_2^2}{n_2}$, tem-se

$$V(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \quad .$$

2.2.1 - Comparação de duas médias quando as variâncias populacionais são conhecidas

Quando as variâncias σ_1^2 e σ_2^2 são conhecidas, para testar a hipótese de que a diferença das médias de duas populações tem um valor especificado θ , isto é, para testar $H_0: \mu_1 - \mu_2 = \theta$, utilizamos a variável:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{V(\bar{X}_1 - \bar{X}_2)}}$$

Substituindo $(\mu_1 - \mu_2)$ por θ e $V(\bar{X}_1 - \bar{X}_2)$ por $[(\sigma_1^2/n_1) + (\sigma_2^2/n_2)]$, temos:

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - \theta}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

que, sob a hipótese da nulidade (H_0), tem distribuição normal reduzida.

Para testar $H_0: \mu_1 - \mu_2 = \theta$, devemos comparar o valor calculado z da variável Z com o valor crítico z_0 , obtido a partir da tabela de distribuição normal reduzida, considerando-se um nível de significância α , tal que:

$$P(Z > z_0) = \frac{\alpha}{2}, \text{ se o teste for bilateral}$$

ou

$$P(Z > z_0) = \alpha, \text{ se o teste for unilateral.}$$

O procedimento do teste consiste em rejeitar a hipótese da nulidade $H_0: \mu_1 - \mu_2 = \theta$ se:

- a) $|z| \geq z_0$ em um teste bilateral, onde a hipótese alternativa é $H_1: \mu_1 - \mu_2 \neq \theta$;
- b) $z \geq z_0$ em um teste unilateral, onde a hipótese alternativa é $H_1: \mu_1 - \mu_2 > \theta$;
- c) $z \leq -z_0$ em um teste unilateral, onde a hipótese alternativa é $H_1: \mu_1 - \mu_2 < \theta$.

Em geral, o valor especificado para θ é zero, isto é, em geral existe interesse em verificar se as médias são iguais. Isto significa que a hipótese da nulidade, em geral, é da forma: $H_0: \mu_1 - \mu_2 = 0$ ou, o que é o mesmo se $H_0: \mu_1 = \mu_2$. Então, se $\sigma_1^2 = \sigma_2^2$, testar $H_0: \mu_1 = \mu_2$ consiste em verificar se as duas amostras são provenientes de uma única população, com média $\mu_1 = \mu_2 = \mu$ e variância $\sigma_1^2 = \sigma_2^2 = \sigma^2$. Assim, a expressão se reduz a:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad [2.2.1]$$

2.2.2 - Comparação de médias quando as variâncias populacionais são desconhecidas mas iguais por pressuposição

Consideremos duas amostras aleatórias independentes com n_1 e n_2 elementos. Nos casos em que as variâncias populacionais são desconhecidas mas consideradas iguais por pressuposição, isto é, nos casos em que $\sigma_1^2 = \sigma_2^2 = \sigma^2$, onde σ^2 é um valor desconhecido, se $\bar{X}_1$ e $\bar{X}_2$ são médias amostrais, temos que:

$$V(\bar{X}_1 - \bar{X}_2) = V(\bar{X}_1) + V(\bar{X}_2) ,$$

uma vez que as amostras são independentes. Como $\sigma_1^2 = \sigma_2^2 = \sigma^2$, temos:

$$V(\bar{X}_1 - \bar{X}_2) = \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right) = \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \sigma^2 \quad [2.2.2]$$

Um estimador não-tendencioso de σ^2 é

$$S^2 = \frac{\sum_{j=1}^{n_1} (x_{1j} - \bar{X}_1)^2 + \sum_{j=1}^{n_2} (x_{2j} - \bar{X}_2)^2}{n_1 + n_2 - 2}$$

onde x_{1j} e x_{2j} são os valores observados das variáveis X_{1j} e X_{2j} nas amostras de tamanho n_1 e n_2 respectivamente. Tal expressão pode ser reescrita como segue:

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad [2.2.3]$$

que $\bar{x}$ é a média ponderada das variâncias das duas amostras, onde:

$$S_1^2 = \frac{\sum_{j=1}^{n_1} (x_{1j} - \bar{x}_1)^2}{n_1 - 1} \quad \text{e} \quad S_2^2 = \frac{\sum_{j=1}^{n_2} (x_{2j} - \bar{x}_2)^2}{n_2 - 1}$$

são os estimadores não-tendenciosos das variâncias populacionais.

Lembremos que a variável Z , definida em [2.2.1], tem distribuição normal. Quando o valor de σ^2 é desconhecido, podemos usar seu estimador, que é dado em [2.2.3]. Neste caso, devemos verificar a distribuição da expressão

$$\frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad [2.2.4]$$

Por um lado sabemos que, se Y tem distribuição normal de média igual a zero e variância igual a um, então a variável Y^2 tem distribuição de Qui-Quadrado (χ^2) com 1 (um) grau de liberdade.

Lembremos os enunciados de dois importantes teoremas ^{1/}:

^{1/} Ver, por exemplo: HOGG e CRAIG (1970, p. 159-160).

Teorema 2:

Sejam $X_1, X_2, \dots, X_n$ elementos de uma amostra aleatória de tamanho n , normalmente distribuídos com média μ e variância σ^2 . Então, a variável aleatória

$$Y = \sum_{i=1}^n \left(\frac{X_i - \mu}{\sigma} \right)^2$$

tem distribuição de χ^2 com n graus de liberdade.

Teorema 3:

Se as variáveis aleatórias $X_1, X_2, \dots, X_n$ são independentes e tem distribuição de χ^2 com $r_1, r_2, \dots, r_n$ graus de liberdade, respectivamente, então a variável:

$$Y = X_1 + X_2 + \dots + X_n$$

tem distribuição de χ^2 com $r_1 + r_2 + \dots + r_n$ graus de liberdade.

Com base nesses teoremas conclui-se que^{1/}:

$$\sum_{j=1}^{n_1} \left(\frac{X_{1j} - \mu_1}{\sigma_1} \right)^2 \sim \chi^2_{(n_1)}$$

e

^{1/} Ver, por exemplo: BROWNLEE (1965, p. 276).

$$\left[\frac{\bar{X}_1 - \mu_1}{\frac{\sigma_1}{\sqrt{n_1}}} \right]^2 \sim \chi^2_{(1)}$$

e ainda,

$$\sum_{j=1}^{n_1} \left(\frac{X_{1j} - \bar{X}_1}{\sigma_1} \right)^2 \sim \chi^2_{(n_1-1)}$$

Consideremos agora a expressão:

$$S_1^2 = \frac{\sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2}{n_1 - 1}$$

temos que

$$(n_1 - 1)S_1^2 = \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2$$

Logo,

$$\frac{(n_1 - 1)S_1^2}{\sigma_1^2} = \sum_{j=1}^{n_1} \left(\frac{X_{1j} - \bar{X}_1}{\sigma_1} \right)^2$$

e então:

$$\frac{(n_1 - 1)S_1^2}{\sigma_1^2} \sim \chi^2_{(n_1-1)}$$

Da mesma maneira, temos que:

$$\frac{(n_2 - 1)S_2^2}{\sigma_2^2} \sim \chi^2_{(n_2 - 1)}$$

Assim, $\frac{(n_1 - 1)S_1^2}{\sigma_1^2}$ tem distribuição de χ^2 com $(n_1 - 1)$ graus de liberdade e $\frac{(n_2 - 1)S_2^2}{\sigma_2^2}$ também tem distribuição de χ^2 com $(n_2 - 1)$ graus de liberdade. Como $\sigma_1^2 = \sigma_2^2 = \sigma^2$, temos que:

$$\frac{(n_1 - 1)S_1^2}{\sigma_1^2} + \frac{(n_2 - 1)S_2^2}{\sigma_2^2} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2}$$

tem distribuição de χ^2 com $(n_1 + n_2 - 2)$ graus de liberdade, com base no Teorema 3 (Teorema da Reprodutividade) da distribuição de qui-quadrado.

A expressão [2.2.4] pode ser reescrita como se segue:

$$\frac{\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}}{\sqrt{\frac{S^2}{\sigma^2}}} \quad [2.2.5]$$

Agora é fácil notar que o numerador tem distribuição normal de média zero e variância um, isto é, $N(0, 1)$.

Vimos que:

$$\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{\sigma^2} \cap \chi^2_{(n_1+n_2-2)}$$

Multiplicando e dividindo esta expressão por (n_1+n_2-2) , temos que:

$$(n_1+n_2-2) \frac{[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2]}{(n_1+n_2-2)\sigma^2} \cap \chi^2_{(n_1+n_2-2)}$$

e como

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1+n_2-2}$$

verificamos que

$$(n_1+n_2-2) \frac{S^2}{\sigma^2} \cap \chi^2_{(n_1+n_2-2)}$$

ou ainda

$$\frac{S^2}{\sigma^2} \cap \frac{\chi^2_{(n_1+n_2-2)}}{(n_1+n_2-2)}$$

Portanto, o denominador de [2.2.5] corresponde à raiz quadrada de uma distribuição de χ^2 , dividida

pelo número de graus de liberdade. A variável:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

onde o numerador e o denominador são independentes, tem distribuição t de Student, com (n_1+n_2-2) graus de liberdade ^{1/}.

Para testar a hipótese da nulidade $H_0: \mu_1 - \mu_2 = \theta$ podemos utilizar a variável:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\frac{1}{n_1} + \frac{1}{n_2})S^2}} \quad [2.2.6]$$

que tem distribuição t de Student, com (n_1+n_2-2) graus de liberdade, onde S^2 é a variância ponderada dada em [2.2.3].

Então, fixado o valor do nível de significância α , para testar $H_0: \mu_1 - \mu_2 = \theta$, comparamos o valor calculado t da variável T com o valor crítico t_0 , obtido na tabela de distribuição t de Student com (n_1+n_2-2) graus de liberdade, tal que:

^{1/} Ver, por exemplo: BROWNLEE (1965, p. 298).

$$P(T \geq t_0) = \frac{\alpha}{2}, \quad \text{se o teste for bilateral}$$

ou

$$P(T \geq t_0) = \alpha, \quad \text{se o teste for unilateral}$$

O procedimento do teste consiste em rejeitar a hipótese da nulidade $H_0: \mu_1 - \mu_2 = 0$ se:

- a) $|t| \geq t_0$ em um teste bilateral onde a hipótese alternativa é $H_1: \mu_1 - \mu_2 \neq 0$;
- b) $t \geq t_0$ em um teste unilateral onde a hipótese alternativa é $H_1: \mu_1 - \mu_2 > 0$;
- c) $t \leq -t_0$ em um teste unilateral onde a hipótese alternativa é $H_1: \mu_1 - \mu_2 < 0$.

Em geral, o valor de θ é igual a zero, isto é, testamos a hipótese da nulidade $H_0: \mu_1 = \mu_2$ contra a hipótese alternativa: $H_1: \mu_1 \neq \mu_2$ ou $H_1: \mu_1 > \mu_2$ ou ainda, $H_1: \mu_1 < \mu_2$.

Vimos como, através de [2.2.6], podemos testar a hipótese $H_0: \mu_1 = \mu_2$, quando se supõe $\sigma_1^2 = \sigma_2^2$ (homocedasticia), embora o valor da variância seja desconhecido. A mesma hipótese, nessas condições, pode ser testada através da análise de variância, obtendo resultados perfeitamente equivalentes.

A análise de variância é, entretanto, um procedimento de aplicação mais geral, podendo ser usado quando há mais de duas médias para comparar. Nesse sentido, o teste t pode ser considerado como um caso particular de análise de variância, quando houver somente duas médias para comparar.

2.2.3 - Comparação de duas médias quando as variâncias populacionais são desconhecidas e não são supostas iguais

Vamos supor que pretendemos comparar as médias de duas amostras aleatórias independentes. Façamos X_{ij} indicar a variável em análise, onde o índice j representa a j -ésima observação ($j=1,2,\dots,n_i$) da i -ésima amostra ($i=1,2$). Vamos supor ainda, que X_{1j} tenha distribuição normal de média μ_1 e variância σ_1^2 e que X_{2j} tenha distribuição normal de média μ_2 e variância σ_2^2 .

Com base em n_1 observações de X_1 e n_2 observações de X_2 , independentes, obtemos $\bar{x}_1$ e $\bar{x}_2$ e ainda s_1^2 e s_2^2 , que são as estimativas não-tendenciosas das médias e variâncias respectivamente.

Para testar a hipótese de que a diferença entre as médias de duas populações tem um valor especificado, isto é, $H_0: \mu_1 - \mu_2 = \theta$, quando as variâncias são

desconhecidas e não são supostas iguais, são utilizados testes aproximados uma vez que a variável:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - \theta}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

não tem, a rigor, distribuição t de Student.

Foram propostos vários procedimentos alternativos a serem utilizados para testar $H_0: \mu_1 - \mu_2 = \theta$ em condições de heterocedasticia, isto é, quando as variâncias são desconhecidas e supostas diferentes.

2.2.3.1 - Procedimentos alternativos baseados no cálculo do número de graus de liberdade

BROWNLEE (1965) descreve o procedimento a ser adotado para testar a hipótese da nulidade $H_0: \mu_1 - \mu_2 = \theta$ quando as variâncias são desconhecidas e não são supostas iguais. Este procedimento consiste na comparação de

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - \theta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

com o valor crítico t_0 , obtido a partir da tabela de distribuição t de Student, com f' graus de liberdade, que é um valor aproximado, a um nível de significância α .

O número de graus de liberdade f' , descrito por BROWNLEE (1965), baseado em trabalhos de WELCH (1937, 1947), é obtido inicialmente para o caso geral em que se supõe S_i^2 ($i=1,2,\dots,k$) variâncias amostrais independentes com f_i graus de liberdade e valores esperados σ_i^2 . O autor define ainda S^2 , como uma combinação linear das variáveis amostrais S_i^2 , isto é:

$$S^2 = a_1 S_1^2 + a_2 S_2^2 + \dots + a_k S_k^2 = \sum_{i=1}^k a_i S_i^2$$

onde a_i são constantes conhecidas.

Desta forma verifica-se que

$$E[S^2] = E\left[\sum_{i=1}^k a_i S_i^2\right] = \sum_{i=1}^k a_i E[S_i^2] = \sum_{i=1}^k a_i \sigma_i^2.$$

Como

$$S_i^2 \sim \frac{\sigma_i^2 \chi^2(f_i)}{f_i} \quad \text{e} \quad V\left[\frac{\chi^2(f_i)}{f_i}\right] = \frac{2}{f_i}$$

tem-se que:

$$V(S_i^2) = V\left[\frac{\sigma_i^2 \chi^2(f_i)}{f_i}\right] = (\sigma_i^2)^2 V\left[\frac{\chi^2(f_i)}{f_i}\right] = \frac{2(\sigma_i^2)^2}{f_i} \quad [2.2.7]$$

e

$$V(S^2) = V\left[\sum_{i=1}^k a_i S_i^2\right] = \sum_{i=1}^k a_i^2 V[S_i^2] = \sum_{i=1}^k a_i^2 \frac{2(\sigma_i^2)^2}{f_i}, \quad [2.2.8]$$

aproximamos a distribuição de S^2 pela distribuição de S'^2 , onde S'^2 tem um certo número de graus de liberdade f' , tal que S^2 e S'^2 tenham a mesma média e variância. Então:

$$E[S'^2] = E[S^2] = \sum_{i=1}^k a_i \sigma_i^2$$

Da expressão [2.2.7] temos a variância da estimativa

$$V[S'^2] = [E(S'^2)]^2 \cdot \frac{2}{f'} = \left(\sum_{i=1}^k a_i \sigma_i^2\right)^2 \cdot \frac{2}{f'} \quad [2.2.9]$$

Igualando-se as expressões [2.2.8] e [2.2.9] pode-se verificar que:

$$f' = \frac{\left(\sum_{i=1}^k a_i \sigma_i^2\right)^2}{\sum_{i=1}^k a_i^2 \frac{(\sigma_i^2)^2}{f_i}}$$

mas como em geral não se conhece os valores de σ_i^2 , substituem-se esses valores pelas suas respectivas estimativas amostrais s_i^2 . Desta forma f' é estimado como se segue:

$$f' \approx \frac{\left(\sum_{i=1}^k a_i s_i^2 \right)^2}{\sum_{i=1}^k a_i^2 \frac{(s_i^2)^2}{f_i}}$$

Para o caso particular de duas amostras temos a combinação linear de duas variâncias s_1^2 e s_2^2 , cujos coeficientes são $a_1 = 1/n_1$ e $a_2 = 1/n_2$ respectivamente, onde n_1 e n_2 representam o número de observações em cada amostra. E, neste caso, realizando as devidas substituições temos então:

$$f' \approx \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\frac{\left(\frac{1}{n_1} \right)^2 (s_1^2)^2}{n_1 - 1} + \frac{\left(\frac{1}{n_2} \right)^2 (s_2^2)^2}{n_2 - 1}}$$

ou ainda:

$$f' \approx \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 - 1}} \quad [2.2.10]$$

Este procedimento também é descrito por ANDERSON e BANCROFT (1952), MORETTIN e BUSSAB (1981), VIEIRA (1981). Segundo ANDERSON e BANCROFT (1952), o número de graus de liberdade f' foi sugerido por SMITH (1936) e posteriormente generalizado por SATTERTHWAIT (1946). Este mesmo procedimento é citado por COCHRAN (1964), que o atribui a SMITH (1936).

Devemos notar que no caso particular onde $s_1^2 = s_2^2$ e $n_1 = n_2 = n$, a expressão [2.2.10] será

$$f' = 2n - 2 .$$

Existe coerência desse resultado com o número de graus de liberdade associado à comparação de duas médias através do teste t quando as variâncias são desconhecidas mas supostas iguais. Se o tamanho de uma das amostras tender ao infinito (por exemplo: $n_1 \rightarrow \infty$) teremos:

$$f' = n_2 - 1 .$$

É claro que, se $n_1 \rightarrow \infty$, $\bar{X}_1 \rightarrow \mu_1$. Logo existe coerência do número de graus de liberdade, porque corresponde ao teste de uma média com um valor especificado.

Um outro procedimento é descrito pelos autores ANDERSON e BANCROFT (1952), DIXON e MASSEY (1957), MONTGOMERY (1976), COSTA NETO (1977), BERQUÕ *et alii* (1980) e que apresenta uma diferença quanto ao cálculo do número de graus de liberdade f . Neste caso o número aproximado de graus de liberdade é igual a

$$f = \frac{\left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right]^2}{\frac{\left(\frac{s_1^2}{n_1} \right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2} \right)^2}{n_2 + 1}} - 2 \quad [2.2.11]$$

que é um caso particular para duas amostras, baseado em trabalho publicado por WELCH (1947). Esse artigo, refere-se a uma generalização do problema quando variâncias populacionais diferentes são envolvidas, obtendo-se então:

$$f = \frac{(\sum \lambda_i s_i^2)^2 - 2 \left(\frac{\sum \lambda_i^2 s_i^4}{f_i + 2} \right)}{\frac{\sum \lambda_i^2 s_i^4}{f_i + 2}}$$

onde $\lambda_i = \frac{1}{n_i}$ e $f_i = n_i - 1$. Para o caso particular de duas amostras, simplificando obtem-se a expressão [2.2.11].

Note-se que no caso particular em que $s_1^2 = s_2^2$ e $n_1 = n_2 = n$, temos:

$$f = 2n - 2 .$$

Neste caso obtemos um número de graus de liberdade superior àquele obtido através da expressão [2.2.10], consequentemente temos um valor crítico menor para o teste, o que aumenta a probabilidade de rejeitar H_0 . Quanto o tamanho de uma das amostras tender ao infinito (por exemplo: $n_1 \rightarrow \infty$) temos:

$$f = n_2 - 1 .$$

Verificamos a coerência com o resultado obtido através da expressão [2.2.10].

2.2.3.2 - Procedimento alternativo baseado no cálculo do valor crítico

Diversos autores, como ANDERSON E BANCROFT (1952), COCHRAN e COX (1957), STEEL e TORRIE (1960), OSTLE (1963), SNEDECOR e COCHRAN (1967), JOHN (1971), HOFFMANN (1980), descrevem um procedimento alternativo, quando as variâncias são desconhecidas e não supostas iguais.

Este procedimento foi apresentado por Cochran e Cox em 1950, sendo posteriormente discutido por COCHRAN (1964). O teste consiste na comparação de

$$t' = \frac{(\bar{x}_1 - \bar{x}_2) - \theta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

onde $\theta = \mu_1 - \mu_2$, com o valor crítico t'_0 aproximado, obtido a partir dos valores de t_1 e t_2 , da tabela da distribuição t de Student, com (n_1-1) e (n_2-1) graus de liberdade, respectivamente, ao nível de significância α , tal que:

$$t'_0 = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

$$\text{onde } w_1 = \frac{s_1^2}{n_1} \text{ e } w_2 = \frac{s_2^2}{n_2}.$$

No caso particular em que $s_1^2 = s_2^2$ e $n_1 = n_2$ temos $w_1 = w_2$. Desta forma obtemos um valor crítico

$$t'_0 = \frac{t_1 + t_2}{2},$$

maior do que nos casos anteriores, o que significa maior probabilidade de não rejeitar H_0 . Quando o tamanho de uma das amostras tender ao infinito (por exemplo: $n_1 \rightarrow \infty$), te

mos $t'_0 = t_2$, onde t_2 é um valor crítico obtido com $(n_2 - 1)$ graus de liberdade. Nestas condições, quando n_1 tender para infinito verificamos uma coerência entre os procedimentos alternativos descritos.

2.2.3.3 - Comparações de médias nos ensaios em "split-plot"

Pode ser necessário estabelecer a comparação de duas médias, com variâncias não-iguais, quando se procede a uma análise de variância em split-plot. Isto ocorre quando o pesquisador quer comparar duas médias de tratamentos principais, no mesmo tratamento secundário. O problema são as variâncias desiguais. Então, para a comparação de médias, é preciso calcular a variância ponderada, usando tanto a estimativa de variância dada pelo $QMRes(a)$, como a estimativa de variância dada pelo $QMRes(b)$. Essas variâncias são, por pressuposição, desiguais. Então, mesmo quando se aplica o teste de Tukey, tem sido indicado o uso da fórmula de Satterthwaite (PIMENTEL GOMES, 1982) ou uma adaptação na forma de proceder ao teste como indicada por Cochran (STEEL e TORRIE, 1960).

2.3 - Simulação de Amostras por Geração de Variáveis

2.3.1 - Geração de variáveis com distribuição normal reduzida

Através do processo de geração de dados foram simulados 216 mil pares de amostras. Para isso, foram utilizados dez computadores pessoais CP-200, da PROLÓGICA, pertencentes ao Centro de Ciências da UNIMEP. Os programas para a geração dos dados, em linguagem BASIC, estão apresentados no Apêndice (Programa 1 e Programa 2).

Para a geração das amostras com distribuição normal foi utilizado o processo baseado no teorema do limite central. De acordo com o teorema, uma soma de variáveis com distribuição não-especificada tende à distribuição normal quando o número de elementos somados tende para infinito. O processo consiste em gerar k números aleatórios r_i , independentes, com distribuição uniforme, cuja função densidade é dada por:

$$f(x) = \begin{cases} 1 & ; \quad 0 < x < .1 \\ 0 & ; \quad \text{para outros casos,} \end{cases}$$

com média $1/2$ e variância $1/12$ e a utilização da seguinte expressão:

$$Z = \left(\frac{12}{k}\right)^{1/2} \left(\sum_{i=1}^k R_i - \frac{k}{2} \right),$$

onde R_i é a variável com distribuição uniforme.

A variável Z assim definida, tem distribuição normal reduzida. Procedeu-se a geração de $k = 12$ números aleatórios r_i pois, segundo SHIMIZU (1975) a aproximação é boa e simplifica os cálculos, o que reduz o tempo e o trabalho de computação. A expressão:

$$Z = \left(\frac{12}{k}\right)^{1/2} \left(\sum_{i=1}^k R_i - \frac{k}{2} \right)$$

para $k = 12$, se reduz a:

$$Z = \sum_{i=1}^{12} R_i - 6.$$

2.3.2 - Geração de pares de amostras de variáveis com distribuição normal

De posse de valores z da variável aleatória Z , com distribuição normal reduzida, é fácil obter valores da variável aleatória X com distribuição normal de média μ e variância σ^2 :

$$X = \mu + Z \cdot \sigma$$

Então, para os 216 mil pares de amostras, tomamos $\mu_2 = 10$. A média da outra amostra (μ_1) ou era igual a média $\mu_2 = 10$ ou foi obtida pelo procedimento descrito em seguida.

Através da tabela de distribuição normal reduzida, obteve-se o valor de z , tal que:

$$P(Z > z) = 0,05 ,$$

o nível de significância adotado para teste. O valor de z dado na tabela é 1,645. Seja W a diferença entre as duas médias em comparação. Calcularam-se então valores de W , através de:

$$z = 1,645 = \frac{W}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

para todas as 12 combinações de σ_1^2 e σ_2^2 , apresentados na Tabela 2.2 e associados a todos os valores de n_1 e n_2 apresentados na Tabela 2.3.

Tabela 2.2 - Valores de σ_2^2 associados a cada um dos 3 valores de σ_1^2 , segundo a razão $\sigma_2^2/\sigma_1^2 = 1, 2, 3$ e 4.

σ_1^2	R A Z Ã O			
	1	2	3	4
4	4	8	12	16
8	8	16	24	32
16	16	32	48	64

Tabela 2.3 - Valores de n_1 e n_2 usados neste trabalho.

$N = n_1 + n_2 = 12$		$N = n_1 + n_2 = 24$		$N = n_1 + n_2 = 36$	
n_1	n_2	n_1	n_2	n_1	n_2
6	6	12	12	18	18
4	8	8	16	12	24
3	9	6	18	9	27

De posse dos valores x , da variável em análise, calcularam-se, para cada caso, os valores de z_i tais que:

$$P(Z < z_i) \approx \beta_i,$$

onde $\beta_i = 0,30; 0,20$ e $0,10$.

Esses valores de z_i conduziram às médias μ_1 , que foram arredondadas com uma única decimal. Tais médias μ_1 , foram então comparadas às médias $\mu_2 = 10$ através de teste estatístico. Quando as variâncias são conhecidas, o teste Z tem poder em torno de 0,70, 0,80 e 0,90.

Para melhor visualizar todo esse procedimento, vamos apresentar um exemplo. Então, dado $\mu_2 = 10$, $\sigma_1^2 = \sigma_2^2 = 4$, $n_1 = 3$ e $n_2 = 9$, temos:

$$z = 1,645 = \frac{w}{\sqrt{\frac{4}{3} + \frac{4}{9}}},$$

onde $w = 2,193$. Para:

$$P(Z < z) = 0,30$$

a tabela de distribuição normal reduzida fornece o valor $z = -0,52$. Então:

$$-0,52 = \frac{2,193 - (\mu_1 - \mu_2)}{\sqrt{\frac{4}{3} + \frac{4}{9}}}$$

onde $\mu_1 - \mu_2 = 2,887$, ou, arredondado, com apenas uma decimal, $\mu_1 - \mu_2 = 2,9$. Com $\mu_2 = 10$, temos $\mu_1 = 12,9$. A Figura 2.2 mostra o procedimento.

Figura 2.2 - Ilustração gráfica do procedimento para obtenção da média μ_1 .

É preciso frisar que mesmo quando se tomou $\mu_1 = \mu_2$, foram geradas as amostras com todas as combinações de variâncias, apresentadas na Tabela 2.2, associadas a todos os pares de valores n_1 e n_2 , apresentados na Tabela 2.3.

Então, dado um par das 9 combinações possíveis de n_1 e n_2 , foram considerados 4 valores de μ_1 , para cada uma das 12 combinações de σ_1^2 e σ_2^2 . Em cada um desses casos, foram gerados 500 pares de amostras. Desta maneira, foram geradas $9 \times 4 \times 12 \times 500 = 216$ mil pares de amostras.

2.3.3 - Estimativa do nível de significância

Considerando $\mu_1 = \mu_2 = 10$, foram gerados 500 pares de amostras, para cada caso em que $\sigma_1^2 = \sigma_2^2$ ($\sigma_1^2 = 4, 8$ e 16), associados aos 9 pares de valores estabelecidos para n_1 e n_2 . Consequentemente, foram gerados $3 \times 9 \times 500 = 13.500$ pares de amostras com $\mu_1 = \mu_2$ e $\sigma_1^2 = \sigma_2^2$.

Se, por construção, as médias são iguais, nos casos em que rejeitamos H_0 cometemos erro tipo I. Para estimar a probabilidade de ocorrência de erro tipo I, procedemos ao teste:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

usando as variâncias estabelecidas "a priori" e procedemos ao teste:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \right] \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

associado a $n_1 + n_2 - 2$ graus de liberdade, estimando a variância ponderada a partir das variâncias calculadas com os dados amostrais.

Depois, para cada conjunto de 500 pares de amostras, calculamos a proporção de vezes em que rejeitamos H_0 . Essa proporção é uma estimativa do nível de significância ($\hat{\alpha}$). Dado que o nível de significância é conhecido, o valor das estimativas permite estabelecer se o número de pares de amostras é suficiente e se os dados gerados estão coerentes com as pressuposições básicas. Então, este procedimento atua como um "ensaio em branco", pois dá condições para verificar a validade do método utilizado.

Considerando $\mu_1 = \mu_2 = 10$, foram gerados 500 pares de amostras, para a razão de variâncias $\sigma_2^2/\sigma_1^2 = 2, 3$ e 4 ($\sigma_1^2 = 4, 8$ e 16), associados aos 9 pares de valores de n_1 e n_2 . Então, foram gerados $3 \times 3 \times 9 \times 500 = 40.500$ pares de amostras com médias iguais e variâncias diferentes.

Como as médias são iguais, cometemos erro tipo I quando rejeitamos H_0 . Para estimar a probabilidade de cometer esse erro, procedemos ao teste:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

usando as variâncias estabelecidas. Procedemos também ao teste:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad ; \quad S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

associado a $n_1 + n_2 - 2$ graus de liberdade, estimando a variância ponderada a partir das variâncias amostrais.

Depois, procedemos ao teste t na forma:

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

estimando as variâncias a partir dos dados amostrais. Associamos, ao teste t, graus de liberdade:

a) na forma:

$$f' = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

b) na forma:

$$f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2$$

Calculamos ainda o valor crítico

$$t'_0 = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2}$$

onde: $w_1 = \frac{s_1^2}{n_1}$, $w_2 = \frac{s_2^2}{n_2}$ e t_1 e t_2 são valores críticos obtidos na tabela de distribuição t com (n_1-1) e (n_2-2) graus de liberdade, respectivamente, a um nível de significância α .

Para cada conjunto de 500 pares de amostras calculamos a proporção de vezes que rejeitamos H_0 , que estima o valor do nível de significância ($\hat{\alpha}$).

2.3.4 - Estimativa do poder do teste

Foram gerados pares de amostras em três situações, isto é, uma com média $\mu_2 = 10$ e outra com média μ_1 calculada para que o poder do teste Z ficasse em torno de 0,70, 0,80 e 0,90, respectivamente. Estabeleceu-se $\sigma_1^2 = \sigma_2^2$ ($\sigma_1^2 = 4, 8, 16$) mas, para cada par de valores foram associados 9 pares de valores para n_1 e n_2 . Como, em cada situação, foram gerados 500 pares de amostras, no total foram obtidas $3 \times 3 \times 9 \times 500 = 40.500$ pares de amostras.

Se, por construção, as médias são diferentes, o poder do teste é a probabilidade de rejeitar H_0 . Para estimar essa probabilidade calculamos:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

usando as variâncias estabelecidas e:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad ; \quad s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$$

associado a $n_1 + n_2 - 2$ graus de liberdade, estimando a variância ponderada a partir das variâncias amostrais.

Para cada conjunto de 500 pares de amostras obtivemos a proporção de vezes em que H_0 foi rejeitada. Esta proporção estima o poder do teste ($\hat{\pi}$).

Depois geramos pares de amostras nas mes-
mas situações, isto é, uma com média $\mu_2 = 10$ e outra com
média μ_1 calculada para que o poder do teste Z estivesse
próximo dos valores 0,70, 0,80 e 0,90, respectivamente. Fo
ram consideradas as razões de variâncias $\sigma_2^2/\sigma_1^2 = 2, 3, 4$
($\sigma_1^2 = 4, 8$ e 16) associados às razões de tamanhos de amos-
tras $n_2/n_1 = 1, 2, 3$ ($N = n_1 + n_2 = 12, 24$ e 36). No to-
tal, foram obtidas $3 \times 3 \times 3 \times 9 \times 500 = 121.500$ pares de amostras.

O poder do teste é obtido da mesma maneira
que no caso de variâncias iguais.

Calculamos ainda:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

usando as variâncias amostrais. Este valor de t pode ser associado aos seguintes graus de liberdade:

$$f' = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$$

ou, então

$$f = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 + 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 + 1}} - 2$$

Obtivemos ainda o valor crítico de t , isto é:

$$t'_0 = \frac{w_1 t_1 + w_2 t_2}{w_1 + w_2} \quad ; \quad w_1 = \frac{s_1^2}{n_1} \quad \text{e} \quad w_2 = \frac{s_2^2}{n_2} .$$

A proporção de vezes que se rejeitou H_0 , para cada conjunto de 500 pares de amostras, dá a estimativa do poder do teste ($\hat{\pi}$).

3. RESULTADOS

Foram propostas diversas formas de proceder ao teste $\bar{t}$ para testar a hipótese de igualdade de duas médias, quando as variâncias são desconhecidas e não são supostas iguais. Neste trabalho fizemos uma comparação desses procedimentos, usando dados de amostras geradas por simulação. É importante frisar que foram simulados 216 mil pares de amostras. Todos esses pares tinham sempre uma amostra com média $\mu_2 = 10$, mas as médias da outra variavam, bem como a razão de variâncias, a razão de tamanhos de amostras e o total de elementos nas duas amostras.

As variáveis geradas por computador, com distribuição uniforme, conduziram às variáveis com distribuição normal que geraram os 216 mil pares de amostras.

Conforme mostram os programas para computador apresentados no Apêndice (Programa 1 e Programa 2), não foram impressos os valores amostrais, devido ao grande volume de valores gerados.

De acordo com a descrição feita no capítulo MATERIAL E MÉTODOS, para cada conjunto de 500 pares de amostras, de médias e variâncias iguais, foram estimados os níveis de significância. Tais estimativas, segundo o valor da variância e a razão de tamanhos de amostras estão na Tabela 1 (Apêndice).

Na Tabela 2 (Apêndice) estão apresentadas as estimativas dos níveis de significância, para cada conjunto de 500 pares de amostras de médias iguais, segundo a razão de tamanho de amostras, razão de variâncias e procedimento de teste utilizado.

Os valores estabelecidos para μ_1 , fixando o poder do teste Z em torno de 0,70, 0,80 e 0,90, segundo a razão dos tamanhos de amostras e os valores iguais das variâncias, estão apresentados na Tabela 3 (Apêndice).

As estimativas de poder do teste, nos casos em que as médias são diferentes mas as variâncias são iguais, estão apresentadas nas Tabelas 4, 5 e 6 do Apêndice, para poder do teste Z em torno de 0,70, 0,80 e 0,90 respectivamente.

Fixando o poder do teste em torno de 0,70, 0,80 e 0,90, foram obtidos os valores de μ_1 , segundo a razão dos tamanhos de amostras e a razão de variâncias. Estes valores estão apresentados na Tabela 7 (Apêndice).

As Tabelas 8, 9 e 10 do Apêndice apresentam as estimativas do poder do teste, para cada conjunto de 500 pares de amostras de médias diferentes, segundo a razão de tamanhos de amostras, razão de variâncias e o procedimento de teste utilizado, considerando poder do teste Z em torno de 0,70, 0,80 e 0,90, respectivamente.

4. DISCUSSÃO

A Tabela 1 (Apêndice) mostra estimativas do nível de significância, considerando pares de amostras com médias e variâncias iguais. Essas estimativas se referem tanto ao cálculo de z , isto é, ao teste de igualdade de médias considerando variâncias conhecidas, como o cálculo do valor usual de t , isto é, do teste de igualdade de médias considerando a estimativa da variância ponderada das variâncias amostrais. As estimativas do nível de significância estão em torno de 0,05, uma vez que procedemos aos testes usando $\alpha = 0,05$. Os resultados comprovam que o número de pares de amostras, utilizadas neste trabalho, foi suficiente para obter boas estimativas de α . Consequentemente, o processo utilizado deve levar a estimati

vas razoáveis de poder do teste. Os valores de $\hat{\alpha}_t$ e $\hat{\alpha}_z$ são muito próximos porque o valor de t é correto, quando as variâncias são iguais.

Na Tabela 2 (Apêndice) estão as estimativas do nível de significância, considerando pares de amostras com médias iguais e variâncias desiguais. Tais estimativas se referem ao cálculo de z , de t na forma usual e de t nas formas propostas por WELCH (1937) e SATTERTHWAIT (1946), por WELCH (1947) e ASPIN (1948) e por COCHRAN (COCHRAN e COX, 1957; COCHRAN, 1964). A Tabela 2 mostra que as estimativas dos níveis de significância, obtidos através do teste t usual e do teste t , nas formas propostas por Welch-Satterthwaite (W-S) e Aspin-Welch (A-W), são bastante próximas aos valores obtidos quando usamos a variável Z , nos casos em que as amostras têm tamanhos iguais. Os níveis de significância, estimados para o teste t , na forma proposta por Cochran (C) são ligeiramente inferiores.

Quando as amostras são de tamanhos diferentes, as estimativas dos níveis de significância, obtidos pelo teste t , nas formas propostas por Welch-Satterthwaite e Aspin-Welch são praticamente iguais e próximas às estimativas obtidas pelo teste Z . As estimativas dos níveis de significância, usando o teste t , na forma proposta por

Cochran, são sempre ligeiramente inferiores. O teste t usual leva a estimativas de nível de significância tanto menores quanto maior é a razão de variâncias. Estes resultados são similares aos obtidos por HSU (1938), citado por SCHEFFÉ (1959), que estimou nível de significância para teste t bilateral, considerando $\alpha = 0,05$ e tamanhos de amostras (15 e 5), (5 e 3) e (7 e 7), com razões de variâncias 1, 2 e 5, além de outros muito diferentes dos utilizados neste trabalho.

As Tabelas 4, 5 e 6 do Apêndice apresentam as estimativas de poder do teste, quando as variâncias são iguais. Devemos lembrar que os valores exatos do poder desses testes, apresentados na Tabela 3 (Apêndice) estão em torno de 0,70, 0,80 e 0,90, respectivamente. Nas Tabelas 8, 9 e 10 do Apêndice apresentamos as estimativas do poder do teste, quando as razões de variância assumem valores 2, 3 e 4. Os valores exatos do poder desses testes constam da Tabela 7 (Apêndice), mas estão sempre em torno de 0,70, 0,80 e 0,90, respectivamente.

Para melhor visualização dos resultados, construímos os gráficos apresentados nas Figuras 4.1, 4.2 e 4.3, para os casos em que o poder do teste está em torno de 0,70, 0,80 e 0,90, respectivamente.

Figura 4.1 - Poder (P) de teste, segundo a razão (R) de variâncias [$\sigma_2^2/\sigma_1^2 = 1, 2, 3, 4$; para os casos em que $\sigma_1^2 = 4$ (A), 8(B) e 16(C)], segundo a razão de tamanho de amostras ($n_2/n_1 = 1, 2, 3$) e número total de elementos das amostras: $N = n_1 + n_2 = 12, 24$ e 36 . ($\pi \approx 0,70$).

Figura 4.1 - continuação

Figura 4.1 - continuação

Figura 4.2 - Poder (P) de teste, segundo a razão (R) de variâncias [$\sigma_2^2/\sigma_1^2 = 1, 2, 3, 4$; para os casos em que $\sigma_1^2 = 4$ (A), 8(B) e 16(C)], segundo a razão de tamanho de amostras ($n_2/n_1 = 1, 2, 3$) e número total de elementos das amostras: $N = n_1 + n_2 = 12, 24$ e 36 . ($\pi \approx 0,80$).

Figura 4.2 - continuação

Figura 4.2 - continuação

Figura 4.3 - Poder (P) de teste, segundo a razão (R) de variâncias [$\sigma_2^2/\sigma_1^2 = 1, 2, 3, 4$; para os casos em que $\sigma_1^2 = 4$ (A), 8(B), e 16(C)], segundo a razão de tamanho de amostras ($n_2/n_1 = 1, 2, 3$) e número total de elementos das amostras: $N = n_1 + n_2 = 12, 24$ e 36 . ($\pi \approx 0,90$).

Figura 4.3 - continuação

Figura 4.3 - continuação

Tais gráficos mostram que, nos casos em que os tamanhos das amostras são iguais, as estimativas de poder de teste, usando o teste t usual, e o teste t tanto na forma proposta por Welch-Satterthwaite (W-S) e Aspin-Welch (A-W), foram praticamente iguais. No entanto, os resultados obtidos usando o teste t , na forma proposta por Cochran (C) foram sempre ligeiramente inferiores. Então, é razoável optar pelo teste t usual, que corresponde a uma análise de variância com um critério de classificação, no caso particular de duas amostras de tamanhos iguais, desde que a razão de variâncias não exceda ao valor 4. Aliás, esta recomendação foi feita por PIMENTEL GOMES (1982). Devemos notar que esta recomendação atende finalidades práticas, mas não está dentro dos pressupostos teóricos estritos de uma análise de variância. De acordo com SCHEFFÉ (1959) as pressuposições mínimas para uma análise de variância são de que os erros sejam variáveis aleatórias independentemente distribuídas e independentes, com média zero e variância σ^2 . A distribuição dessas variáveis deve ser normal ou aproximadamente normal. Os efeitos da não-normalidade da distribuição da variável e da desigualdade de variâncias sobre os resultados da análise são discutidos pelo autor.

Os gráficos apresentados nas Figuras 4.1, 4.2 e 4.3 também mostram que os resultados para poder de teste melhoram quando o total de elementos, nas duas amostras, aumenta. Isto era esperado, uma vez que aumentando de maneira adequada o tamanho das amostras obtêm-se melhores estatísticas.

Nos casos em que a razão dos tamanhos de amostras é 2 ou 3, os resultados das estimativas de poder do teste, quando se aplica o teste t, na forma proposta por Welch-Satterthwaite e Aspin-Welch mostram ligeira superioridade da forma apresentada por Aspin-Welch, mas não são afetadas pelo aumento da razão de variâncias. Essa superioridade no entanto diminui com o aumento do total de elementos nas duas amostras. Parece razoável esperar que, com o aumento do tamanho das amostras, o poder dos testes, obtido pelo cálculo dessas duas estatísticas, tenha a mesma tendência. Então, de acordo com os resultados obtidos, podemos proceder ao teste t, na forma proposta por Aspin-Welch, quando dispomos de amostras pequenas, com razão de tamanho de amostras $n_2/n_1 = 3$, o que implica num teste mais poderoso, ou seja, um teste em que a probabilidade de cometer erro tipo II seja menor.

Os resultados obtidos para o teste t, na forma proposta por Cochran, sugerem que este teste tem po

der menor do que os testes t , nas formas propostas por Welch-Satterthwaite e Aspin-Welch. No entanto, o poder do teste t , na forma proposta por Cochran, melhora quando aumenta o número total de elementos nas duas amostras.

O comportamento da estatística t , na forma usual, é altamente afetado pela razão de tamanhos de amostras. Se $n_2/n_1 = 2, 3$, as estimativas do poder de teste são baixas e tendem a diminuir com o aumento da razão de variâncias. Então, de acordo com os resultados obtidos neste trabalho, quando as variâncias são desiguais e os tamanhos são diferentes, não deve ser aplicado o teste t , na forma usual, aos dados coletados, na forma original. Poderíamos aplicar o teste após transformação de variáveis, desde que essa transformação estabilize a variância. O procedimento para transformação de variáveis é apresentado, por exemplo, em STEEL e TORRIE (1960), SNEDECOR e COCHRAN (1967) e PIMENTEL GOMES (1982).

O principal enfoque deste trabalho é estimar probabilidades de erro tipo I e erro tipo II, quando se prócede ao teste t , em diferentes formas, e comparar os resultados. Então é importante discutir o problema de decisão, envolvido na escolha do teste.

Para isso, vamos lembrar que erro tipo I consiste em rejeitar a hipótese da nulidade H_0 quando H_0

é verdadeira. No caso particular de comparação de dois grupos — um grupo tratado com um grupo controle, H_0 verdadeira significa que o tratamento não tem efeito. A aplicação de um teste estatístico, neste caso, a um determinado nível de significância α , pode levar a rejeitar H_0 . Esta decisão estaria errada, mas levaria à conclusão de que o tratamento tem efeito. Esta conclusão implicaria em gastos ou mudança de procedimento, totalmente desnecessários.

Por outro lado, devemos lembrar que erro tipo II consiste em não rejeitar H_0 quando H_0 é falsa. Na comparação de um grupo tratado com um grupo controle, H_0 falsa significa que o tratamento tem efeito. No entanto a aplicação de um teste estatístico, a um dado nível de significância, pode levar à não rejeição de H_0 . Esta decisão errada conduziria à conclusão de que o tratamento não tem efeito. Tal conclusão implicaria em perda real de benefício.

Conforme discute MEIER (1975), é elucidativo, no estudo de problemas que utilizam inferência estatística, fazer distinção entre decisões e conclusões. O autor lembra que é possível incorporar uma das palavras na idéia expressa pela outra. Assim, por exemplo, podemos decidir que uma conclusão não é razoável ou concluir que a decisão não foi correta. Entretanto, é preciso manter a

distinção entre decisão e conclusão.

É claro que, quando procedemos a um teste estatístico, estamos interessados na conclusão. No entanto, dado um problema, em geral podemos optar por um teste, entre os diversos testes conhecidos. Ora, os testes têm poder diferente. Então, optar por um teste também significa optar por maior ou menor poder.

É preciso lembrar que poder do teste significa rejeitar H_0 quando H_0 é falsa. Então, um teste com maior poder leva a rejeitar H_0 maior número de vezes, fixado o nível de significância.

Na escolha do teste é preciso levar em consideração os prejuízos que podem advir, dado o problema, se incorrermos em erro tipo I e erro tipo II, conforme salienta HADLEY (1967).

Então, no caso de comparação de médias de duas amostras, considerando razão de variâncias $\sigma_2^2/\sigma_1^2 = 2, 3, 4$ e razão de tamanhos de amostras $n_2/n_1 = 2, 3$, a escolha deve ser feita em função da importância relativa do erro tipo I e erro tipo II. O teste t na forma proposta por Aspin-Welch tem maior poder de teste. Entretanto, o teste t na forma proposta por Welch-Satterthwaite está associado a menor probabilidade de cometer erro tipo I.

O teste t , na forma proposta por Cochran é mais seguro, nos casos em que incorrer em erro tipo I levar a maior prejuízo, se as amostras tiverem o mesmo tamanho. Quando a razão de tamanhos de amostras são de ordem $n_2/n_1 = 2, 3$, o teste t usual apresenta menor probabilidade de cometer erro tipo I do que a forma proposta por Cochran, conforme os gráficos apresentados na Figura 4.4. No entanto, este teste tem poder de teste muito menor do que o teste na forma proposta por Cochran.

Finalmente, podemos sugerir que a comparação, feita neste trabalho, seja estendida para valores de razão de variâncias, superiores a 4, e outros valores de razão de tamanhos de amostras, considerando variâncias diferentes.

Figura 4.4 - Nível de significância (N.S.), segundo a razão (R) de variâncias [$\sigma_2^2/\sigma_1^2 = 1, 2, 3, 4$; para os casos em que $\sigma_1^2 = 4(A), 8(B)$ e $16(C)$], segundo a razão de tamamos ($n_2/n_1 = 1, 2, 3$) e numero total de elementos das amostras: $N = n_1 + n_2 = 12, 24$ e 36 .

Figura 4.4 - continuação

Figura 4.4 - continuação

5. CONCLUSÕES

Neste trabalho comparamos os resultados da aplicação do teste t , na forma usual, com os resultados da aplicação do teste t , nas formas propostas por Welch-Satterthwaite, Aspin-Welch e Cochran, quando a razão de variâncias assume valores iguais a 2, 3 e 4.

Essas comparações permitem as seguintes conclusões:

- 5.1 - Se as amostras são de mesmo tamanho, os resultados obtidos através do teste t usual, e do teste t , nas formas propostas por Welch-Satterthwaite e Aspin-Welch são praticamente iguais. Então, nesses casos, é recomendável proceder ao teste t , na forma usual, porque os cálculos são mais simples.

- 5.2 - Se a razão de tamanhos de amostras for da ordem de 2 e 3, o teste t, na forma usual, não leva a bons resultados. Nesses casos não é recomendável proceder ao teste t usual, a menos que a aplicação do teste seja precedida por uma transformação de variáveis que estabilize a variância.
- 5.3 - Se a razão de tamanhos de amostras for da ordem de 2 e 3, o teste t, na forma proposta por Aspin-Welch, leva a resultados ligeiramente melhores em termos de poder de teste aqueles obtidos quando se procede ao teste t, na forma proposta por Welch-Satterthwaite, o que implica em um teste de maior poder, consequentemente menor probabilidade de cometer erro tipo II. Essa diferença, entre as duas formas de teste, diminui com o aumento do total de elementos nas duas amostras. No entanto, se houver interesse em um teste associado a menor probabilidade de cometer erro tipo I, é razoável proceder ao teste t, na forma proposta por Welch-Satterthwaite, o que não implica em perda considerável do poder.

5.4 - Se a razão de tamanhos de amostra for da ordem de 2 e 3, o teste t, na forma proposta por Cochran, tem poder menor, que cresce com o aumento do total de elementos nas duas amostras. Este teste, no entanto, pode ser mais seguro se, no problema dado, existe prejuízo real em cometer erro tipo I.

6. LITERATURA CITADA

ALTMAN, D.G., 1982. Statistics and ethics in medical research: Collecting and screening data. In: GORE, S.M. e D.G. ALTMAN. Statistics in Practice. Londres, British Medical Association. 100 p.

ANDERSON, R.L. e T.A. BANCROFT, 1952. Statistical Theory in Research. Nova York, McGraw-Hill Book Company. 399 p.

ARMITAGE, P., 1982. The role of randomization in clinical trials. Statistics in Medicine, 1: 345-52.

ASPIN, A.A., 1948. An examination and further development of a formula arising in the problem of comparing two mean values. Biometrika, 35: 88-96.

- BERQUÔ, E.S.; J.M.P. SOUZA e S.L.D. GOTLIEB, 1980. Bio-estatística. São Paulo, E.P.U. 325 p.
- BOX, J.F., 1981. Gosset, Fisher, and the t distribution. The American Statistician, 35: 61-6.
- BROWN, B.W. e M. HOLLANDER, 1977. Statistics: A Biomedical Introduction. Nova York, John Wiley. 456 p.
- BROWNLIE, K.A., 1965. Statistical Theory and Methodology in Science and Engineering. Nova York, John Wiley. 590 p.
- CAMPOS, H., 1983. Estatística Experimental Não-Paramétrica. 4ª ed. Piracicaba, São Paulo, USP/ESALQ. 349 p.
- COCHRAN, W.G. e G.M. COX, 1957. Experimental Designs. 2ª ed. Nova York, John Wiley. 611 p.
- COCHRAN, W.G., 1964. Approximate significance levels of the Behrens-Fisher test. Biometrics, 20: 191-5.
- COSTA NETO, P.L.O., 1977. Estatística. São Paulo, Edgar Blucher. 264 p.
- DIXON, W.J. e F.J. MASSEY, 1957. Introduction to Statistical Analysis. 2ª ed. Nova York, McGraw-Hill Book Company. 488 p.
- EISENHART, C., 1979. On the transition from "Student's" z to "Student's" t. The American Statistician, 33: 6-10.

- FINNEY, D.J., 1955. Experimental Design and its Statistical Basis. Chicago, The University of Chicago Press. 169 p.
- FISHER, R.A., 1932. Statistical Methods for Research Workers. 4ª ed. Edinburgh, Oliver e Boyd.
- GEHAN, E. e E. FREIREICH, 1974. Non-randomized controls in cancer clinical trials. New England Journal of Medicine, 290: 198-203.
- GIBRA, I.N., 1973. Probability and Statistical Inference for Scientists and Engineers. New Jersey, Prentice-Hall Inc. 596 p.
- HADLEY, G., 1967. Introduction to Probability and Statistical Decision Theory. San Francisco, Holden-Day.
- HOFFMANN, R., 1980. Estatística para Economistas. São Paulo, Livraria Pioneira Editora. 379 p.
- HOLLANDER, M. e D.A. WOLFE, 1973. Nonparametric Statistical Methods. Nova York, John Wiley. 503 p.
- HSU, P.L., 1938. Contribution to the theory of Student's t-test as applied to the problem of two samples. Stat. Research Memoirs, 2: 1-24. Apud: SCHEFFÉ, H., 1959.
- JOHN, P.W.M., 1971. Statistical Design and Analysis of Experiments. Nova York, The MacMillan Company. 356 p.

- LEHMANN, E.L., 1975. Nonparametrics: Statistical Methods Based on Ranks. San Francisco, Holden-Day. 457 p.
- MEIER, P., 1975. Statistical and medical experimentation. Biometrics, 31: 511-29.
- MONTGOMERY, D.C., 1976. Design and Analysis of Experiments. Nova York, John Wiley. 418 p.
- MORETTIN, P.A. e BUSSAB, W.O., 1981. Métodos Quantitativos para Economistas e Administradores. São Paulo, Atual. 308 p.
- OSTLE, B., 1963. Statistics in Research. 2ª ed., Ames, Iowa, The Iowa State University Press. 588 p.
- PIMENTEL GOMES, F., 1982. Curso de Estatística Experimental. 10ª ed. Piracicaba, São Paulo, USP/ESALQ.
- SATTERTHWAITE, F.E., 1946. An approximate distribution of variance components. Biometrics Bull., 2: 110-4.
- SCHEFFÉ, H., 1959. The Analysis of Variance. Nova York, John Wiley, 477 p.
- SHIMIZU, T., 1975. Simulação em Computador Digital. São Paulo. Edgar Blucher. 107 p.
- SMITH, H.F., 1936. The problem of comparing the results of two experiments with unequal errors. Journal of the Council of Scientific and Industrial Research, 9: 211-2.

- SNEDECOR, G.W. e W.G. COCHRAN, 1967. Statistical Methods. 6ª ed. Ames, Iowa, The Iowa State University Press. 593 p.
- STEEL, R.G.D. e J.H. TORRIE, 1960. Principles and Procedures of Statistics. Nova York, McGraw-Hill Book Company. 481 p.
- STUDENT, 1908. The probable error of a mean. Biometrika, 6: 1-25.
- STUDENT, 1931. The Lanarkshire milk experiment. Biometrika, 23: 398-406.
- VIEIRA, S., 1981. Introdução à Bioestatística. Rio de Janeiro, Campus. 294 p.
- WALLACE, D.L., 1980. The Behrens-Fisher and Fieller-Creasy Problem. In: FIENBERG, S.E. e D.V. HINKLEY, eds. R.A. FISHER: An Appreciation. Nova York, Springer-Verlag. p. 119-47. (Lecture notes in Statistics, 1).
- WELCH, B.L., 1937. The significance of the difference between two means when the population variances are unequal. Biometrika, 29: 350-62.
- WELCH, B.L., 1947. The generalization of "Student's problem when several different population variances are involved. Biometrika, 34: 28-35.

7. APÊNDICE

PROGRAMAS EM BASIC E TABELAS

PROGRAMA 1 (PROG1)

```
5  REM "PROG1"
6  PRINT "ESTIMATIVAS DO NIVEL DE SIGNIFICANCIA E"
7  PRINT "PODER DO TESTE PARA O CASO DE VARIANCIAS"
8  PRINT "IGUAIS."
10 LET IZ=0
20 LET IT=0
30 PRINT "N. DE ELEMENTOS DA AMOSTRA 1 - N1"
40 INPUT N1
50 PRINT N1
60 PRINT "N. DE ELEMENTOS DA AMOSTRA 2 - N2"
70 INPUT N2
80 PRINT N2
90 PRINT "VARIANCIA 1"
100 INPUT SA
110 PRINT SA
120 PRINT "VARIANCIA 2"
130 INPUT SB
140 PRINT SB
150 PRINT "MEDIA 1"
160 INPUT M1
170 PRINT M1
180 PRINT "MEDIA 2"
190 INPUT M2
200 PRINT M2
210 PRINT "N. DE PARES DE X E DE Y"
220 INPUT KX
230 PRINT KX
240 PRINT "T-CRITICO COM";(N1+N2-2);"G.L."
250 INPUT T0
260 PRINT T0
270 LET Z0=1.645
275 RAND
280 FAST
290 FOR K=1 TO KX
300 LET W=N1
310 LET VN=SQR(SA)
320 LET MN=M1
330 GOSUB 600
```

```
340 LET XB=YB
350 LET S1=S2
360 LET W=N2
370 LET VN=SQR(SB)
380 LET MN=M2
390 GOSUB 600
420 LET DF=XB-YB
430 LET N=N1+N2-2
440 LET Z=DF/SQR(SA/N1+SB/N2)
450 IF Z<Z0 THEN GOTO 470
460 LET IZ=IZ+1
470 LET T=DF/SQR(((S1+S2)/N)*(1/N1*1/N2))
480 IF T<T0 THEN GOTO 500
490 LET IT=IT+1
510 NEXT K
520 PRINT "N. DE VEZES QUE SE REJEITA H0"
530 PRINT "TEST Z :";IZ
535 PRINT "TEST T :";IT
540 STOP
600 REM SIMULACAO
610 LET SN=0
615 LET QN=0
620 FOR L=1 TO W
630 LET SQ=0
640 FOR I=1 TO 12
650 LET R=RND
660 LET SR=SR+R
670 NEXT I
680 LET Y=(SR-6)*VN+MN
690 LET SN=SN+Y
700 LET QN=QN+Y*Y
710 NEXT L
720 LET YB=SN/W
730 LET S2=(QN-SN*SN/W)
740 RETURN
```

PROGRAMA 2 (PROG2)

```
5  REM "PROG2"
6  PRINT "ESTIMATIVAS DO NIVEL DE SIGNIFICANCIA E"
7  PRINT "PODER DO TESTE PARA O CASO DE VARIANCIAS"
8  PRINT "DESIGUAIS."
9  DIM T(40)
10 LET IT=0
11 LET IZ=0
12 LET IB=0
14 LET IC=0
20 LET BW=0
30 LET DW=0
50 PRINT "N. DE ELEMENTOS DA AMOSTRA 1 - N1"
60 INPUT N1
70 PRINT N1
80 PRINT "N. DE ELEMENTOS DA AMOSTRA 2 - N2"
90 INPUT N2
100 PRINT N2
110 PRINT "VARIANCIA 1"
120 INPUT SA
130 PRINT SA
140 PRINT "VARIANCIA 2"
150 INPUT SB
160 PRINT SB
170 PRINT "MEDIA 1"
180 INPUT M1
190 PRINT M1
200 PRINT "MEDIA 2"
210 INPUT M2
220 PRINT M2
230 PRINT "N. DE PARES DE X E DE Y"
240 INPUT KX
250 PRINT KX
251 LET T(1)=6.3138
252 LET T(2)=2.9200
253 LET T(3)=2.3534
254 LET T(4)=2.1318
255 LET T(5)=2.0150
256 LET T(6)=1.9432
257 LET T(7)=1.8946
258 LET T(8)=1.8595
259 LET T(9)=1.8331
```

```
260 LET T(10)=1.8125
261 LET T(11)=1.7959
262 LET T(12)=1.7823
263 LET T(13)=1.7709
264 LET T(14)=1.7613
265 LET T(15)=1.7531
266 LET T(16)=1.7459
267 LET T(17)=1.7396
268 LET T(18)=1.7341
269 LET T(19)=1.7291
270 LET T(20)=1.7247
271 LET T(21)=1.7207
272 LET T(22)=1.7171
273 LET T(23)=1.7139
274 LET T(24)=1.7109
275 LET T(25)=1.7081
276 LET T(26)=1.7056
277 LET T(27)=1.7033
278 LET T(28)=1.7011
279 LET T(29)=1.6991
280 LET T(30)=1.6973
281 LET T(31)=1.6955
282 LET T(32)=1.6939
283 LET T(33)=1.6924
284 LET T(34)=1.6909
285 LET T(35)=1.6896
286 LET T(36)=1.6883
287 LET T(37)=1.6871
288 LET T(38)=1.6860
289 LET T(39)=1.6849
290 LET T(40)=1.6839
300 LET Z0=1.645
310 RAND
320 FAST
330 FOR K=1 TO KX
340 LET NN=N1
350 LET VN=SQR(SA)
360 LET MN=M1
370 GOSUB 840
380 LET XB=YB
390 LET S1=S2
400 LET NN=N2
410 LET VN=SQR(SB)
420 LET MN=M2
430 GOSUB 840
```

```

440 LET DF=XB-YB
450 LET N=N1+N2-2
460 LET Z=DF/SQR(SA/N1+SB/N2)
470 IF Z<Z0 THEN GOTO 490
480 LET IZ=IZ+1
490 LET TC=DF/SQR(S1/N1+S2/N2)
500 LET W1=S1/N1
510 LET W2=S2/N2
520 LET W=W1+W2
530 LET G1=N1-1
540 LET G2=N2-1
550 LET WQ=W*W
555 LET TI=DF/SQR(((G1*S1+G2*S2)/N)*(1/N1+1/N2))
556 IF TI<T(N) THEN GOTO 560
557 LET IT=IT+1
560 LET DB=(W1**2/G1)+(W2**2/G2)
570 LET B=INT((WQ/DB)+0.5)
580 LET DD=W1**2/(N1+1)+W2**2/(N2+1)
590 LET D=INT(((WQ/DD)-2)+0.5)
600 LET T1=T(N1-1)
610 LET T2=T(N2-1)
620 LET T0=(W1*T1+W2*T2)/W
630 IF TC<T(B) THEN GOTO 650
640 LET IB=IB+1
650 IF TC<T(D) THEN GOTO 670
660 LET ID=ID+1
670 IF TC<T0 THEN GOTO 685
680 LET IC=IC+1
685 NEXT K
690 PRINT "N. DE VEZES QUE SE REJEITA H0"
700 PRINT "TESTE Z :";IZ
702 PRINT "TEST T : ";IT
704 PRINT "WELCH-SATTERTHWAITE :";IB
706 PRINT "ASPIN-WELCH :";ID
708 PRINT "COCHRAN :";IC
740 STOP
840 REM SIMULACAO
850 LET SN=0
860 LET QN=0
870 FOR L=1 TO NN
880 LET SR=0
890 FOR I=1 TO 12
900 LET R=RND
910 LET SR=SR+R
920 NEXT I
930 LET Y=(SR-6)*VN+MN
940 LET SN=SN+Y
950 LET QN=QN+Y*Y
960 NEXT L
970 LET YB=SN/NN
980 LET S2=(QN-SN*SN/NN)/(NN-1)
990 RETURN

```


Tabela 1. - Estimativas do nível de significância para conjuntos de 500 pares de amostras, de médias iguais, segundo a variância e a razão de tamanho de amostras.

N	$\sigma_1^2 = \sigma_2^2$	n_1	n_2/n_1	$\hat{\alpha}_z$	$\hat{\alpha}_t$
12	4	6	1	0,060	0,058
		4	2	0,040	0,036
		3	3	0,054	0,050
		6	1	0,048	0,038
	8	4	2	0,064	0,066
		3	3	0,040	0,034
		6	1	0,056	0,062
	16	4	2	0,054	0,040
		3	3	0,052	0,048
		12	1	0,048	0,042
		4	8	0,044	0,042
		6	3	0,040	0,038
24	8	12	1	0,054	0,056
		8	2	0,044	0,048
		6	3	0,064	0,058
		12	1	0,056	0,052
	16	8	2	0,054	0,048
		6	3	0,064	0,066
		18	1	0,054	0,042
		4	12	0,056	0,066
	4	9	3	0,070	0,070
		18	1	0,038	0,038
		12	2	0,054	0,052
		9	3	0,056	0,046
36	8	18	1	0,052	0,052
		12	2	0,058	0,046
		9	3	0,046	0,052
		12	2	0,058	0,046
	16	9	3	0,046	0,052

Tabela 2 - Estimativas do nível de significância para conjuntos de 500 pares de amostras, de médias iguais, segundo a razão de tamanhos de amostras, razão de variâncias e procedimento de teste utilizado.

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\alpha}_z$	$\hat{\alpha}_t$	$\hat{\alpha}_{W-S}$	$\hat{\alpha}_{A-W}$	$\hat{\alpha}_C$
12	6	1	8	2	0,050	0,050	0,048	0,048	0,036
				4	3	0,062	0,066	0,064	0,046
				4	0,064	0,066	0,056	0,064	0,048
				2	0,062	0,054	0,042	0,048	0,032
				3	0,052	0,044	0,044	0,044	0,038
				4	0,060	0,064	0,058	0,062	0,044
				2	0,040	0,044	0,042	0,042	0,028
			16	3	0,046	0,058	0,054	0,054	0,034
				4	0,058	0,064	0,058	0,064	0,040
				2	0,038	0,026	0,030	0,032	0,018
				4	3	0,052	0,026	0,040	0,032
				4	0,052	0,018	0,042	0,046	0,028
12	4	2	8	2	0,054	0,034	0,044	0,044	0,032
				3	0,056	0,032	0,046	0,050	0,038
				4	0,048	0,022	0,036	0,038	0,026
				2	0,050	0,026	0,044	0,048	0,032
			16	3	0,056	0,024	0,042	0,044	0,026
				4	0,056	0,020	0,042	0,044	0,030
				2	0,048	0,022	0,052	0,060	0,032
				4	3	0,046	0,012	0,036	0,018
				4	0,058	0,014	0,038	0,046	0,026
				2	0,050	0,026	0,042	0,052	0,026
			16	3	0,044	0,018	0,040	0,052	0,030
				4	0,058	0,012	0,080	0,080	0,050
				2	0,046	0,026	0,062	0,066	0,042
12	3	3	8	3	0,050	0,018	0,048	0,052	0,032
				4	0,040	0,008	0,040	0,044	0,020

(continua)

Tabela 2 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\alpha}_z$	$\hat{\alpha}_t$	$\hat{\alpha}_{W-S}$	$\hat{\alpha}_{A-W}$	$\hat{\alpha}_C$
24	12	1	8	2	0,060	0,068	0,066	0,066	0,054
				4	3	0,048	0,052	0,052	0,044
				4	0,040	0,044	0,042	0,044	0,038
				2	0,060	0,056	0,056	0,056	0,044
				3	0,048	0,040	0,036	0,038	0,036
				4	0,042	0,048	0,046	0,048	0,046
				2	0,048	0,040	0,040	0,040	0,034
				16	3	0,038	0,034	0,034	0,032
				4	0,054	0,050	0,048	0,048	0,042
				2	0,056	0,028	0,046	0,048	0,038
				4	3	0,050	0,028	0,050	0,038
				4	0,032	0,018	0,030	0,030	0,026
24	8	2	8	2	0,036	0,022	0,030	0,032	0,028
				3	0,048	0,028	0,036	0,036	0,032
				4	0,066	0,026	0,056	0,056	0,046
				2	0,034	0,020	0,030	0,030	0,026
				16	3	0,036	0,018	0,036	0,028
				4	0,050	0,022	0,048	0,048	0,044
				2	0,058	0,030	0,058	0,060	0,044
				4	3	0,038	0,018	0,038	0,040
				4	0,056	0,024	0,044	0,046	0,036
				2	0,058	0,026	0,056	0,058	0,044
				3	0,042	0,012	0,042	0,042	0,036
				4	0,052	0,012	0,042	0,042	0,034
24	6	3	8	2	0,056	0,022	0,052	0,054	0,038
				16	3	0,044	0,016	0,040	0,044
				4	0,034	0,006	0,032	0,032	0,022

(continua)

Tabela 2 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\alpha}_z$	$\hat{\alpha}_t$	$\hat{\alpha}_{W-S}$	$\hat{\alpha}_{A-W}$	$\hat{\alpha}_C$
36	18	1	4	2	0,046	0,050	0,050	0,050	0,048
				3	0,042	0,038	0,038	0,038	0,038
				4	0,052	0,046	0,044	0,044	0,042
				2	0,046	0,042	0,042	0,042	0,038
				3	0,036	0,036	0,036	0,036	0,036
				4	0,056	0,054	0,050	0,050	0,050
			16	2	0,054	0,054	0,054	0,054	0,054
				3	0,056	0,054	0,052	0,054	0,048
				4	0,050	0,040	0,038	0,038	0,032
				2	0,058	0,040	0,064	0,064	0,060
				3	0,040	0,018	0,030	0,030	0,028
				4	0,054	0,020	0,048	0,050	0,046
36	12	2	8	2	0,052	0,026	0,056	0,056	0,046
				3	0,056	0,032	0,058	0,058	0,054
				4	0,050	0,024	0,056	0,056	0,046
				2	0,050	0,044	0,048	0,048	0,048
			16	3	0,046	0,024	0,046	0,048	0,042
				4	0,058	0,028	0,058	0,058	0,058
				2	0,050	0,026	0,048	0,048	0,044
			4	3	0,042	0,016	0,046	0,046	0,040
				4	0,064	0,018	0,054	0,054	0,048
				2	0,056	0,028	0,052	0,052	0,046
				3	0,052	0,016	0,052	0,052	0,048
36	9	3	8	4	0,056	0,020	0,050	0,050	0,044
				2	0,054	0,030	0,054	0,054	0,044
				16	3	0,048	0,024	0,048	0,046
				4	0,052	0,006	0,042	0,042	0,036

W-S = WELCH (1937) e SATTERTHWAITE (1946)

A-W = ASPIN (1948) e WELCH (1947)

C = COCHRAN (1957)

Tabela 3 - Valores estabelecidos para μ_1 , fixando o poder do teste em torno de 0,7, 0,8 e 0,9, segundo a razão dos tamanhos de amostras e os valores da variância.

N	n_1	n_2/n_1	σ_1^2	μ_1	$\pi \approx 0,70$	μ_2	$\pi \approx 0,80$	μ_2	$\pi \approx 0,90$
12	6	1	4	12,6	0,729	12,9	0,807	13,4	0,903
			8	13,6	0,712	14,1	0,807	14,8	0,901
			16	15,1	0,712	15,8	0,807	16,8	0,903
		2	4	12,7	0,712	13,1	0,813	13,6	0,901
			8	13,8	0,708	14,4	0,815	15,1	0,903
			16	15,4	0,712	16,1	0,802	17,2	0,901
	3	3	4	12,9	0,701	13,4	0,818	14,0	0,913
			8	14,1	0,701	14,7	0,802	15,6	0,908
			16	15,8	0,701	16,7	0,807	17,9	0,906
		1	4	11,8	0,712	12,1	0,823	12,4	0,901
			8	12,6	0,729	12,9	0,807	13,4	0,903
			16	13,6	0,712	14,1	0,807	14,8	0,901
24	8	2	4	11,9	0,708	12,2	0,815	12,6	0,913
			8	12,7	0,712	13,1	0,813	13,6	0,901
			16	13,8	0,708	14,4	0,815	15,1	0,903
		3	4	12,1	0,719	12,4	0,815	12,8	0,908
			8	12,9	0,701	13,4	0,818	14,0	0,913
			16	14,1	0,701	14,7	0,802	15,6	0,908
	18	1	4	11,5	0,729	11,7	0,818	12,0	0,913
			8	12,1	0,719	12,4	0,815	12,8	0,908
			16	12,9	0,701	13,4	0,818	14,0	0,913
		2	4	11,6	0,732	11,8	0,815	12,1	0,908
			8	12,2	0,712	12,5	0,805	13,0	0,913
			16	13,1	0,708	13,6	0,815	14,2	0,908
36	12	2	4	11,7	0,712	12,0	0,828	12,3	0,909
			8	12,4	0,712	12,8	0,823	13,2	0,901
			16	13,4	0,712	13,9	0,813	14,6	0,909

Tabela 4 - Estimativas do poder de teste para conjuntos de 500 pares de amostras de médias diferentes e variâncias iguais, segundo a razão de tamanho de amostras. ($\pi \approx 0,70$).

N	n_1	n_2/n_1	$\sigma_1^2 = \sigma_2^2$	$\hat{\pi}_z$	$\hat{\pi}_t$
12	6	1	4	0,708	0,666
			8	0,712	0,644
			16	0,706	0,656
		2	4	0,674	0,624
			8	0,704	0,650
			16	0,698	0,640
	3	3	4	0,706	0,660
			8	0,696	0,624
			16	0,724	0,678
		1	4	0,694	0,672
			8	0,726	0,706
			16	0,736	0,700
24	8	2	4	0,714	0,696
			8	0,712	0,658
			16	0,692	0,670
		3	4	0,708	0,686
			8	0,670	0,650
			16	0,708	0,700
	6	3	4	0,696	0,668
			8	0,710	0,680
			16	0,694	0,682
		1	4	0,720	0,698
			8	0,708	0,692
			16	0,720	0,690
36	12	2	4	0,720	0,692
			8	0,708	0,692
			16	0,720	0,692
	9	3	8	0,724	0,690
			16	0,742	0,712

Tabela 5 - Estimativas do poder do teste para conjuntos de 500 pares de amostras de médias diferentes e variâncias iguais, segundo a razão de tamanho de amostras ($\Pi \approx 0,80$).

N	n_1	n_2/n_1	$\sigma_1^2 = \sigma_2^2$	$\hat{\Pi}_Z$	$\hat{\Pi}_t$
12	6	1	4	0,794	0,738
			8	0,808	0,758
			16	0,824	0,780
			4	0,804	0,738
			8	0,834	0,790
			16	0,818	0,754
	4	2	4	0,826	0,752
			8	0,810	0,758
			16	0,804	0,750
			4	0,814	0,774
			8	0,794	0,768
			16	0,816	0,790
24	12	1	4	0,826	0,808
			8	0,810	0,776
			16	0,836	0,786
			4	0,828	0,792
			8	0,832	0,798
			16	0,810	0,796
	8	2	4	0,822	0,808
			8	0,712	0,784
			16	0,824	0,812
			4	0,818	0,790
			8	0,792	0,780
			16	0,822	0,802
36	12	2	4	0,844	0,814
			8	0,832	0,808
			16	0,804	0,790
	9	3	8	0,832	0,808
			16	0,804	0,790
			4	0,844	0,814

Tabela 6 - Estimativas do poder do teste para conjunto de 500 pares de amostras de médias diferentes e variâncias iguais, segundo a razão de tamanho de amostras. ($\pi = 0,90$).

N	n_1	n_2/n_1	$\sigma_1^2 = \sigma_2^2$	$\hat{\pi}_2$	$\hat{\pi}_t$
12	6	1	4	0,912	0,872
			8	0,880	0,844
			16	0,884	0,848
		2	4	0,914	0,864
			8	0,890	0,850
			16	0,898	0,874
	3	3	4	0,928	0,874
			8	0,904	0,874
			16	0,892	0,842
		1	4	0,914	0,890
			8	0,890	0,862
			16	0,912	0,896
24	8	2	4	0,916	0,896
			8	0,896	0,864
			16	0,900	0,868
		3	4	0,908	0,884
			8	0,890	0,890
			16	0,902	0,884
	6	3	4	0,902	0,884
			8	0,898	0,894
			16	0,928	0,900
		1	4	0,902	0,898
			8	0,896	0,890
			16	0,934	0,922
36	12	2	4	0,906	0,892
			8	0,920	0,896
			16	0,908	0,890
	9	3	8	0,920	0,896
			16	0,908	0,890
			4	0,906	0,892

Tabela 7 - Valores estabelecidos para μ_1 fixando o poder do teste em torno de 0,70, 0,80 e 0,90, segundo a razão de tamanhos de amostras e a razão de variâncias.

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	μ_1	$\Pi \approx 0,70$	μ_1	$\Pi \approx 0,80$	μ_1	$\Pi \approx 0,90$
12	4		4	2	13,1	0,708	13,6	0,815	14,2	0,908
				3	13,6	0,712	14,1	0,807	14,8	0,901
				4	14,0	0,708	14,6	0,810	15,4	0,904
				2	14,4	0,712	15,0	0,805	15,9	0,904
	6	1	8	3	15,1	0,712	15,8	0,807	16,8	0,903
				4	15,7	0,712	16,5	0,807	17,6	0,903
				2	16,2	0,708	17,1	0,807	18,3	0,901
				3	17,1	0,701	18,2	0,807	19,6	0,901
			16	4	18,0	0,708	19,1	0,802	20,7	0,901
				2	13,1	0,708	13,6	0,815	14,2	0,906
				3	13,5	0,715	14,0	0,813	14,7	0,908
				4	13,8	0,708	14,4	0,815	15,1	0,903
	2	8	2	14,4	0,712	15,0	0,805	15,9	0,904	
			3	14,9	0,708	15,6	0,805	16,6	0,904	
			4	15,4	0,712	16,1	0,802	17,2	0,901	
			2	16,2	0,708	17,1	0,807	18,3	0,901	
		16	3	16,9	0,705	17,9	0,802	19,3	0,903	
			4	17,6	0,708	18,7	0,807	20,2	0,903	
			2	13,3	0,715	13,8	0,815	14,4	0,904	
			3	13,6	0,712	14,1	0,807	14,8	0,901	
			4	13,9	0,715	14,4	0,802	15,2	0,903	
			2	14,6	0,705	15,3	0,807	16,2	0,903	
			3	15,1	0,712	15,8	0,807	16,8	0,903	
			4	15,5	0,712	16,3	0,810	17,4	0,906	
3	3	8	2	16,5	0,705	17,5	0,807	18,8	0,904	
			3	17,1	0,701	18,2	0,807	19,6	0,901	
			4	17,7	0,705	18,8	0,802	20,4	0,903	

(continua)

Tabela 7 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	μ_1	$\Pi \approx 0,70$	μ_1	$\Pi \approx 0,80$	μ_1	$\Pi \approx 0,90$
24	8		4	2	12,2	0,712	12,5	0,805	13,0	0,913
				3	12,6	0,729	12,9	0,807	13,4	0,903
				4	12,9	0,725	13,3	0,818	13,8	0,903
				2	13,1	0,708	13,6	0,815	14,2	0,906
		1	8	3	13,6	0,712	14,1	0,807	14,8	0,901
				4	14,0	0,708	14,6	0,810	15,4	0,904
				2	14,4	0,712	15,0	0,805	15,9	0,904
				3	15,1	0,712	15,8	0,807	16,8	0,903
			16	4	15,7	0,712	16,5	0,807	17,6	0,903
				2	12,2	0,712	12,5	0,805	13,0	0,913
				3	12,5	0,722	12,8	0,805	13,3	0,904
				4	12,7	0,712	13,1	0,813	13,6	0,901
		2	8	2	13,1	0,708	13,6	0,815	14,2	0,908
				3	13,5	0,715	14,0	0,813	14,7	0,908
				4	13,8	0,708	14,4	0,815	15,1	0,903
				2	14,4	0,712	15,0	0,805	15,9	0,904
			16	3	14,9	0,708	15,6	0,807	16,6	0,904
				4	15,4	0,712	16,1	0,802	17,2	0,901
				2	12,3	0,705	12,7	0,821	13,1	0,903
				3	12,6	0,729	12,9	0,807	13,4	0,903
		3	8	3	13,3	0,715	13,8	0,815	14,4	0,904
				4	13,6	0,712	14,1	0,807	14,8	0,901
				2	13,9	0,715	14,4	0,802	15,2	0,903
				3	14,6	0,705	15,3	0,807	16,2	0,903
			16	3	15,1	0,712	15,8	0,807	16,8	0,903
				4	15,5	0,712	16,3	0,810	17,4	0,906

(continua)

Tabela 7 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	μ_1	$\Pi \approx 0,70$	μ_1	$\Pi \approx 0,80$	μ_1	$\Pi \approx 0,90$
36	18	1	4	2	11,8	0,712	12,1	0,823	12,4	0,901
				3	12,1	0,719	12,4	0,815	12,8	0,908
				4	12,3	0,705	12,7	0,821	13,1	0,903
				2	12,6	0,729	12,9	0,807	13,4	0,903
		1	8	3	12,9	0,701	13,4	0,818	14,0	0,913
				4	13,3	0,715	13,8	0,815	14,4	0,904
				2	13,6	0,712	14,1	0,807	14,8	0,901
				3	14,1	0,701	14,7	0,802	15,6	0,908
	12	2	16	4	14,6	0,705	15,3	0,807	16,2	0,903
				2	11,8	0,712	12,1	0,823	12,4	0,901
				3	12,0	0,708	12,3	0,810	12,7	0,904
				4	12,2	0,712	12,5	0,805	13,0	0,913
		2	8	2	12,6	0,729	12,9	0,807	13,4	0,903
				3	12,9	0,725	13,3	0,818	13,8	0,903
				4	13,1	0,708	13,6	0,815	14,2	0,908
				2	13,6	0,712	14,1	0,807	14,8	0,901
36	9	3	16	3	14,0	0,708	14,6	0,810	15,4	0,904
				4	14,4	0,712	15,0	0,805	15,9	0,904
				2	11,9	0,712	12,2	0,818	12,6	0,916
				3	12,1	0,719	12,4	0,815	12,8	0,908
		3	8	4	12,3	0,729	12,6	0,818	13,0	0,903
				2	12,7	0,715	13,1	0,815	13,6	0,904
				3	12,9	0,701	13,4	0,818	14,0	0,913
				4	13,2	0,719	13,6	0,805	14,3	0,909
	6	2	16	2	13,8	0,712	14,3	0,802	15,1	0,906
				3	14,1	0,701	14,7	0,802	15,6	0,908
				4	14,5	0,712	15,1	0,805	16,0	0,903

Tabela 8 - Estimativas do poder do teste para conjuntos de 500 pares de amostras de médias diferentes, segundo a razão do tamanho de amostras, razão de variâncias e procedimento de teste utilizado ($\pi \approx 0,70$).

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_Z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
12	6	1	4	2	0,730	0,676	0,664	0,670	0,626
				3	0,778	0,684	0,672	0,676	0,628
				4	0,710	0,626	0,612	0,620	0,568
			8	2	0,712	0,650	0,636	0,652	0,592
				3	0,738	0,676	0,656	0,666	0,612
				4	0,714	0,664	0,638	0,652	0,588
			16	2	0,694	0,662	0,638	0,650	0,586
				3	0,708	0,664	0,636	0,648	0,586
				4	0,674	0,626	0,598	0,610	0,556
			4	2	0,686	0,552	0,616	0,634	0,558
				3	0,754	0,550	0,670	0,690	0,596
				4	0,724	0,492	0,666	0,682	0,592
12	4	2	8	2	0,708	0,580	0,646	0,660	0,574
				3	0,726	0,562	0,670	0,678	0,606
				4	0,740	0,524	0,680	0,698	0,610
			16	2	0,736	0,576	0,664	0,672	0,562
				3	0,686	0,514	0,600	0,608	0,540
				4	0,754	0,522	0,664	0,674	0,588
			4	2	0,694	0,518	0,568	0,608	0,468
				3	0,720	0,446	0,596	0,616	0,476
				4	0,716	0,392	0,608	0,628	0,514
			8	2	0,716	0,540	0,596	0,624	0,500
				3	0,698	0,452	0,598	0,626	0,476
				4	0,680	0,392	0,596	0,642	0,478
12	3	3	8	2	0,718	0,514	0,596	0,640	0,472
				3	0,718	0,438	0,626	0,654	0,494
				4	0,722	0,422	0,608	0,634	0,502
			16	3	0,718	0,438	0,626	0,654	0,494

(continua)

Tabela 8 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
24	12	1	8	2	0,698	0,676	0,672	0,674	0,646
				4	3	0,724	0,706	0,696	0,682
				4	0,728	0,698	0,692	0,696	0,678
				2	0,708	0,694	0,686	0,690	0,668
				3	0,718	0,668	0,656	0,658	0,644
				4	0,692	0,682	0,670	0,670	0,646
				2	0,716	0,674	0,668	0,672	0,652
			16	3	0,736	0,686	0,678	0,680	0,662
				4	0,718	0,704	0,690	0,692	0,666
				2	0,750	0,600	0,680	0,684	0,640
				4	3	0,724	0,584	0,706	0,676
				4	0,694	0,524	0,656	0,658	0,636
24	8	2	8	2	0,708	0,584	0,670	0,674	0,636
				3	0,718	0,590	0,696	0,698	0,660
				4	0,708	0,548	0,664	0,668	0,634
				2	0,710	0,578	0,664	0,670	0,622
			16	3	0,704	0,546	0,670	0,680	0,640
				4	0,710	0,532	0,678	0,686	0,654
				2	0,722	0,606	0,702	0,714	0,652
				4	3	0,718	0,506	0,668	0,680
				4	0,750	0,498	0,706	0,714	0,674
				2	0,740	0,580	0,676	0,684	0,642
			16	3	0,718	0,532	0,684	0,692	0,640
				4	0,694	0,428	0,652	0,654	0,614
				2	0,702	0,566	0,634	0,648	0,610
				3	0,728	0,514	0,688	0,692	0,646
24	6	3	8	4	0,742	0,464	0,682	0,684	0,642

(continua)

Tabela 8 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_Z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
36	18	1	8	2	0,706	0,702	0,700	0,702	0,680
				4	3	0,692	0,692	0,686	0,676
				4	4	0,700	0,698	0,692	0,684
				2	0,742	0,706	0,706	0,706	0,678
				3	0,660	0,654	0,646	0,648	0,628
				4	0,718	0,688	0,682	0,682	0,668
				2	0,726	0,684	0,680	0,680	0,670
			16	3	0,696	0,682	0,682	0,682	0,672
				4	0,676	0,682	0,678	0,678	0,664
				2	0,686	0,612	0,660	0,660	0,642
				4	0,698	0,576	0,680	0,680	0,648
			4	4	0,726	0,544	0,698	0,700	0,680
36	12	2	8	2	0,724	0,642	0,700	0,700	0,694
				3	0,696	0,584	0,680	0,682	0,664
				4	0,718	0,558	0,688	0,690	0,670
				2	0,710	0,632	0,700	0,700	0,682
			16	3	0,720	0,578	0,688	0,688	0,666
				4	0,708	0,548	0,680	0,680	0,662
				2	0,700	0,556	0,656	0,656	0,630
				4	0,704	0,504	0,668	0,668	0,644
				4	0,724	0,460	0,688	0,690	0,656
				2	0,718	0,580	0,682	0,688	0,662
			16	3	0,700	0,512	0,684	0,694	0,660
				4	0,728	0,458	0,702	0,702	0,672
				2	0,722	0,588	0,690	0,692	0,652
				3	0,716	0,472	0,672	0,674	0,648
36	9	3	8	4	0,718	0,482	0,702	0,704	0,674

Tabela 9 - Estimativas do poder do teste para conjuntos de 500 pares de amostras de médias diferentes, segundo a razão do tamanho de amostras, razão de variâncias e procedimento de teste utilizado. ($\Pi \approx 0,80$).

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\Pi}_Z$	$\hat{\Pi}_t$	$\hat{\Pi}_{W-S}$	$\hat{\Pi}_{A-W}$	$\hat{\Pi}_C$
12	6	1	4	2	0,830	0,792	0,786	0,790	0,718
				3	0,814	0,766	0,748	0,754	0,706
				4	0,818	0,748	0,722	0,730	0,686
				2	0,812	0,750	0,738	0,742	0,680
			8	3	0,808	0,730	0,712	0,722	0,690
				4	0,794	0,750	0,726	0,734	0,688
				2	0,830	0,772	0,740	0,746	0,700
				3	0,786	0,724	0,700	0,708	0,640
			16	4	0,814	0,784	0,756	0,770	0,720
				2	0,824	0,704	0,736	0,756	0,660
				3	0,780	0,622	0,728	0,730	0,652
				4	0,810	0,606	0,742	0,754	0,684
12	4	2	8	2	0,826	0,696	0,756	0,778	0,678
				3	0,786	0,622	0,748	0,752	0,662
				4	0,766	0,576	0,696	0,704	0,634
				2	0,826	0,706	0,746	0,756	0,684
			16	3	0,778	0,610	0,732	0,742	0,676
				4	0,782	0,588	0,732	0,738	0,668
				2	0,838	0,674	0,704	0,740	0,598
				3	0,820	0,568	0,684	0,716	0,578
			4	4	0,796	0,536	0,704	0,728	0,584
				2	0,816	0,656	0,688	0,736	0,572
				3	0,794	0,560	0,666	0,696	0,568
				4	0,802	0,522	0,700	0,728	0,576
12	3	3	8	2	0,782	0,618	0,660	0,694	0,538
				3	0,782	0,542	0,670	0,704	0,564
				4	0,830	0,496	0,732	0,758	0,604
			16	3	0,782	0,542	0,670	0,704	0,564
				4	0,830	0,496	0,732	0,758	0,604

(continua)

Tabela 9 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
24	12	1	8	2	0,820	0,778	0,778	0,778	0,748
				3	0,822	0,798	0,792	0,794	0,770
				4	0,798	0,764	0,762	0,762	0,748
				2	0,818	0,790	0,786	0,786	0,770
				3	0,776	0,742	0,738	0,740	0,730
				4	0,826	0,802	0,794	0,794	0,782
			16	2	0,824	0,782	0,780	0,782	0,762
				3	0,824	0,798	0,788	0,792	0,770
				4	0,826	0,786	0,766	0,768	0,752
				2	0,786	0,690	0,740	0,740	0,714
				3	0,818	0,674	0,774	0,774	0,740
				4	0,786	0,642	0,760	0,766	0,736
24	8	2	8	2	0,828	0,752	0,808	0,810	0,790
				3	0,832	0,704	0,806	0,814	0,782
				4	0,816	0,646	0,798	0,802	0,768
				2	0,804	0,722	0,770	0,772	0,756
			16	3	0,816	0,690	0,784	0,790	0,764
				4	0,798	0,658	0,776	0,776	0,760
				2	0,814	0,680	0,728	0,734	0,700
			4	3	0,796	0,604	0,766	0,774	0,734
				4	0,826	0,584	0,798	0,800	0,762
				2	0,826	0,692	0,780	0,780	0,736
			8	3	0,796	0,598	0,756	0,762	0,734
				4	0,794	0,558	0,764	0,770	0,734
				2	0,832	0,692	0,792	0,798	0,756
			16	3	0,808	0,606	0,748	0,756	0,722
				4	0,790	0,536	0,742	0,742	0,714

(continua)

Tabela 9 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
36	18	1	4	2	0,820	0,796	0,794	0,794	0,786
				3	0,820	0,782	0,792	0,792	0,766
				4	0,802	0,800	0,802	0,800	0,782
				2	0,806	0,794	0,794	0,794	0,784
				3	0,802	0,778	0,774	0,776	0,756
				4	0,812	0,790	0,784	0,784	0,776
			16	2	0,816	0,788	0,788	0,788	0,768
				3	0,786	0,782	0,774	0,774	0,772
				4	0,792	0,780	0,770	0,770	0,768
				2	0,822	0,712	0,788	0,788	0,764
				3	0,816	0,700	0,798	0,798	0,776
				4	0,812	0,654	0,752	0,752	0,738
36	12	2	8	2	0,810	0,732	0,786	0,786	0,778
				3	0,828	0,710	0,806	0,806	0,786
				4	0,834	0,678	0,802	0,804	0,784
				2	0,798	0,726	0,780	0,780	0,764
				3	0,802	0,686	0,788	0,790	0,762
				4	0,810	0,660	0,784	0,784	0,760
			16	2	0,838	0,712	0,800	0,804	0,782
				3	0,818	0,632	0,794	0,794	0,774
				4	0,822	0,600	0,790	0,790	0,774
				2	0,806	0,702	0,796	0,800	0,780
				3	0,816	0,624	0,784	0,786	0,760
				4	0,830	0,580	0,786	0,786	0,768
36	9	3	8	2	0,802	0,700	0,786	0,788	0,756
				3	0,816	0,630	0,778	0,778	0,762
				4	0,816	0,562	0,774	0,776	0,758
				2	0,802	0,700	0,786	0,788	0,756

Tabela 10 - Estimativas do poder do teste para conjuntos de 500 pares de amostras de médias diferentes, segundo a razão do tamanho de amostras, razão de variâncias e procedimento de teste utilizado. ($\pi \approx 0,90$).

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_Z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
12	6	1	4	2	0,918	0,870	0,860	0,874	0,828
				3	0,900	0,856	0,840	0,844	0,806
				4	0,898	0,854	0,836	0,838	0,814
			8	2	0,900	0,862	0,856	0,860	0,818
				3	0,872	0,826	0,822	0,822	0,786
				4	0,876	0,836	0,804	0,808	0,762
			16	2	0,914	0,878	0,866	0,872	0,830
				3	0,898	0,844	0,822	0,830	0,798
				4	0,908	0,854	0,840	0,844	0,806
			4	2	0,916	0,828	0,854	0,860	0,790
				3	0,918	0,788	0,860	0,870	0,826
				4	0,880	0,694	0,796	0,808	0,762
12	4	2	8	2	0,870	0,768	0,798	0,810	0,752
				3	0,924	0,800	0,866	0,872	0,812
				4	0,882	0,716	0,808	0,822	0,776
			16	2	0,908	0,830	0,854	0,864	0,804
				3	0,882	0,714	0,800	0,810	0,744
				4	0,882	0,748	0,844	0,852	0,794
			4	2	0,902	0,780	0,776	0,806	0,678
				3	0,906	0,682	0,798	0,810	0,688
				4	0,916	0,648	0,828	0,844	0,740
			8	2	0,906	0,788	0,794	0,828	0,648
				3	0,906	0,718	0,808	0,838	0,726
				4	0,920	0,702	0,824	0,848	0,732
12	3	3	16	2	0,908	0,766	0,774	0,818	0,696
				3	0,912	0,714	0,818	0,844	0,706
				4	0,898	0,652	0,832	0,852	0,744

(continua)

Tabela 10 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_Z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
24	12	1	8	2	0,918	0,892	0,892	0,892	0,882
				3	0,920	0,882	0,874	0,876	0,866
				4	0,908	0,856	0,850	0,854	0,840
				2	0,910	0,898	0,898	0,898	0,886
				3	0,900	0,894	0,894	0,894	0,876
				4	0,912	0,888	0,884	0,884	0,874
			16	2	0,898	0,880	0,876	0,880	0,866
				3	0,924	0,888	0,888	0,888	0,872
				4	0,902	0,876	0,866	0,868	0,860
				2	0,924	0,842	0,892	0,894	0,874
				3	0,914	0,784	0,872	0,876	0,850
				4	0,900	0,758	0,864	0,864	0,846
24	8	2	8	2	0,916	0,856	0,894	0,894	0,886
				3	0,924	0,816	0,906	0,910	0,894
				4	0,892	0,774	0,878	0,880	0,858
				2	0,910	0,830	0,868	0,874	0,850
			16	3	0,900	0,786	0,874	0,876	0,856
				4	0,900	0,798	0,880	0,880	0,868
				2	0,886	0,792	0,854	0,862	0,828
			4	3	0,914	0,768	0,880	0,882	0,862
				4	0,900	0,714	0,866	0,872	0,850
				2	0,926	0,808	0,864	0,870	0,820
				3	0,912	0,750	0,890	0,892	0,854
26	6	3	8	4	0,888	0,710	0,858	0,864	0,828
				2	0,902	0,804	0,850	0,858	0,816
			16	3	0,928	0,768	0,878	0,884	0,842
				4	0,910	0,704	0,896	0,896	0,878

(continua)

Tabela 10 - continuação

N	n_1	n_2/n_1	σ_1^2	σ_2^2/σ_1^2	$\hat{\pi}_z$	$\hat{\pi}_t$	$\hat{\pi}_{W-S}$	$\hat{\pi}_{A-W}$	$\hat{\pi}_C$
36	18	1	4	2	0,898	0,892	0,890	0,890	0,886
				3	0,922	0,910	0,910	0,910	0,904
				4	0,908	0,890	0,884	0,884	0,878
				2	0,912	0,882	0,882	0,882	0,870
			8	3	0,918	0,904	0,900	0,900	0,898
				4	0,912	0,900	0,892	0,892	0,884
				2	0,904	0,888	0,886	0,886	0,884
				16	3	0,918	0,906	0,906	0,896
				4	0,918	0,902	0,902	0,902	0,896
			4	2	0,914	0,850	0,898	0,900	0,894
				3	0,924	0,806	0,900	0,900	0,890
				4	0,910	0,830	0,904	0,904	0,894
				2	0,896	0,830	0,870	0,870	0,852
			8	3	0,906	0,808	0,880	0,880	0,864
				4	0,920	0,814	0,896	0,896	0,890
				2	0,916	0,820	0,876	0,876	0,858
36	12	2	16	3	0,922	0,820	0,890	0,890	0,878
				4	0,906	0,808	0,854	0,894	0,874
				2	0,930	0,846	0,892	0,894	0,886
				4	0,914	0,774	0,892	0,898	0,882
			4	4	0,910	0,714	0,890	0,890	0,874
				2	0,916	0,826	0,886	0,888	0,868
				3	0,912	0,780	0,898	0,898	0,884
				4	0,916	0,724	0,892	0,894	0,870
			8	2	0,900	0,810	0,872	0,876	0,858
				3	0,904	0,782	0,888	0,892	0,878
				4	0,910	0,726	0,888	0,890	0,870
			16	3	0,904	0,782	0,888	0,892	0,878
				4	0,910	0,726	0,888	0,890	0,870