

## Comparação de testes paramétricos e não paramétricos aplicados em delineamentos experimentais

Gustavo Mello Reis (UFV) [gustavo\\_epr@yahoo.com.br](mailto:gustavo_epr@yahoo.com.br)

José Ivo Ribeiro Júnior (UFV) [jivo@dpi.ufv.br](mailto:jivo@dpi.ufv.br)

*RESUMO: Para comparar mais de dois níveis de um fator de interesse sob os delineamentos inteiramente (DIC) e em blocos casualizados (DBC), podem ser usados os testes não paramétricos de Kruskal-Wallis e Friedman, respectivamente, além do teste F da análise de variância (ANOVA). Com base na simulação de mil amostras com 5, 10 e 25 repetições por amostra, foram estimados os alarmes falso ( $\alpha$ ) e verdadeiro (poder) sob as distribuições normal, lognormal e binomial, para  $H_0$  verdadeira e falsa, respectivamente. As estimativas desses alarmes se referem às proporções dos p-valores menores ou iguais a cada valor de  $\alpha$  teórico, no intervalo de  $0 \leq \alpha \leq 1$  nas duas hipóteses. Desse modo, para um DIC, foram plotados num gráfico de dispersão, as estimativas de  $\alpha$  pelos testes F da ANOVA e de Kruskal-Wallis em função do  $\alpha$  teórico, sendo o mesmo feito para as estimativas do poder (Pd) destes testes. Para um DBC, de acordo com os testes F da ANOVA e de Friedman, foram adotados os mesmos procedimentos. De acordo com os gráficos construídos, concluiu-se que o teste F, tanto para o DIC quanto para o DBC, proporcionou estimativas do Pd maiores que a dos testes não paramétricos.*

*Palavras-chave: Testes de hipóteses; Comparação; Delineamentos experimentais.*

### 1. Introdução

A formulação de hipóteses tem sido muito empregada em pesquisas de diversas áreas do conhecimento. Para decidir se uma determinada hipótese é confirmada por um conjunto de dados, é necessário ter um procedimento objetivo para aceitar ou rejeitar a hipótese (SIEGEL & CASTELLAN, 2006).

Segundo Shimakura, métodos estatísticos são utilizados para o planejamento e condução de um estudo, descrição dos dados e para tomada de decisões, onde pode-se citar os testes de hipóteses que se baseiam nos riscos associados às mesmas.

Quando formulamos uma decisão sobre  $H_0$  podem ocorrer dois erros distintos. O primeiro, designado por erro tipo I, consiste em rejeitar a hipótese nula quando ela é verdadeira. O segundo, designado por erro tipo II, consiste em aceitar  $H_0$  quando ela é falsa. A estes erros estão associados uma probabilidade:  $P(\text{rejeitar } H_0 \mid H_0 \text{ verdadeira}) = \alpha$ ;  $P(\text{aceitar } H_0 \mid H_0 \text{ falsa}) = \beta$  (CÂMARA & SILVA, 2001).

Os testes de hipóteses se dividem em paramétricos e não paramétricos. Os paramétricos são aqueles que utilizam os parâmetros da distribuição, ou uma estimativa destes, para o cálculo de sua estatística. Normalmente, estes testes são mais rigorosos e possuem mais pressuposições para sua validação. Já os não paramétricos utilizam, para o cálculo de sua estatística, postos atribuídos aos dados ordenados e são livres da distribuição de probabilidades dos dados estudados.

Segundo Moore & McCabe (2002), os testes de hipóteses estão entre os tipos mais comuns de inferência. Percebe-se que os testes de hipóteses paramétricos são os mais utilizados, muitas vezes devido ao não conhecimento dos seus concorrentes não paramétricos.

No entanto, segundo Siegel & Castellan (2006), a validação dos resultados dos testes paramétricos depende da verificação de suas pressuposições, como por exemplo a normalidade dos dados, pressuposição básica para aplicação da maioria dos testes paramétricos.

Desse modo, é interessante verificar até que ponto os resultados dos testes paramétricos serão prejudicados quando a pressuposição de normalidade não venha a ser satisfeita.

## 2. Testes de hipóteses aplicáveis em delineamentos experimentais

O teste paramétrico, sob normalidade, mais usual em experimentos com um fator, com mais de dois níveis de interesse, é o F da análise de variância (ANOVA), que pode ser realizado segundo os delineamentos inteiramente casualizado (DIC) e em blocos casualizados (DBC).

O teste não paramétrico de Kruskal-Wallis foi introduzido por estes autores, Kruskal e Wallis, em 1952, como um competidor ou um substituto do teste F da ANOVA segundo um IDC (CAMPOS, 1976). Assim como a maioria dos testes não paramétricos, este também dispõe as respostas dos tratamentos que serão comparados na forma de postos. Quanto maior for a diferença entre a soma dos postos, maior será a evidência de que exista diferença entre os mesmos. Mesmo não precisando da exigência de normalidade ou de outra distribuição qualquer para as populações estudadas, o teste exige que a distribuição dos erros a mesma para todos os níveis. Como alternativa não paramétrica ao teste F da ANOVA, segundo o DBC, existe o teste de Friedman.

## 3. Objetivo

O presente trabalho teve como objetivo principal comparar a eficiência entre os métodos paramétricos e não paramétricos aplicáveis em comparações múltiplas, para dados sob a presença de normalidade ou não. A comparação foi feita através das estimativas do erro  $\alpha$  e do poder do teste (Pd), em que  $\alpha$  é a probabilidade de rejeitar a hipótese  $H_0$  quando ela é verdadeira (alarme falso) e Pd é a probabilidade de rejeitar a hipótese  $H_0$  quando ela é falsa (alarme verdadeiro).

## 4. Metodologia empregada

### 4.1. Simulação dos dados

Para estimar o  $\alpha$  e o Pd, foi analisado o quadrado do parâmetro phi ( $\Phi^2$ ) que, segundo Montgomery & Runger (2003), está relacionado ao parâmetro delta ( $\delta$ ) de não centralidade. Foi adotado  $\Phi^2 = 0$  ( $H_0$  verdadeira) para estudar o  $\alpha$  estimado e  $\Phi^2 = 4$  ( $H_0$  falsa) para estudar o Pd estimado, de acordo com a seguinte fórmula:

$$\Phi^2 = \frac{n \sum_{i=1}^a \tau_i^2}{a\sigma^2}, \text{ em que:}$$

$n$  = nº de repetições (DIC) = nº de blocos (DBC);

$\tau_i$  = efeito do tratamento  $i$ ;

$a$  = nº de médias a serem testadas;

$\sigma^2$  = variância do erro experimental.

Os parâmetros  $n$  (nº de repetições) e  $a$  (nº de médias a serem testadas) foram especificados em  $a = 5$  (níveis) e  $n = 5, 10$  ou  $25$ . Já os parâmetros  $\tau_i$  e  $\sigma^2$  foram estabelecidos de forma a obter  $\Phi^2 = 0$  e  $\Phi^2 = 4$ , para as hipóteses  $H_0$  falsa e verdadeira, respectivamente.

No primeiro caso, para obter  $\Phi^2 = 0$ , basta adotar  $\tau_i = 0, \forall i, i = \{1, 2, 3, 4, 5\}$ . Já no segundo caso, para obter  $\Phi^2 = 4$ , deve-se combinar os valores dos  $\tau_i$ 's com os valores de  $\sigma^2$ .

Logo, pode-se obter o valor do  $\sum_{i=1}^a \tau_i^2$  através da equação de  $\Phi^2$ , da seguinte forma:

$$\Phi^2 = \frac{n \sum_{i=1}^a \tau_i^2}{a\sigma^2} \quad \Rightarrow \quad \sum_{i=1}^a \tau_i^2 = \frac{a\sigma^2 \Phi^2}{n}$$

Neste estudo foram simulados dados por meio da imposição de que  $\tau_i \neq 0$  para somente um tratamento. Logo, para  $a = 5$  e  $\Phi^2 = 4$ , tem-se  $\tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ , sendo a equação utilizada para estimar  $\tau_5$  igual a:

$$(0^2 + 0^2 + 0^2 + 0^2 + \tau_5^2) = \frac{5\sigma^2 4}{n} \quad \Rightarrow \quad \tau_5 = \sqrt{\frac{20\sigma^2}{n}} \quad \Rightarrow \quad \tau_5 = \sigma \sqrt{\frac{20}{n}}$$

No caso do DBC, primeiro foi realizada a simulação da distribuição com os parâmetros de interesse, assim como no DIC, para depois somar a estes, os efeitos dos blocos e, em seguida, o erro aleatório relacionado a cada valor. O efeito do bloco foi representado pela adição de 40 unidades em relação ao bloco anterior.

Os efeitos do erro aleatório foram simulados de acordo com as seguintes distribuições:

Os parâmetros que regem as três distribuições e também o tamanho da amostra ( $n$ ), foram analisados em diferentes situações com o objetivo de verificar qual o efeito que estes parâmetros têm sobre o  $\alpha$  e o Pd estimados.

Para os dados simulados a partir da distribuição lognormal, foram considerados os parâmetros  $\mu_1 = \ln(\mu)$  e  $\sigma_1 = \sigma/\mu$ , de forma aproximada.

A confiabilidade deste estudo se baseia na quantidade de amostras simuladas. Para a construção de cada gráfico ilustrado foram simuladas quinhentas mil amostras e aplicados duzentos mil testes. Desconsiderando as simulações em fase de teste e outras que tiveram que ser refeitas para alguma alteração, para a realização deste estudo foram simuladas dez milhões e quinhentas mil amostras e aplicados quatro milhões e duzentos mil testes.

## 4.2. Realização dos Testes

Como  $\alpha$  é a probabilidade de rejeitar  $H_0$  quando esta é verdadeira ( $\Phi^2 = 0$ ) e o Pd é a probabilidade de rejeitar  $H_0$  quando esta é falsa ( $\Phi^2 = 4$ ), a estimação destes foi realizada por meio de 1000 amostras simuladas de cada distribuição estudada. Em seguida foram aplicados os testes estatísticos paramétricos, F da ANOVA DIC (ADIC) e da ANOVA DBC (ADBC), e não paramétricos, de Kruskal-Wallis (K-W) e de Friedman (Fr). Os parâmetros utilizados na simulação dos dados foram aqueles utilizados na formulação das hipóteses  $H_0$ 's. Para aplicação do teste binomial, foram realizadas algumas manipulações nos dados simulados. Primeiro foi realizada uma subtração dos valores da amostra simulada pela média populacional ( $\mu$ ). Em seguida foi atribuído o valor 1 para os valores positivos e 0 para os valores negativos, sendo os valores nulos descartados. Por fim, o teste binomial foi aplicado à nova amostra de zeros e uns, sendo  $p = 0,5$  a hipótese nula ( $H_0$ ).

### 4.3. Programa R

As simulações dos dados, a aplicação dos testes e as estimativas do  $\alpha$  e do Pd, foram realizados utilizando-se o programa gratuito R. No programa, foram programadas funções para gerar os dados de acordo com os parâmetros de interesse e para aplicar os testes sobre os dados criados, armazenando o p-valor de cada teste, para as 1000 amostras de cada distribuição.

Com os 1000 p-valores retornados de cada teste e de acordo com o  $\alpha$  de referência variando de 0 a 1 em intervalos de 0,001, foram contabilizados quantas amostras foram rejeitadas no respectivos valores de  $\alpha$  que, pela definição, é o  $\alpha$  estimado ( $\hat{\alpha}$ ), para  $\Phi^2 = 0$ , ou seja, a probabilidade de rejeitar a hipótese nula quando esta é verdadeira (alarme falso) e, o Pd estimado ( $\hat{Pd}$ ), para  $\Phi^2 = 4$ , ou seja, a probabilidade rejeitar a hipótese nula quando esta é falsa (alarme verdadeiro).

Para obter valores mais precisos do  $\alpha$  e do Pd estimados, a função foi programada, também, para repetir o processo de simulação dos dados, aplicação dos testes e estimação do  $\alpha$  e do Pd, por 100 vezes e, em seguida, calcular a média dos 100 valores encontrados para cada  $\alpha$  de referência. Assim, com mais precisão, a média encontrada, em cada caso, será o  $\alpha$  e o Pd estimados, que foram salvos em arquivos do tipo csv.

### 4.4. Gráfico de dispersão

Para melhor avaliar o  $\alpha$  e o Pd estimados, foram plotados gráficos de dispersão da probabilidade acumulada, onde para  $\Phi^2 = 0$  permitiu estudar o  $\alpha$  estimado (eixo y) em função do  $\alpha$  de referência (eixo x) e, para  $\Phi^2 = 4$ , foi estudado o Pd estimado (eixo y) em função do  $\alpha$  de referência (eixo x), para todos os testes aplicados.

Sendo a distância no eixo x ( $\alpha$  de referência), entre um par ordenado e outro, de apenas 0,001, estes pares foram ligados por linhas para reproduzir a continuidade destes valores, ou seja, ao invés de mostrar inúmeros pontos consecutivos que distam entre si de apenas 0,001, foi exibida uma linha contínua que percorre todo o digrama de dispersão.

Os gráficos de dispersão da probabilidade acumulada, foram construídos no programa comercial Microsoft Excel, contendo uma linha representativa de cada teste em estudo, para cada uma das três distribuições.

### 4.5. Análise dos gráficos

Para a análise do erro  $\alpha$ , primeiro foi comparada uma situação com o aumento do desvio padrão ( $\sigma$ ), para as distribuições de probabilidades normal e lognormal. Para a binomial, fez-se referência ao coeficiente de variação (CV), ao invés do  $\sigma$ , porque a média nem sempre era mantida. Já para o CV ( $\sigma/\mu$ ), esta variação na média será considerada. Neste caso, tem-se  $\sigma = \sqrt{nb \times p \times (1-p)}$  e  $\mu = nb \times p$ . Note que para indicar o número de itens avaliados na distribuição binomial, foi utilizado o símbolo nb, com objetivo de diferenciá-lo do símbolo n que, neste estudo, representa o tamanho da amostra simulada. Em seguida, foi avaliado um aumento do tamanho da amostra simulada (n).

Para análise do  $\hat{Pd}$ , teve uma comparação a mais em relação ao  $\alpha$ , que é a comparação do desvio para cima da média ( $\tau$  positivo) com o desvio para baixo da média ( $\tau$  negativo), sendo as outras iguais às utilizadas para o  $\hat{\alpha}$ .

As discussões para  $\hat{\alpha}$ , será feita com base em  $\alpha$ 's de referência no intervalo de 0 a 0,20, o qual é exibido no gráfico. Já para o Pd, os  $\alpha$ 's de referência exibidos estarão no

intervalo de 0 a 0,30. Estes intervalos foram escolhidos de forma a obter uma melhor comparação visual dos testes, e, também, por abranger a maior parte dos estudos.

Para facilitar a indicação dos parâmetros simulados em cada gráfico, considere como apresentado na Tabela 1, para i variando de 1 a 4.

TABELA 1. Parâmetros das distribuições

Distribuição Normal	Distribuição LogNormal	Distribuição Binomial
$n_i = n_5 = n$	$n_i = n_5 = n$	$n_i = n_5 = n$
$\mu_i = \mu; \mu_5 = \mu + \tau$	$\mu_{i1} = \mu_1 = \ln(\mu); \mu_{i5} = \ln(\mu + \tau)$	$nb_i = nb_5 = nb$
$\sigma_i = \sigma$	$\sigma_{i1} = \sigma_1 \approx \sigma/\mu; \sigma_{i5} \approx \sigma/(\mu + \tau)$	$p_i = p; p_5 = p + \tau_p$

Onde:  $\tau = \sigma \sqrt{\frac{20}{n}}$  e,  $\tau_p = \sigma_p \sqrt{\frac{20}{n}}$ , sendo  $\sigma_p = \sqrt{\frac{p \times (1-p)}{nb}}$ .

Em que  $\sigma_p$  é o desvio padrão do parâmetro p da distribuição binomial.

Note que quando  $\tau = 0$ , para o estudo de alfa, tem-se  $\mu_i = \mu_5$  (normal),  $\mu_{i1} = \mu_{i5}$  e  $\sigma_{i1} = \sigma_{i5}$  (lognormal) e quando  $\tau_p = 0$ ,  $p_i = p_5$  (binomial).

## 5. Resultados e Discussões

### 5.1. Alfa

Observando os gráficos a seguir, pode-se perceber certa semelhança entre os resultados obtidos para o DIC e DBC, sendo as especificações dos testes de Kruskal-Wallis (K-W), e Friedman (Fr), semelhantes para as distribuições normal, lognormal e binomial.

Ao aumentar o  $\sigma$  ou o CV, observa-se nas Figuras 1 e 2 (normal), 4 e 5 (lognormal) e 7 e 8 (binomial), que não houve grandes alterações para o  $\alpha$  estimado em todos os testes.

Percebe-se também, que ao aumentar o n de 5 para 25, observa-se nas Figuras 1 e 3 (normal), 4 e 6 (lognormal) e 7 e 9 (binomial), que há um aumento nas estimativas de  $\alpha$  para os testes não paramétricos, K-W e Fr, tanto que estes passam a ter  $\alpha$ 's estimados semelhantes aos observados para a ANOVA (ADIC e ADBC), sendo esta pouco influenciada pelo aumento de n. De todo modo, todas as estimativas de  $\alpha$  estão próximas aos valores teóricos.

#### 5.1.1. Distribuição Normal

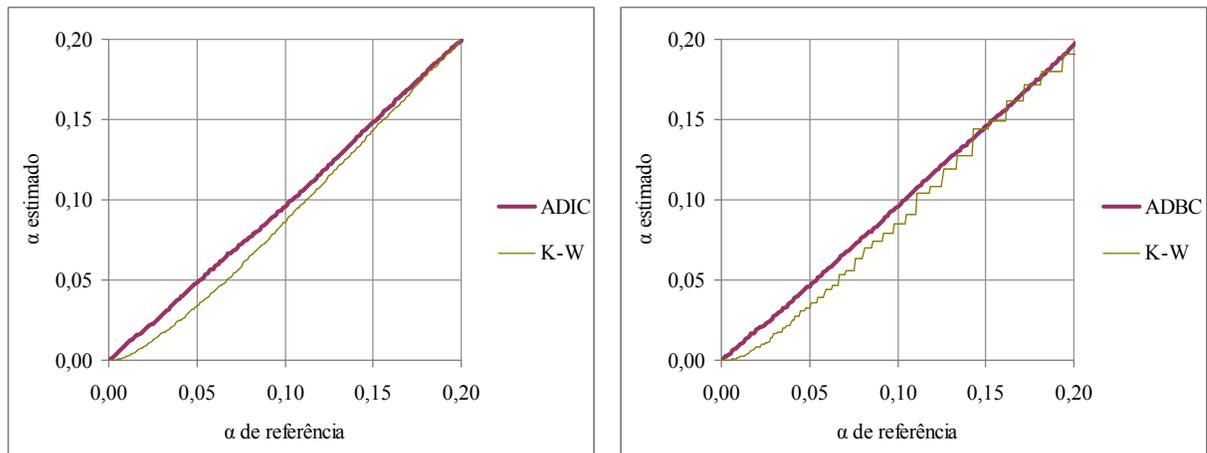


FIGURA 1 - Amostras simuladas com n = 5,  $\mu = 1000$ ,  $\sigma = 30$ ,  $\tau = 0$

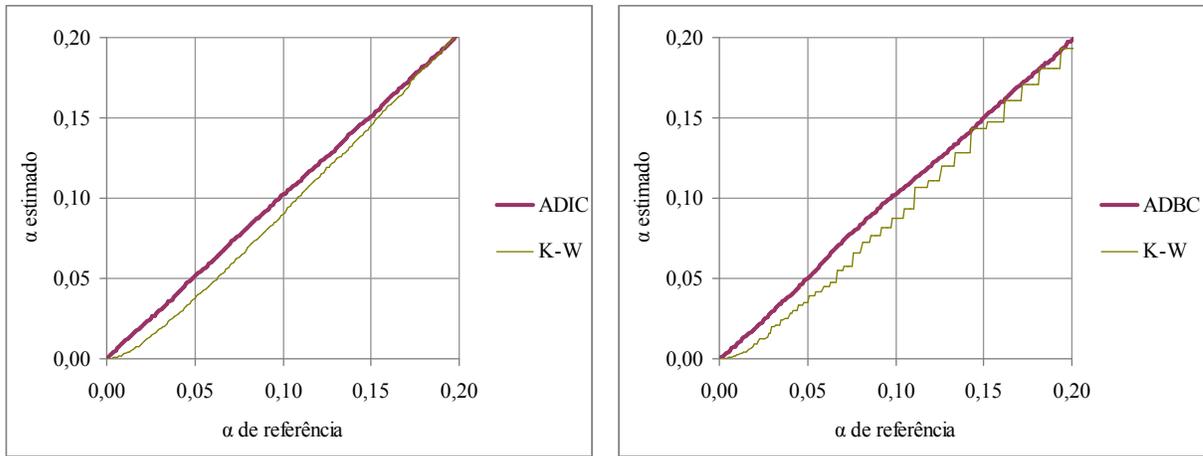


FIGURA 2 - Amostras simuladas com  $n = 5$ ,  $\mu = 1000$ ,  $\sigma = 150$ ,  $\tau = 0$

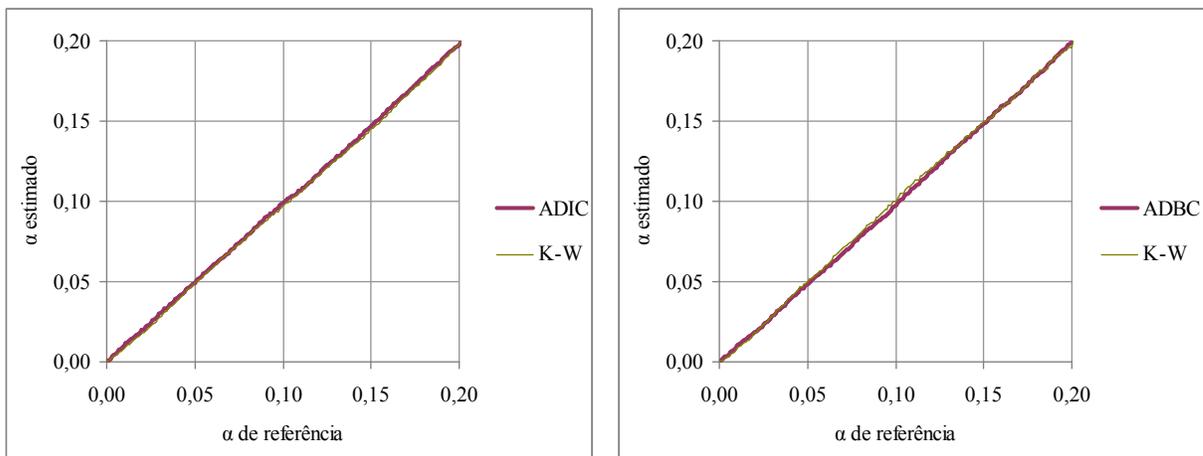


FIGURA 3 - Amostras simuladas com  $n = 25$ ,  $\mu = 1000$ ,  $\sigma = 30$ ,  $\tau = 0$

**5.1.2. Distribuição LogNormal**

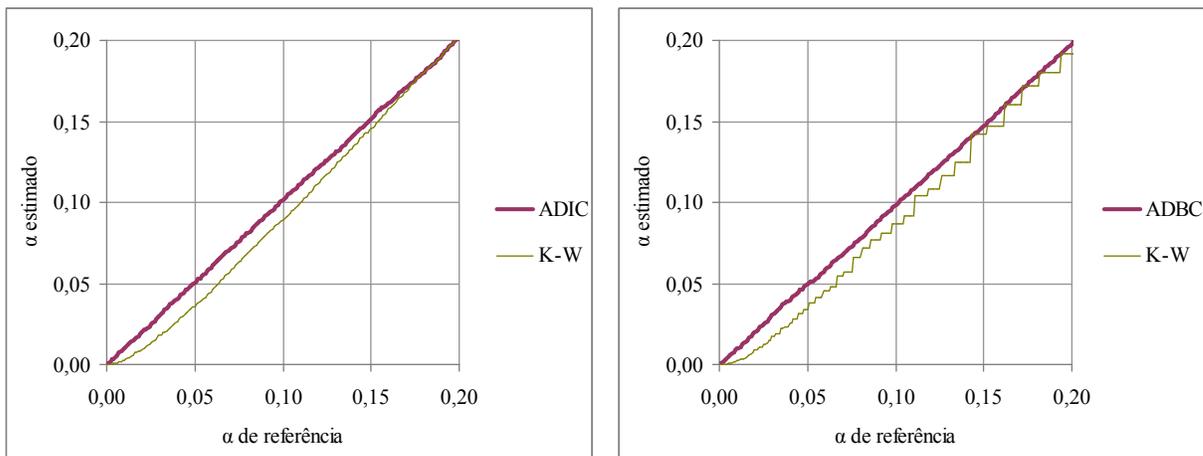


FIGURA 4 - Amostras simuladas com  $n = 5$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 30/1000$ ,  $\tau = 0$

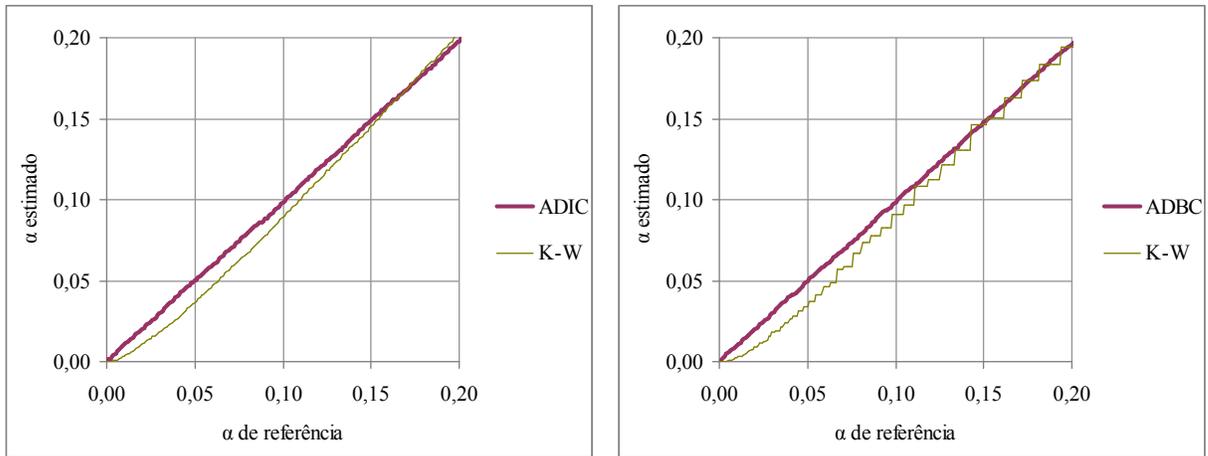


FIGURA 5 - Amostras simuladas com  $n = 5$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 150/1000$ ,  $\tau = 0$

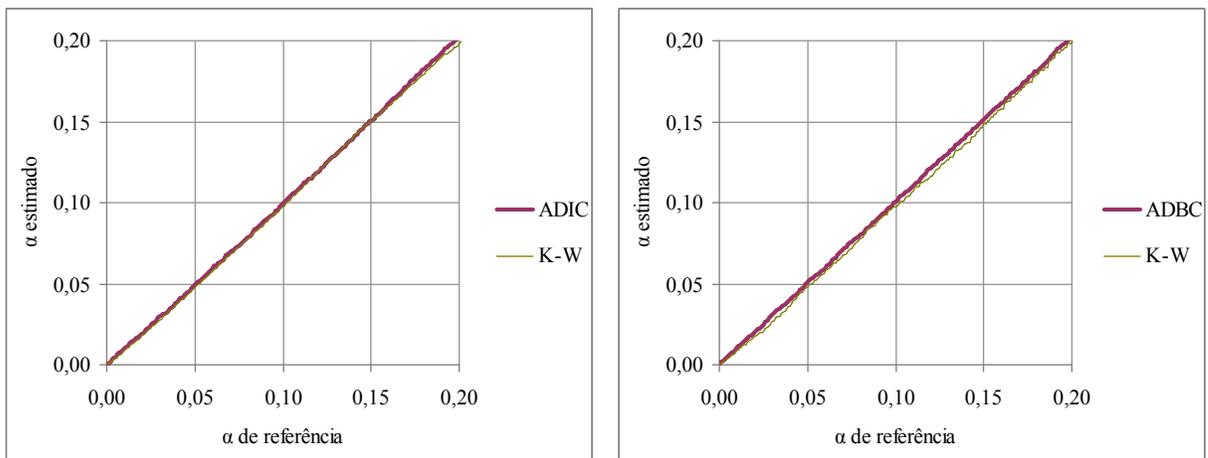


FIGURA 6 - Amostras simuladas com  $n = 25$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 30/1000$ ,  $\tau = 0$

### 5.1.3. Distribuição Binomial

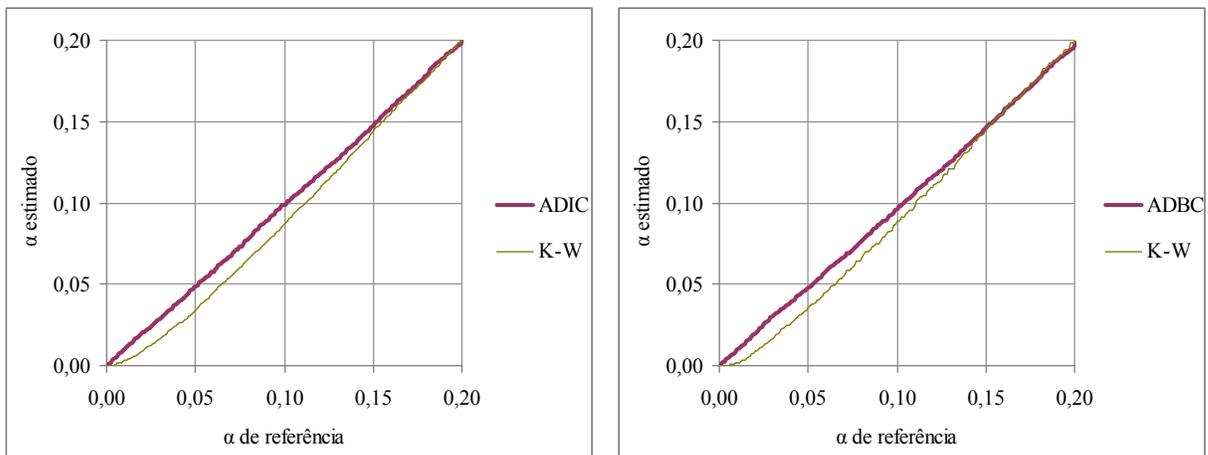
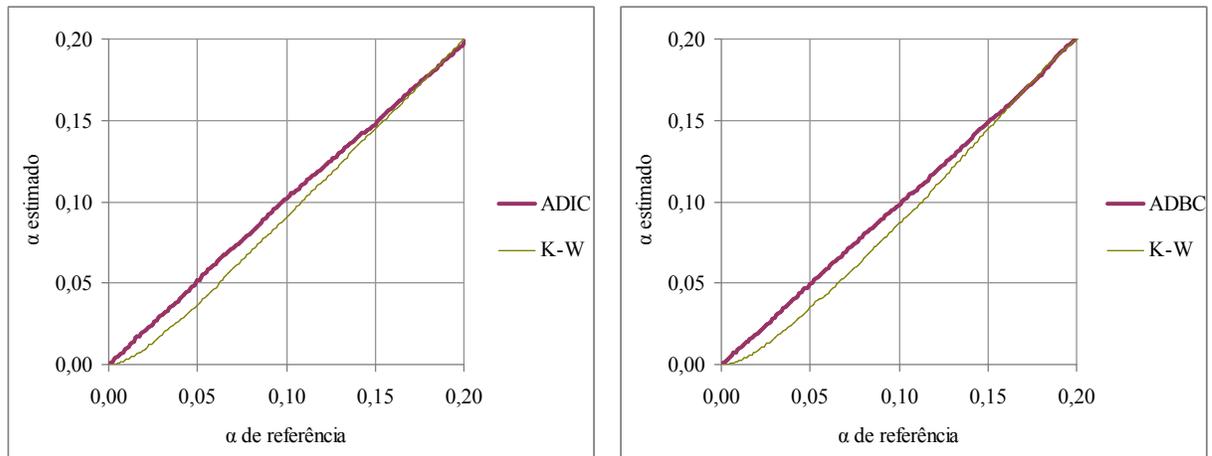
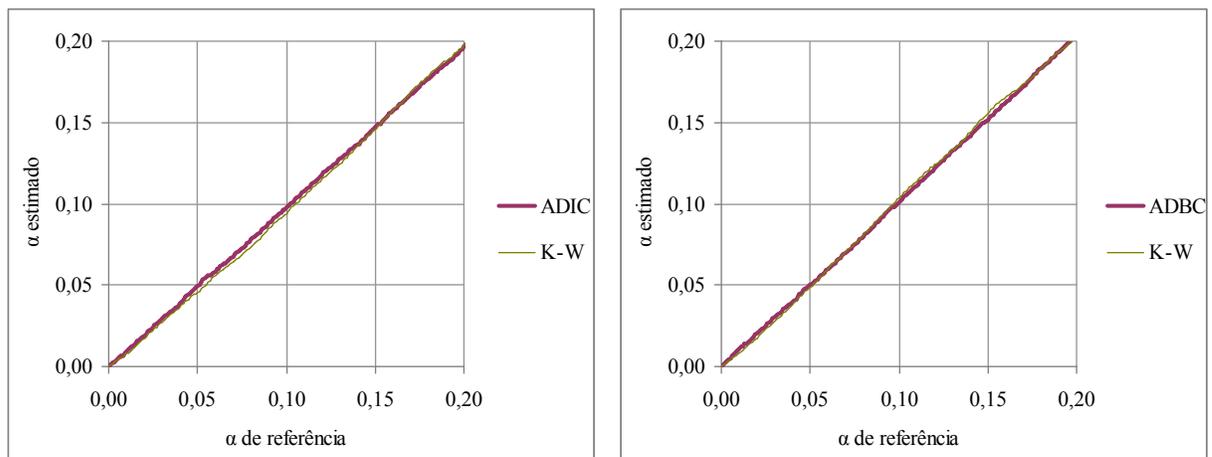


FIGURA 7 - Amostras simuladas com  $n = 5$ ,  $n_b = 1000$ ,  $p = 0,1$ ,  $\tau_p = 0$

FIGURA 8 - Amostras simuladas com  $n = 5$ ,  $nb = 100$ ,  $p = 0,1$ ,  $\tau_p = 0$ FIGURA 9 - Amostras simuladas com  $n = 25$ ,  $nb = 1000$ ,  $p = 0,1$ ,  $\tau_p = 0$ 

## 5.2. Poder

Para a estimação do PD, a semelhança observada entre o DIC e DBC está relacionada ao comportamento dos testes, quando se altera as variáveis em estudo. De forma geral, as estimativas do Pd dos testes utilizados no DBC foram menores que aquelas obtidas no DIC, em todas as situações estudadas. Entretanto, essa diferença foi maior entre os testes não paramétrico.

Entre as distribuições, a semelhança observada é com relação ao aumento de  $n$ , de 5 para 10, em que as estimativas do Pd dos testes aumentaram significativamente, sendo os aumentos ocorridos nos testes não paramétricos, maiores que nos paramétricos, mas não suficiente para alcançar o Pd dos últimos que foram sempre superiores em todos os casos estudados.

As mudanças no desvio da média, para baixo e para cima, provocado em um dos tratamentos ( $\tau$ ), não alterou o Pd dos testes sob distribuição normal (Figuras 10 e 11), como já era de se esperar, já que é uma distribuição simétrica. Para a distribuição lognormal (Figuras 14 e 15), observa-se uma pequena queda no Pd dos testes não paramétricos, K-W e

Fr, ao provocar um desvio para baixo da média. Já para a distribuição binomial (Figuras 18 e 19) percebe-se um maior Pd para ambos os testes ao provocar um desvio para baixo da média.

Para um aumento no  $\sigma$  na distribuição normal (Figuras 10 e 12), não houve mudança significativa na estimativa do Pd dos testes. Já para a distribuição lognormal (Figuras 14 e 16), percebeu-se uma ligeira diminuição no Pd do teste F da ADIC e da ADBC e um ligeiro aumento no Pd dos testes K-W e Fr. Esta diferença pode ser explicada pelo maior distanciamento da distribuição normal ao aumentar o  $\sigma_1$  da distribuição lognormal. No entanto, o teste F continuou tendo maior Pd para  $\alpha < 0,20$ , que abrange a maior parte dos estudos. Já para a distribuição binomial (Figuras 18 e 20), houve uma diminuição no Pd dos testes ao aumentar o CV. Novamente, esta alteração se deve ao maior distanciamento da distribuição normal ao diminuir o nb (aumento de CV) da distribuição binomial. No entanto, a distância entre o Pd do teste F (ADIC e ADBC) e os testes não paramétricos, K-W e Fr, aumentou. Portanto, este distanciamento da normalidade beneficiou o teste paramétrico, diferentemente do esperado.

### 5.2.1. Distribuição Normal

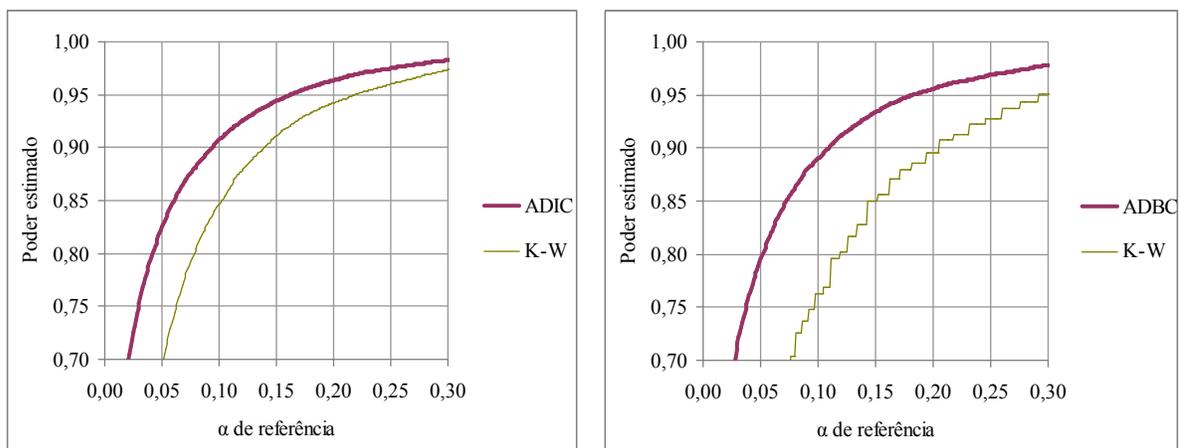


FIGURA 10 - Amostras simuladas com  $n = 5$ ,  $\mu = 1000$ ,  $\sigma = 30$ ,  $\tau = 60$

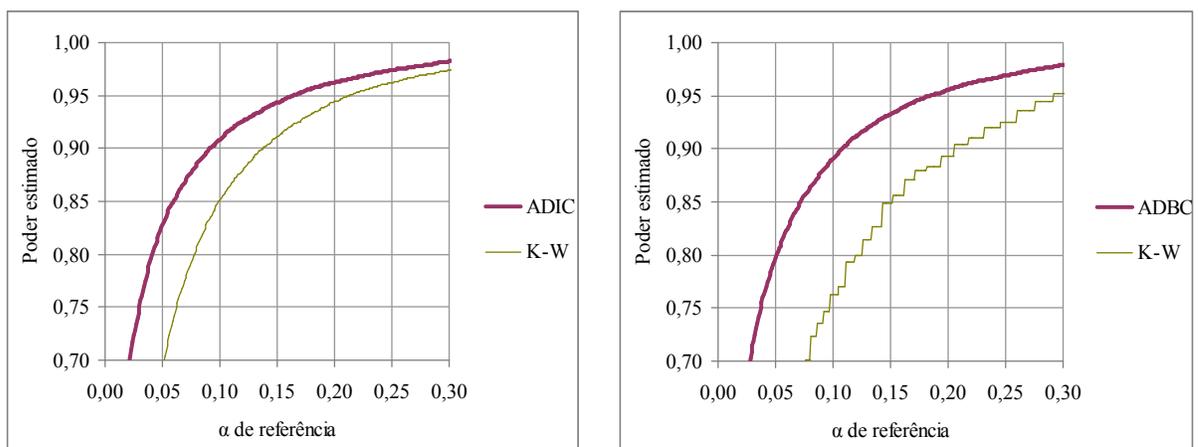


FIGURA 11 - Amostras simuladas com  $n = 5$ ,  $\mu = 1000$ ,  $\sigma = 30$ ,  $\tau = -60$

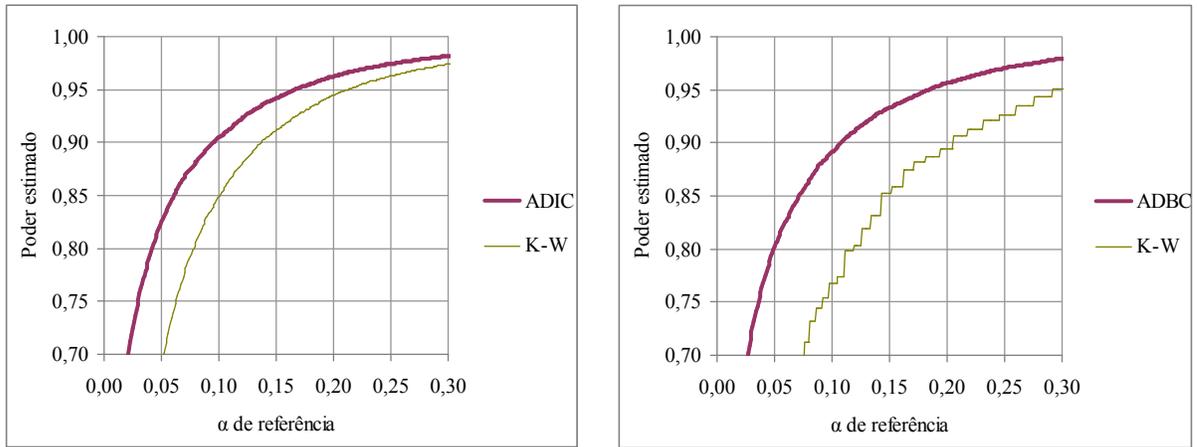


FIGURA 12 - Amostras simuladas com  $n = 5$ ,  $\mu = 1000$ ,  $\sigma = 150$ ,  $\tau = 300$

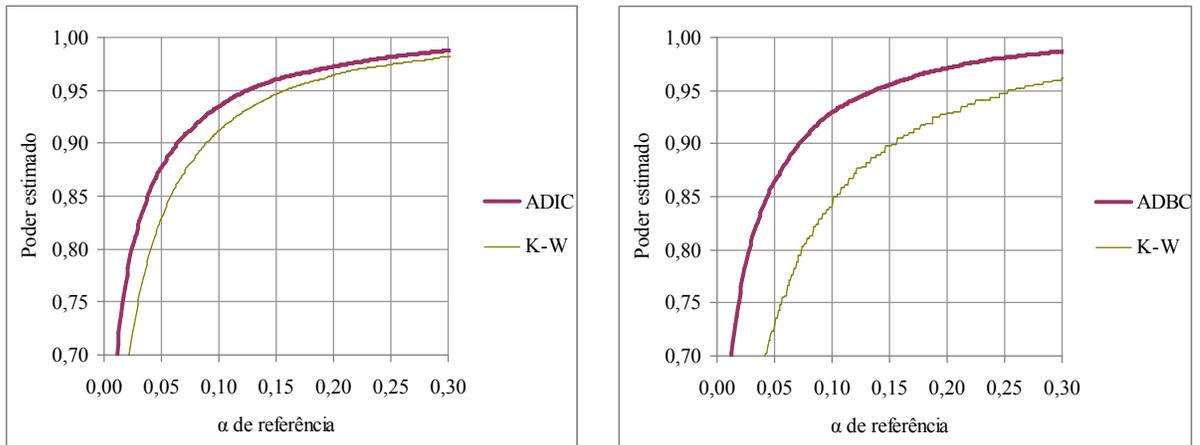


FIGURA 13 - Amostras simuladas com  $n = 10$ ,  $\mu = 1000$ ,  $\sigma = 30$ ,  $\tau = 30\sqrt{2}$

### 5.2.2. Distribuição LogNormal

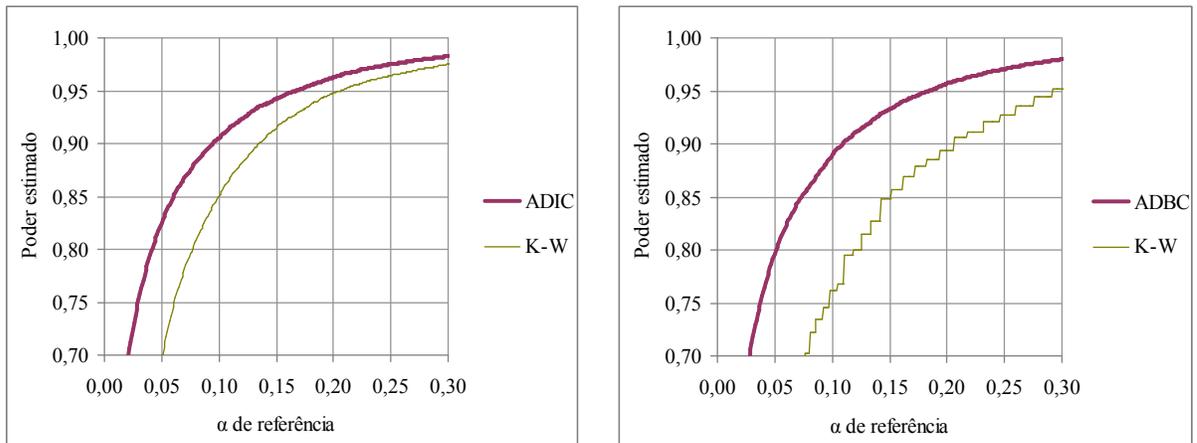


FIGURA 14 - Amostras simuladas com  $n = 5$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 30/1000$ ,  $\tau = 60$

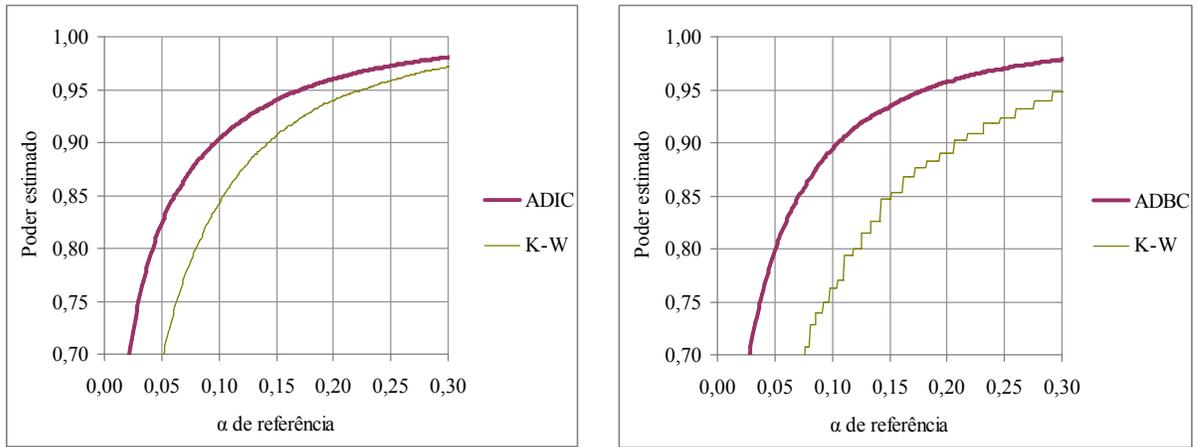


FIGURA 15 - Amostras simuladas com  $n = 5$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 30/1000$ ,  $\tau = -60$

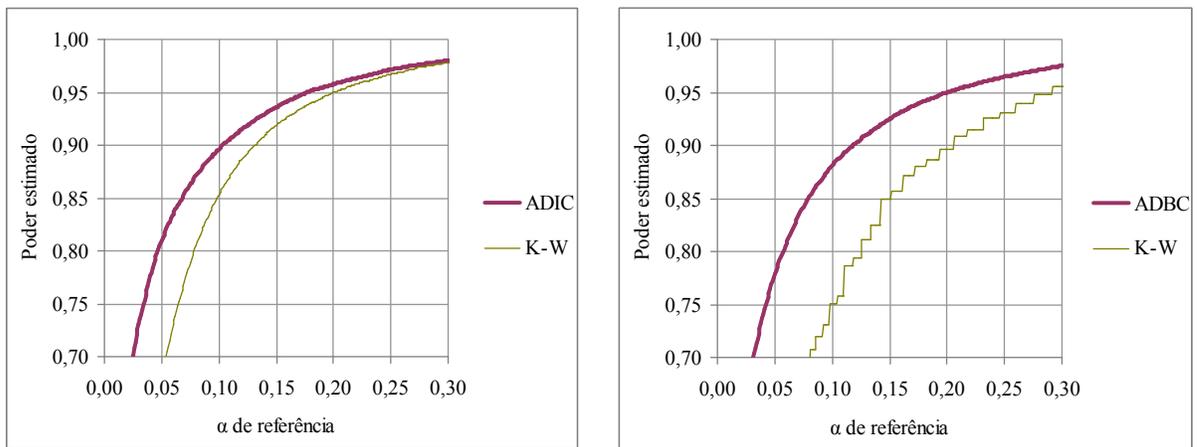


FIGURA 16 - Amostras simuladas com  $n = 5$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 150/1000$ ,  $\tau = 300$

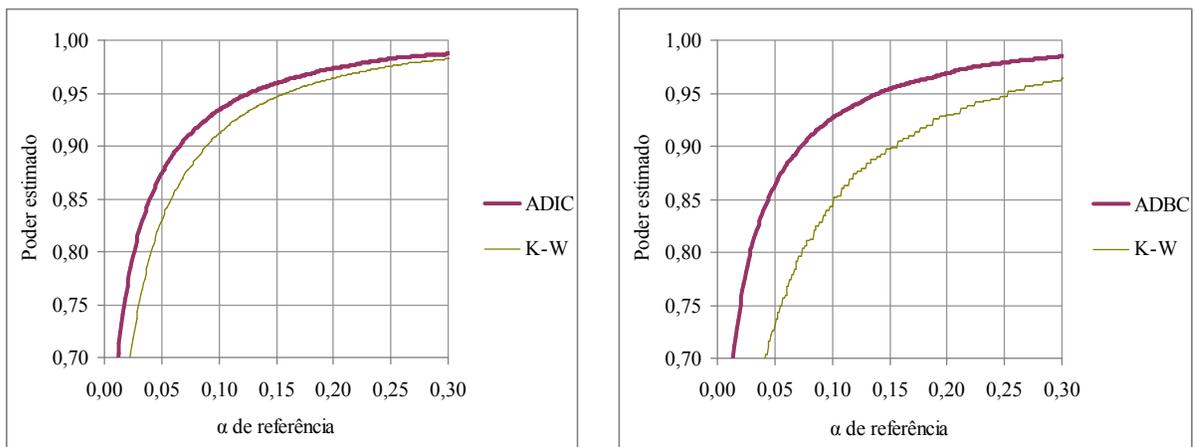


FIGURA 17 - Amostras simuladas com  $n = 10$ ,  $\mu_1 = \ln(1000)$ ,  $\sigma_1 = 30/1000$ ,  $\tau = 30\sqrt{2}$

### 5.2.3. Distribuição Binomial

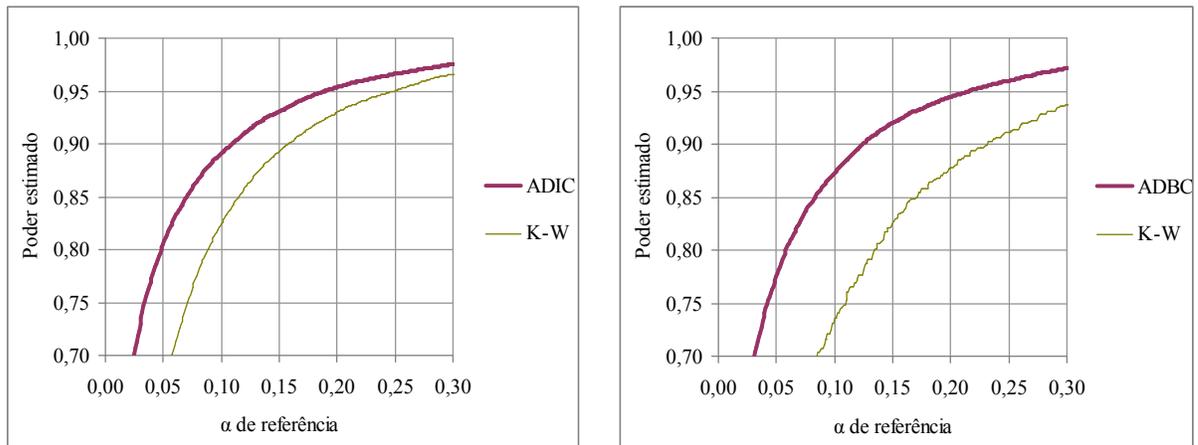


FIGURA 18 - Amostras simuladas com  $n = 5$ ,  $nb=1000$ ,  $p=0,1$ ,  $\tau_p = \sqrt{\frac{0,36}{1000}}$

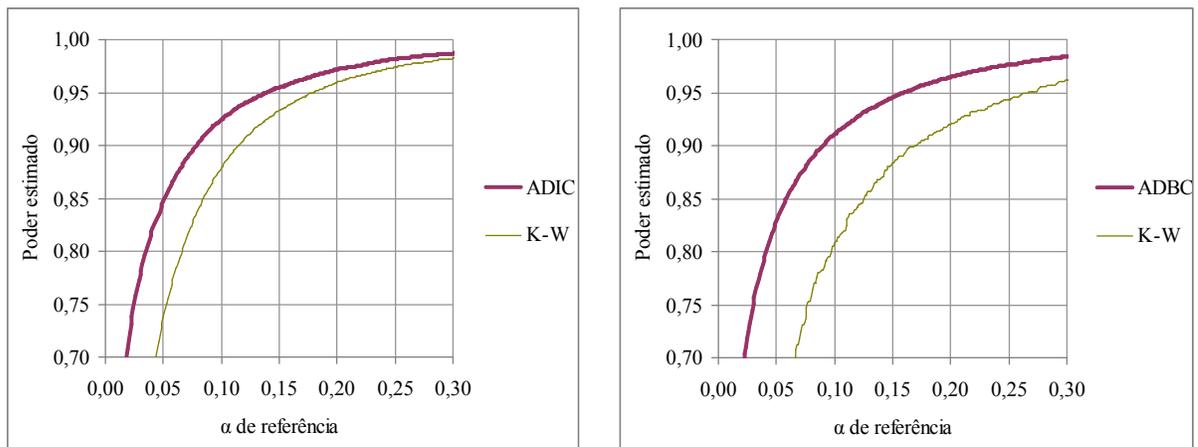


FIGURA 19 - Amostras simuladas com  $n = 5$ ,  $nb=1000$ ,  $p=0,1$ ,  $\tau_p = -\sqrt{\frac{0,36}{1000}}$

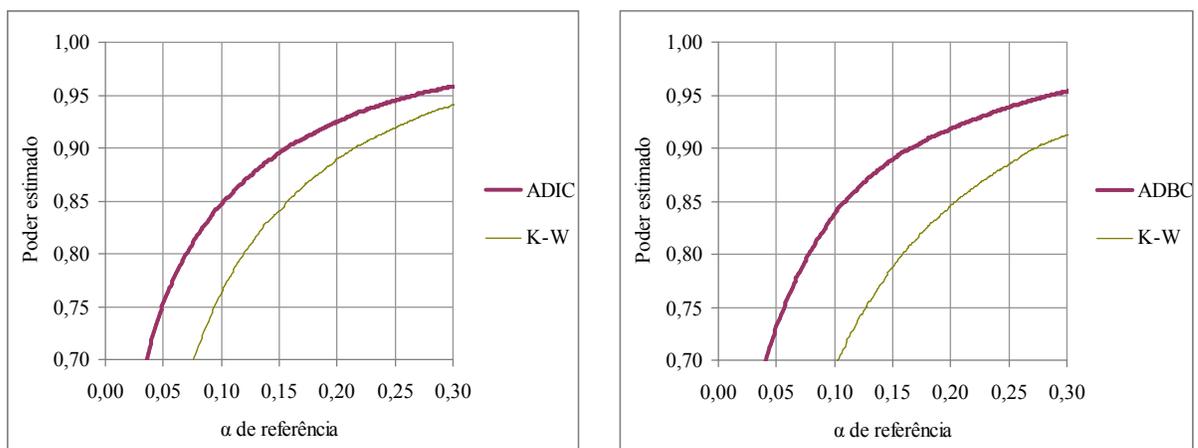


FIGURA 20 - Amostras simuladas com  $n = 5$ ,  $nb=100$ ,  $p=0,1$ ,  $\tau_p = \sqrt{\frac{0,36}{100}}$

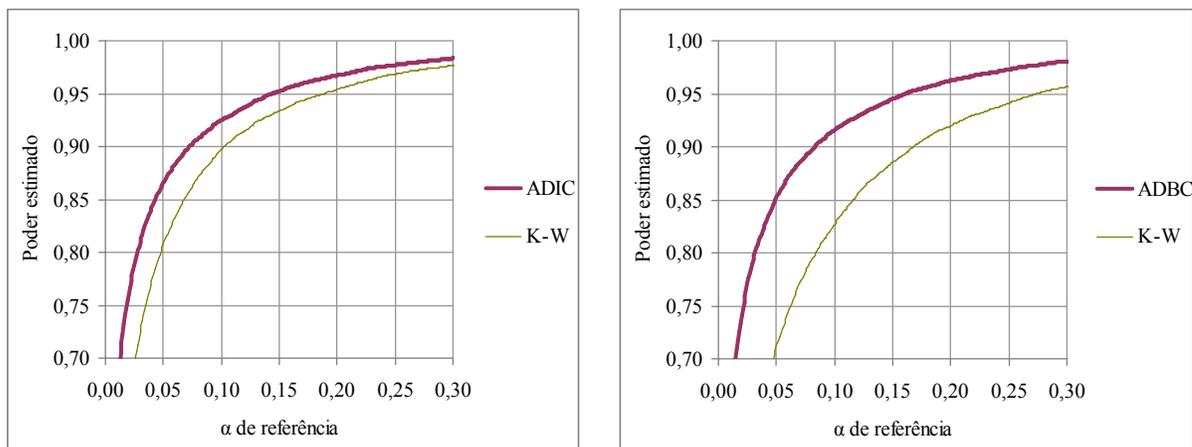


FIGURA 21 - Amostras simuladas com  $n = 10$ ,  $nb=1000$ ,  $p=0,1$ ,  $\tau_p = \sqrt{\frac{0,18}{1000}}$

## 6. Conclusões

O teste paramétrico F da ANOVA obteve resultados mais satisfatórios que os testes não paramétricos de Kruskal-Wallis e de Friedman, mesmo para situações que se distanciam da normalidade. Pode-se observar um maior Pd para o teste F no intervalo estudado ( $\alpha \leq 0,20$ ), além de que os valores de  $\alpha$  estimados por este teste foram muito próximos ao  $\alpha$  de referência, em todos os casos estudados.

Resultado semelhante foi obtido por Feir & Toothaker (1974), onde eles concluíram que o teste de Kruskal-Wallis tem seu Pd afetado pela diminuição do tamanho da amostra e, também, pela diminuição da variância, para baixos níveis de significância ( $\alpha$ ), sendo o teste F da ANOVA capaz de se adaptar melhor às diferentes situações.

A mudança dos parâmetros das distribuições lognormal e binomial provocam maiores alterações no comportamento dos testes em relação à distribuição normal, principalmente para os testes não paramétricos.

Em síntese, não há necessidade de substituir o teste paramétrico F pelos seus respectivos competidores não paramétricos na ausência de normalidade.

## Referências

- CÂMARA, F. G. & SILVA, O. Estatística Não Paramétrica - Testes de Hipóteses e Medidas de Associação. Departamento de Matemática - Universidade dos Açores, Ponta Delgada, 2001.
- CAMPOS, H. Estatística experimental não paramétrica. 2ª ed. Escola Superior de Agricultura “Luiz de Queiroz”, USP. Piracicaba, 1976.
- FEIR, BETTY J. & TOOTHAKER, LARRY E. The ANOVA F-Test Versus The Kruskal-Wallis Test: A Robustness Study. Artigo publicado em: Annual Meeting of the American Educational Research Association. Chicago, 1974.
- MONTGOMERY, D.C. & RUNGER, G.C. Estatística aplicada e probabilidade para engenheiros. Tradução de Verônica Calado. 2ª ed. Rio de Janeiro: Editora LTC, 2003.
- MOORE, D. S. & MCCABE, G. P. Introdução à prática da estatística. 3ª ed. Rio de Janeiro: Editora LTC, 2002.
- SIEGEL, S. & CASTELLAN JR., N. J. Estatística não-paramétrica para ciências do comportamento. Tradução de Sara Ianda Correa Carmona. 2ª ed. Porto Alegre: Editora Artmed, 2006.
- SHIMAKURA, S. E. Bioestatística A. Departamento de Estatística, UFPR. Disponível em: <http://leg.ufpr.br/~shimakur/CE055/>. Acesso em: 29/08/2007.