

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

João Carlos Poloniato Ferreira

O corte do FBST em modelos de alta dimensionalidade

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DES-UFSCar, como parte dos requisitos para obtenção do título de Mestre em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.
VERSÃO REVISADA

Área de Concentração: Estatística

Orientador: Prof. Dr. Rafael Bassi Stern

São Carlos
Janeiro de 2019

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

João Carlos Poloniato Ferreira

The FBST cutoff in high dimensional models

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DES-UFSCar, in partial fulfillment of the requirements for the degree of the Master Interagency Program Graduate in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Rafael Bassi Stern

São Carlos
January 2019



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato João Carlos Poloniato Ferreira, realizada em 03/12/2018:

Rafael Izbicki

Prof. Dr. Rafael Izbicki
UFSCar

Luis Gustavo Esteves

Prof. Dr. Luis Gustavo Esteves
USP

Paulo do Canto Hubert Junior

Prof. Dr. Paulo do Canto Hubert Junior
POLI

À minha irmã, Livia Poloniato, para lhe incentivar a busca pelo conhecimento.
Ao meu orientador e amigo, Rafael Stern, para que continue a lecionar, pesquisar e orientar
com maestria.

AGRADECIMENTOS

Agradeço aos meus pais, Aparecido e Kerli, pelo apoio incondicional.

Agradeço aos professores do DEs-UFSCar e do ICMC-USP, por me transmitirem conhecimento. Em especial, os professores Adriano Polpo, Rafael Izbicki, Rafael Stern por estarem próximos e me acompanharem nesta jornada.

Agradeço aos meus amigos pelo convívio dentro e fora da universidade.

Agradeço à banca pelo tempo disponibilizado e pelas críticas que são sempre bem-vindas.

Agradeço ao programa PIPGEs, pela infra-estrutura oferecida e à CAPES, pelo auxílio financeiro.

“Se eu vi mais longe, foi por estar sobre ombro de gigantes.”
(Isaac Newton)

RESUMO

POLONIATO, J. C. **O corte do FBST em modelos de alta dimensionalidade**. 2019. 47 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

O problema de controlar o nível significância do teste FBST (*Full Bayesian Significant Test*) é estudado no contexto de modelos bayesianos para densidade. Assim, é mostrado um método bayesiano que trabalha com estimação da densidade e como deve ser conduzido o FBST com este método quando deseja-se testar se uma população pode ser dita de determinada distribuição ou testar igualdade de duas populações. Para isso é apresentada a definição do e-valor modificado que é uma maneira alternativa de cálculo da medida de evidência do FBST. Por fim, é feito um estudo de simulação com diferentes distribuições de densidade e analisado o comportamento da função poder do teste nos casos de uma e duas populações.

Palavras-chave: Testes bayesianos, hipóteses precisas, função poder.

ABSTRACT

POLONIATO, J. C. **The FBST cutoff in high dimensional models**. 2019. 47 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

The problem of controlling the significance level of the FBST (Full Bayesian Significant Test) test is studied in the context of Bayesian models for density, thus, a Bayesian method is shown that works with density estimation estimation and how the FBST should be conducted in that situation with this method when it is desired to test if one population has certain density distribution or equality test of two populations. For this, a modified e-value definition is presented that is an alternative to calculate the FBST measure. At end a simulation study with different density distributions and analysis the power function of the test in cases of one and two populations.

Keywords: Bayesian test, precise hypothesis, power function.

LISTA DE ILUSTRAÇÕES

Figura 1 – Comportamento do e-valor de acordo com a quantidade de parâmetros testados	20
Figura 2 – Testando Normal(0,1)	30
Figura 3 – Testando Exp(1)	30
Figura 4 – Testando Normal(0,1) gerando de mistura de 2 normais	31
Figura 5 – Testando Normal(0,1) gerando de mistura de 3 normais	31
Figura 6 – Testando Beta(0,5;0,5)	32
Figura 7 – Testando Beta(1,1)	32
Figura 8 – Exemplo densidade trigonométrica	33
Figura 9 – Função poder testando U(0,1) gerando de distribuição trigonométrica	33
Figura 10 – Exemplo densidade onda quadrada	34
Figura 11 – Função poder testando U(0,1) gerando de distribuição onda quadrada	34
Figura 12 – Testando igualdade de 2 populações com amostras proveniente de uma distribuição normal e $n_1 = n_2 = 100$	36
Figura 13 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição normal e $n_1 = n_2 = 500$	36
Figura 14 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição exponencial e $n_1 = n_2 = 100$	36
Figura 15 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição exponencial e $n_1 = n_2 = 500$	36
Figura 16 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 2 normais e $n_1 = n_2 = 100$	37
Figura 17 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 2 normais e $n_1 = n_2 = 500$	37
Figura 18 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 2 normais e outra amostra de uma Normal(0,1) com $n_1 = n_2 = 100$	37
Figura 19 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 2 normais e outra amostra de uma Normal(0,1) com $n_1 = n_2 = 500$	37
Figura 20 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 3 normais e $n_1 = n_2 = 100$	38
Figura 21 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 3 normais e $n_1 = n_2 = 500$	38

Figura 22 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 3 normais e outra amostra de uma Normal(0,1) com $n_1 = n_2 = 100$	38
Figura 23 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 3 normais e outra amostra de uma Normal(0,1) com $n_1 = n_2 = 500$	38
Figura 24 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (0.5,0.5) e $n_1 = n_2 = 100$	39
Figura 25 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (0.5,0.5) e $n_1 = n_2 = 500$	39
Figura 26 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (1,1) e $n_1 = n_2 = 100$	40
Figura 27 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (1,1) e $n_1 = n_2 = 500$	40
Figura 28 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição trigonométrica e $n_1 = n_2 = 100$	41
Figura 29 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição trigonométrica e $n_1 = n_2 = 500$	41
Figura 30 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição trigonométrica e $n_1 = n_2 = 100$	41
Figura 31 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição trigonométrica e $n_1 = n_2 = 500$	41
Figura 32 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição onda quadrada e $n_1 = n_2 = 100$	42
Figura 33 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição onda quadrada e $n_1 = n_2 = 500$	42
Figura 34 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição onda quadrada e $n_1 = n_2 = 100$	43
Figura 35 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição onda quadrada e $n_1 = n_2 = 500$	43

SUMÁRIO

1	INTRODUÇÃO	19
2	MÉTODOS	21
2.1	<i>Full Bayesian Significance Test</i>	21
2.1.1	<i>Posteriori Normal multivariada</i>	22
2.2	Modelo bayesiano para uma densidade	23
2.3	Cálculo do e-valor no modelo bayesiano para densidade	24
2.4	Teste de igualdade de duas populações	26
2.4.1	<i>Cálculo do e-valor via bootstrap</i>	27
3	SIMULAÇÕES COM UMA POPULAÇÃO	29
3.1	Distribuição normal	29
3.2	Distribuição exponencial	30
3.3	Distribuição mistura de normais	31
3.4	Distribuição beta	32
3.5	Distribuição trigonométrica	32
3.6	Distribuição onda quadrada	33
4	SIMULAÇÕES COM DUAS POPULAÇÕES	35
4.1	Distribuição normal	35
4.2	Distribuição exponencial	36
4.3	Distribuição mistura de normais	36
4.3.1	<i>Mistura de 2 normais</i>	37
4.3.2	<i>Mistura de 3 normais</i>	38
4.4	Distribuição beta	39
4.4.1	<i>Beta (0.5,0.5)</i>	39
4.4.2	<i>Beta(1,1)</i>	39
4.5	Distribuição trigonométrica	40
4.5.1	<i>Duas amostras de uma distribuição trigonométrica</i>	40
4.5.2	<i>Uma amostra da distribuição uniforme e outra da distribuição trigonométrica</i>	41
4.6	Distribuição onda quadrada	41
4.6.1	<i>Duas amostras da distribuição onda quadrada</i>	42

4.6.2	<i>Uma amostra da distribuição uniforme e outra da distribuição onda</i>	
	<i>quadrada</i>	42
5	CONSIDERAÇÕES FINAIS	45
	REFERÊNCIAS	47

INTRODUÇÃO

Em problemas estatísticos é comum deparar-se com o questionamento do pesquisador sobre o comportamento de determinada quantidade populacional. Para fazer inferências sobre esta quantidade de interesse, denominada parâmetro, um dos possíveis procedimentos estatísticos que podem ser usados é o teste de hipótese que mede a consistência dos dados com determinada suspeita relevante chamada de hipótese.

As hipóteses, definidas de acordo com o foco da pesquisa, são escritas como

$$\begin{cases} H_0 : \theta \in \Theta_0, \\ H_1 : \theta \in \Theta_1, \end{cases}$$

ou seja, pergunta-se se o parâmetro de interesse θ pertence a uma espaço Θ_0 ou não.

Tradicionalmente, os testes de hipóteses bayesianos são feitos usando a probabilidade a posteriori do parâmetro que está sendo testado. Nestes testes, assim como na abordagem clássica, deseja-se escolher entre rejeitar ou não a hipótese H_0 . O que ocorre com frequência é o pesquisador estar interessado em testar uma hipótese precisa, isto é, testar uma hipótese tal que $P(\Theta_0) = 0$ ¹. Isso ocorre, por exemplo, ao definir uma hipótese pontual em um espaço paramétrico contínuo. Assim, nestes casos de hipóteses precisas, obtém-se uma posteriori tal que para todo x , $P(\Theta_0|x) = 0$, e desse modo, a decisão será sempre rejeitar H_0 , por ter probabilidade 0 de ocorrer.

A partir da necessidade de testar hipóteses precisas em problemas bayesianos, [Pereira e Stern \(1999\)](#) propuseram o *Full Bayesian Significance Test* (FBST), procedimento de teste de hipótese bayesiano que soluciona esta dificuldade. No FBST, é calculado o e-valor, uma medida de evidência a favor da hipótese nula, medida similar ao p-valor dos testes frequentistas, no

¹ A hipótese precisa também pode ser definida como $\lambda(\Theta_0) = 0$, ou seja, quando o volume do espaço da hipótese nula é 0.

entanto o p-valor é calculado no espaço amostral restrito a Θ_0 enquanto o e-valor é computado considerando todo o espaço paramétrico. A partir desta medida de evidência, é possível tomar a decisão de rejeitar ou não H_0 .

Como motivação para este trabalho tome a Figura 1. Nesta, veja um resultado de simulação de e-valores calculados via FBST com dados gerados sob H_1 . Claramente nota-se que a medida que o número de parâmetros testados conjuntamente aumenta, o e-valor fica mais concentrado próximo de 1 indicando uma decisão favorável a H_0 enquanto na realidade H_0 é falso, levando a uma decisão errônea.

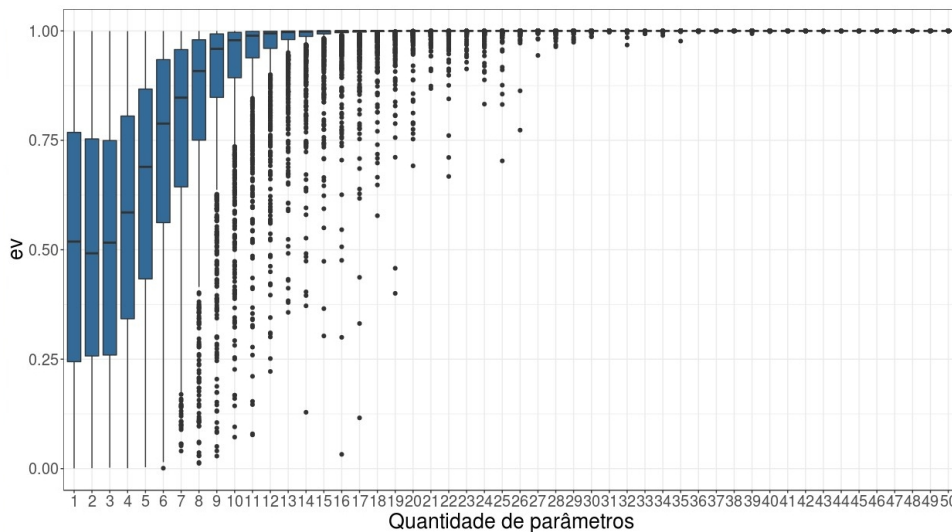


Figura 1 – Comportamento do e-valor de acordo com a quantidade de parâmetros testados

Com isto, o foco deste trabalho é estudar o comportamento do FBST em problemas de alta dimensão, isto é, como calcular a medida de evidência quando deseja-se testar muitos parâmetros ao mesmo tempo e também como definir um corte c , que controla o nível de significância do teste e a usar este corte para tomar a decisão de rejeitar ou não rejeitar H_0 baseado no e-valor.

Neste texto, no Capítulo 2 é explicado mais detalhadamente o procedimento de teste do FBST bem como deve ser calculado o e-valor em uma situação específica de posteriori normal multivariada. Em seguida, é apresentado um cenário em que ocorre uma posteriori de alta dimensão, que é o caso do modelo de estimação bayesiana de uma função de densidade. Nesse, capítulo ainda é discutido porque este cálculo pode ter alguns problemas e é mostrada uma alternativa para calcular o e-valor corrigindo essa situação. Em seguida, nos Capítulos 3 e 4, são apresentados os resultados do estudo de simulação em diversas distribuições para avaliar o desempenho da função poder do FBST e, por fim, no Capítulo 5, são ditas as conclusões tidas até o momento e as propostas futuras.

MÉTODOS

Neste capítulo, é apresentado o FBST e exemplificado como calcular a medida de evidência a favor de H_0 em uma posteriori normal multivariada. Em seguida, é mostrado o método bayesiano usado para atribuir uma posteriori à função densidade de probabilidade e discutido o cálculo do e-valor em casos de uma ou duas populações.

2.1 *Full Bayesian Significance Test*

Para a realização deste teste, primeiramente define-se T_φ como um subconjunto do espaço paramétrico Θ em que a densidade a posteriori é maior ou igual a φ , $\varphi > 0$.

$$T_\varphi = \{\theta \in \Theta | \pi(\theta|x) > \varphi\}.$$

A credibilidade de T_φ é sua própria probabilidade a posteriori,

$$\kappa = \int_{T_\varphi} \pi(\theta|x) d\theta = \int_{\Theta} \pi_\varphi(\theta) d\theta,$$

em que $\pi_\varphi(\theta) = \pi(\theta|x)$ se $\pi(\theta|x) > \varphi$ e zero caso contrário.

Agora define-se π^* como o máximo da densidade a posteriori sob a hipótese nula

$$\theta^* = \arg \max_{\theta \in \Theta_0} \pi(\theta|x), \quad \pi^* = \pi(\theta^*|x).$$

Pode-se definir então $T^* = T_{\pi^*}$ como o conjunto tangente à hipótese nula H_0 com credibilidade κ^* dada por

$$\kappa^* = \int_{T_{\pi^*}} \pi(\theta|x) d\theta$$

A medida de evidência, chamada também de e-valor, a favor de H_0 , é 1 menos a probabilidade à posteriori do espaço T^* (PEREIRA; STERN; WECHSLER, 2008). Assim, a evidência da hipótese nula é

$$ev(H_0|x) = 1 - \kappa^* \quad \text{ou} \quad 1 - \pi(T^*|x).$$

Se a probabilidade da região T^* for “grande”, há indícios que o conjunto tangente tem muitos pontos mais plausíveis que a hipótese nula, evidenciando contra esta hipótese. Por outro lado, se a probabilidade de T^* é “pequena”, então a evidência é favorável à H_0 .

2.1.1 Posteriori Normal multivariada

Em um problema com uma posteriori Normal multivariada, $\mu|X \sim N_d(\mu_d, \Sigma_d)$, suponha que seja de interesse testar

$$\begin{cases} H_0 : \mu = \mu_0, \\ H_1 : \mu \neq \mu_0 \end{cases}$$

Nota-se que esta hipótese é pontual em \mathbb{R}^d , portando o máximo da posteriori $\pi(\mu|x)$ restrita a H_0 é $\pi(\mu_0|x)$. O plano tangente é então o conjunto de pontos que produzem uma posteriori maior ou igual a $\pi(\mu_0|x)$. Para obter o e-valor, primeiramente, deve-se definir as curvas de nível de uma distribuição Normal multivariada que são expressas por

$$R_c(\mu) = (\mu - \mu_d)^T \Sigma_d^{-1} (\mu - \mu_d) \leq c^2.$$

Em seguida, será usado que $R(\mu) = (\mu - \mu_d)^T \Sigma_d^{-1} (\mu - \mu_d) \sim \chi_d^2$ (JONSON; WILKINSON; CHERN, 1992).

Agora, para o cálculo do e-valor, define-se $c_0^2 = (\mu_0 - \mu_d)^T \Sigma_d^{-1} (\mu_0 - \mu_d)$. Assim a curva de nível do máximo a posteriori restrito a H_0 é

$$R_{c_0}(\mu) = (\mu - \mu_d)^T \Sigma_d^{-1} (\mu - \mu_d) \leq (\mu_0 - \mu_d)^T \Sigma_d^{-1} (\mu_0 - \mu_d).$$

Por fim, usando que $R(\mu) \sim \chi_d^2$, o e-valor se resume em

$$ev = 1 - P(\chi_d^2 \leq c_0^2), \tag{2.1}$$

em que $c_0^2 = (\mu_0 - \mu_d)^T \Sigma_d^{-1} (\mu_0 - \mu_d)$.

2.2 Modelo bayesiano para uma densidade

O modelo bayesiano usado para obter uma posteriori da densidade supõe que existem β_0, β_1, \dots tais que

$$f(x) = f_0(x) \sum_{r=0}^{\infty} \beta_r \phi_r(x) \quad (2.2)$$

em que, $f(x)$ é a densidade dos dados a ser estimada, $f_0(x)$ é um chute inicial para $f(x)$ e $\phi_r(x)$ é uma sequência de funções ortonormais a $f_0(x)$.

Brunk (1978) propõe atribuir uma priori nos coeficientes β 's da decomposição proposta em Whittle (1958). Pela Equação (2.2), isto implica atribuir uma priori para $f(x)$. Para que $E[f(x)] = f_0(x)$, deve-se ter que $E[\beta_r] = 0$. Define-se ainda que $VAR(\beta_r) = \sigma_r^2$. Usando a ideia que a suavidade de ϕ_r diminui a medida que o índice r aumenta, deseja-se que os coeficientes β_r que acompanham ϕ_r de alta grandeza sejam próximos de 0, e além disso, que os coeficientes sejam independentes entre si. Assim, define-se $COV(\beta_r, \beta_s) = \frac{\delta_{rs}}{\pi_r}$, com $\pi_r = \frac{1}{\sigma_r^2}$ com

$$\delta_{rs} = \begin{cases} 1, & \text{se } r = s, \\ 0, & \text{se } r \neq s. \end{cases}$$

Escolhendo uma priori normal que atenda as condições acima, ou seja, $\beta_r \sim N(0, \sigma_r^2)$ e a partir da amostra podemos calcular $\bar{\phi}_r = \frac{\sum_{i=1}^n \phi_r(x_i)}{n}$, assintoticamente a verossimilhança é $\bar{\phi}_r \sim N(\beta_r, \frac{1}{n})$ então, pela família conjugada normal normal, tem-se que

$$\beta_r | \bar{\phi}_r \sim N\left(\frac{n}{n + \pi_r} \bar{\phi}_r, \frac{1}{n + \pi_r}\right) \quad (2.3)$$

Na prática, a escolha de uma função $f_0(x)$ ortonormal com uma sequência de funções $\phi_r(x)$ é limitada. Para solucionar isso, é feita uma adaptação usando

$$F_0(x) = \int_{-\infty}^x f_0(t) dt.$$

Assim, fazendo a transformação $\tilde{z} = F_0(\tilde{x})$ implica que $\tilde{z} \in [0; 1]$. Daí a densidade $P(z)$ é

$$P(z) = \frac{f(F_0^{-1}(z))}{f_0(F_0^{-1}(z))} = \frac{f(x)}{f_0(x)} = \sum_{r=0}^{\infty} \beta_r \phi_r(x). \quad (2.4)$$

2.3 Cálculo do e-valor no modelo bayesiano para densidade

Obtidos os dados, é relevante testar se estas observações são provenientes de determinada distribuição de interesse. Para isso deve ser feito um teste de hipótese. Na inferência clássica, existe o teste de Kolmogorov-Smirnov (JR, 1951) que usa como estatística de teste a maior distância entre a função distribuição empírica e a função distribuição da hipótese nula. Uma maneira de conduzir isto de forma bayesiana é encontrar a posteriori da densidade usando o método do Brunk, e a partir desta posteriori como na Equação (2.3), aplicar o FBST para os parâmetros β 's.

Para maior facilidade de empregar o FBST, é feita a transformação $\tilde{z} = F_0(\tilde{x})$ nos dados e com isso a densidade de \tilde{z} é como a Equação (2.4). Assim, se a verdadeira distribuição que originou os dados for F_0 , então Z é uniforme e então todos os β 's da decomposição em séries ortogonais deve ser 0, com exceção do primeiro que é sempre 1 (ver Equação (2.2)). Desse modo a partir da posteriori (2.3) é possível encontrar a expressão do e-valor nesta situação.

Lema 1. Em um problema de densidade ao testar

$$\begin{cases} H_0 : \beta = 0, \\ H_1 : \beta \neq 0 \end{cases}$$

pelo FBST o e-valor é dado por

$$ev_d(x) = 1 - F_{\chi_d^2}(S_{d,n}^2) \quad (2.5)$$

em que

$$S_{d,n}^2 = \sum_{i=1}^d \left(\frac{n\bar{\phi}_i}{\sqrt{n + \pi_i}} \right)^2 \quad (2.6)$$

Demonstração. Usando (2.1) e no cálculo de c_0 substituindo $\mu_0 = 0$ obtém-se $c_0 = (-\mu_d)^T \Sigma_d^{-1} (-\mu_d)$.

Mas usando que a posteriori tem vetor de médias $\mu = \frac{n\bar{\phi}_r}{n + \pi_r}$ e matriz de variância $\Sigma = (n + \pi_r)\mathbb{I}_d$, e portanto $\Sigma^{-1} = \frac{1}{(n + \pi_r)}\mathbb{I}_d$. Dessa forma, $c_0 = S_{d,n}^2$ e o cálculo do e-valor fica como na Equação (2.1).

□

O problema do cálculo exato do e-valor é que em grandes dimensões o e-valor resulta em 1, assim a hipótese nula nunca é rejeitada. Para contornar essa situação o que pode ser feito é mudar a distribuição da estatística de teste.

Teorema 1. Seja $T(X)$ qualquer estatística de teste para H_0 , e seja $g : \mathbb{R} \rightarrow \mathbb{R}$ uma função estritamente monótona. Para todo $0 < \alpha < 1$, considere a função de teste $\phi_{g,\alpha}(x) = \mathbb{I}(g(T(X)) \leq \alpha)$. Então $\phi_{g,\alpha}(x)$ tem significância α para todo $0 < \alpha < 1$ se, e somente se, $g(T) = F_{T|H_0}(T)$ se $g(T)$ for crescente e $g(T) = 1 - F_{T|H_0}(T)$ se $g(T)$ for decrescente, em que $F_{T|H_0}$ é a distribuição acumulada de $T(X)$ sob H_0 .

Demonstração. Para toda $F_{T|H_0}$ contínua, fixado $0 < \alpha < 1$, divide nos casos em que $g(x)$ é crescente e decrescente.

Com $g(x)$ crescente

$$P_{H_0}(\phi_{g,\alpha}(x) = 1) = P_{H_0}(g(T(x)) \leq \alpha) = P_{H_0}(T(x) \leq g^{-1}(\alpha)) = F_{T|H_0}(g^{-1}(\alpha)),$$

então segue que $P_{H_0}(\phi_{g,\alpha}(x) = 1) = \alpha$, isto é, o teste só é rejeitado com probabilidade α se, e somente se $g^{-1} = F_{T|H_0}^{-1}(\alpha)$, implicando que $g(\alpha) = F_{T|H_0}(\alpha)$, para todo α fixado.

Com $g(x)$ decrescente

$$P_{H_0}(\phi_{g,\alpha}(x) = 1) = P_{H_0}(g(T(x)) \leq \alpha) = P_{H_0}(T(x) \geq g^{-1}(\alpha)) = 1 - F_{T|H_0}(g^{-1}(\alpha)),$$

Então para que $P_{H_0}(\phi_{g,h}(x) = 1) = \alpha$ tem que

$$\begin{aligned} 1 - F_{T|H_0}(g^{-1}(\alpha)) &= \alpha \\ F_{T|H_0}(g^{-1}(\alpha)) &= 1 - \alpha \\ g^{-1}(\alpha) &= F_{T|H_0}^{-1}(1 - \alpha). \end{aligned} \tag{2.7}$$

Uma candidata que satisfaz (2.7) é $g(\alpha) = 1 - F_{T|H_0}(\alpha)$, assim,

$$\begin{aligned} g^{-1}(\alpha) &= y = 1 - F_{T|H_0}(\alpha) \\ -1 + y &= -F_{T|H_0}(\alpha) \\ 1 - y &= F_{T|H_0}(\alpha) \\ F_{T|H_0}^{-1}(1 - y) &= \alpha \end{aligned}$$

então $g^{-1}(\alpha) = F_{T|H_0}^{-1}(1 - \alpha) = (1 - F_{T|H_0}(\alpha))^{-1}$ se, e somente se $g(\alpha) = 1 - F_{T|H_0}(\alpha)$.

Resultado que vale para todo α fixado. \square

Corolário 1. A única transformação estritamente crescente do e-valor que leva a um nível α para todo $0 < \alpha < 1$ é dado por $F_{ev(x)|H_0}(ev(x))$

A partir desse corolário, pode ser definida uma maneira de aplicar uma função $g(T)$ no e-valor a fim de controlar o nível de significância do teste.

Definição 1. O e-valor modificado de ordem d é

$$ev_d^*(x) = F_{ev_d|\theta=0}(ev_d(x)). \quad (2.8)$$

Finalmente, sob as mesmas condições do Lema [1](#), segue como é o cálculo do e-valor modificado:

Teorema 2. Ao testar

$$\begin{cases} H_0 : \beta = 0, \\ H_1 : \beta \neq 0 \end{cases}$$

em um problema de densidade pelo FBST o e-valor modificado é

$$ev_d^*(x) = 1 - P(Q_{d,n}^T \leq S_{d,n}^2) \quad (2.9)$$

e novamente com $S_{d,n}^2 = \sum_{i=1}^d \left(\frac{n\bar{\phi}_i}{\sqrt{n+\pi_i}} \right)^2$ e a distribuição sob H_0 de $S_{d,n}^2$ é

$$Q_{d,n}^T = \sum_{i=1}^d \left(\frac{n\varepsilon_i}{\sqrt{n+\pi_i}} \right)^2, \quad (2.10)$$

com $\varepsilon_i \sim N(0, 1)$.

Demonstração. Note que o e-valor é uma função $h(x)$ decrescente (ver Equação [\(2.5\)](#)). O e-valor modificado é uma função $g(x)$ crescente aplicada no e-valor (Equação [\(2.8\)](#)). Portanto a composição $g \circ h = j$ é uma função decrescente. Usando o Teorema [1](#) para funções decrescentes chega que $ev_d^*(x) = 1 - F_{S_{d,n}^2|H_0}(s_{d,n}^2)$, e sob H_0 , $S_{d,n}^2 = Q_{d,n}^T$, isto é, sob H_0 $\bar{\phi}_i \sim N(0, 1)$. \square

2.4 Teste de igualdade de duas populações

Para trabalhar com o teste de igualdade de duas populações considere X e Y duas populações com distribuição $f_x(x|\theta)$ e $f_y(y|\theta)$ respectivamente. Então, a partir de duas amostras (x_1, x_2, \dots, x_n) e (y_1, y_2, \dots, y_m) será testada a hipótese

$$\begin{cases} H_0 : f_x(x|\theta) = f_y(y|\theta) \\ H_1 : f_x(x|\theta) \neq f_y(y|\theta), \end{cases}$$

Para isso, por se tratar de duas populações, o teste mostrado na Seção [2.3](#) não se aplica. Diante disso foram buscadas diversas maneiras de conduzir o teste. Por exemplo, buscou-se encontrar uma estatística de teste exata considerando uma posteriori conjunta para as duas

populações, fazer uma diferenciação dos elementos $\bar{\phi}$ das duas populações e testar como no caso de uma população, mas esses caminhos não mostraram resultados eficientes.

A maneira que mostrou melhores resultados foi usar a distribuição empírica $\hat{F}(y)$ da amostra (y_1, y_2, \dots, y_m) para fazer a transformação da Equação (2.4) aplicando na amostra (x_1, x_2, \dots, x_n) . Foi observado que no caso em que $m \gg n$, ou seja, o tamanho da amostra de Y é muito maior que da amostra de X , o nível de significância é controlado e o poder do teste aumenta ao distanciar de H_0 .

Mas na prática raramente acontece de ter uma amostra muito maior que outra. Chegou-se então no método proposto a seguir (Seção 2.4.1) em que é possível controlar o nível de significância do teste e garantir um poder alto usando de simulação *bootstrap*.

2.4.1 Cálculo do e-valor via bootstrap

Lembre-se que no caso de uma população quando se desejava testar se a amostra provinha de determinada distribuição $f_0(x)$, o que deve ser feito é aplicar a função distribuição acumulada $F_0(x)$ na amostra para em seguida testar se os dados transformados tinham distribuição Uniforme(0,1). Essa ideia pode ser estendida para quando tiver duas amostras (x_1, x_2, \dots, x_n) e (y_1, y_2, \dots, y_m) . Pois se H_0 é verdadeiro, essa transformação faz com que os dados obtidos pela transformação sejam aproximadamente uniformes isto motiva usar o teste de uniformidade.

O que é feito então é formular um teste como se fosse testar a adequabilidade de uma população (que gerou a amostra x) à determinada distribuição. A diferença é que esta distribuição em teste é a empírica obtida através da amostra y .

A partir da amostra (y_1, y_2, \dots, y_m) deve-se encontrar uma estimativa da função distribuição de Y através da empírica $\hat{F}(y)$. Em seguida aplicamos $\hat{F}(y)$ na amostra (x_1, x_2, \dots, x_n) . Observe que agora os dados estão transformados como na Equação (2.4). Agora basta prosseguir com o teste detalhado na Seção 2.3. Ao tomar a decisão de rejeitar H_0 significa que existem evidências para dizer que a distribuição $f(x|\theta)$ é diferente da distribuição $f(y|\theta)$. Caso contrário, as amostras podem ser consideradas de populações iguais.

Acontece que por se tratar de duas amostras a estatística de teste (Equação (2.6)) sob H_0 com distribuição mostrada como na Equação (2.10) não é mais válida. Foi proposto então fazer uma simulação *bootstrap* da distribuição da estatística $S_{d,n}^2$ sob H_0 , isto é, como seria a distribuição da estatística de teste se as duas amostras fossem da mesma distribuição.

Unindo as amostras (x_1, x_2, \dots, x_n) e (y_1, y_2, \dots, y_m) será formado um novo vetor $w = (x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m)$. Em seguida é feita uma amostra $x_{(1)}^*$ com reposição de tamanho n do vetor w e outra amostra $y_{(1)}^*$ com reposição de tamanho m também do vetor w . Perceba que desta maneira, as amostras $x_{(1)}^*$ e $y_{(1)}^*$ vieram da mesma população. Depois disso é estimada a distribuição empírica de $\hat{F}_{(1)}^*(y)$ e aplicado na amostra $x_{(1)}^*$. Por fim o cálculo da estatística de teste $S_{d,n}^2$ baseado na amostra transformada pela estimativa da empírica.

Observe que esse processo pode ser repetido B vezes, tendo assim uma amostra de estatísticas de testes calculadas $B_{d,n,m}^2 = (S_{(1)}^2, S_{(2)}^2, \dots, S_{(b)}^2)$. Note que estas estatísticas foram calculadas com H_0 verdadeiro, pois as amostras são de w (que é uma união das amostras x e y).

Agora basta prosseguir o cálculo do e-valor como na Equação (2.9), mas substituindo a distribuição $Q_{d,n}^T$ pela a distribuição $B_{d,n,m}^2$ encontrada via simulação *bootstrap*. Ou seja, o cálculo do e-valor no problema de teste de igualdade de duas populações é

$$ev_d^*(x) = 1 - P(B_{d,n,m}^2 \leq S_{d,n}^2)$$

SIMULAÇÕES COM UMA POPULAÇÃO

Neste capítulo são apresentados os resultados obtidos de um estudo de simulação feito com diferentes distribuições. A partir de uma amostra de dados deseja-se saber se estes são provenientes de determinada distribuição ou não. Desta forma, o que é feito é testar a hipótese

$$\begin{cases} H_0 : X \sim f(x|\theta_0) \\ H_1 : X \sim g(x|\theta), \end{cases}$$

com $g(x) \in L^2$, isto é, $\int g(x|\theta)^2 dx < \infty$.

Para isto, sempre na mesma amostra, foram aplicados dois testes: o FBST, calculando o e-valor modificado e o teste de Kolmogorov-Smirnov que foi escolhido por ser um teste não-paramétrico bastante usado quando se deseja testar igualdade de distribuições de probabilidade.

Foi definido um corte $c = 0,05$ para o qual todo e-valor modificado ou p-valor menor que c o teste rejeita H_0 . Daí, variando o parâmetro da distribuição em questão foram geradas 1000 amostras de tamanho 100, e verificado se o teste rejeita ou não H_0 em cada uma das amostras, com isto, por fim, foi calculada a probabilidade de rejeitar H_0 e então traçada a função poder dos testes. A sequência de funções ortonormais usadas foi $\phi_r = \sqrt{2} \cos(\pi r x)$ e os pesos a priori são $\pi_r = 5^r$

3.1 Distribuição normal

Com observações geradas de uma distribuição Normal de variância 1 e mudando a média μ no intervalo $[-1; 1]$, foi testado se os dados são de uma distribuição normal padrão. A Figura 2a mostra a função poder do FBST e do KS e para melhor visualização foi calculada a diferença entre os poderes (Figura 2b). Pode-se observar que neste caso o poder dos testes são bem semelhantes e não há indícios que algum sobressaia sobre o outro. Nota-se ainda que quando

$\mu = 0$, que é a hipótese em teste, o poder é 0,05, confirmando que o nível de significância de ambos estão sendo controlados.

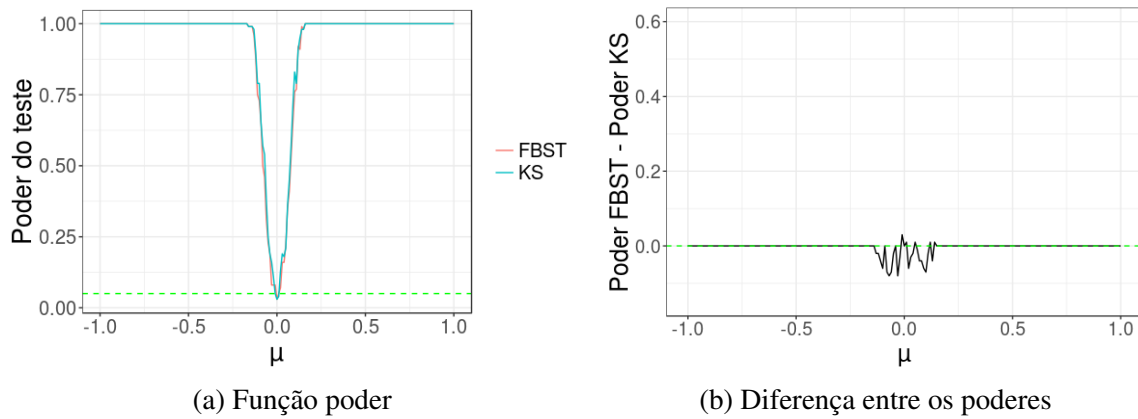


Figura 2 – Testando Normal(0,1)

3.2 Distribuição exponencial

Em seguida, foram simulados dados de uma distribuição exponencial variando λ de 0,01 a 2. A hipótese em teste é se as observações provêm de uma distribuição Exponencial com $\lambda = 1$. É possível ver na Figura 3a que as funções poder de ambos os testes estão visualmente sobrepostas. Já na Figura 3b observa-se que existe uma pequena diferença variando no intervalo de -0,1 a 0,05, confirmando que os poderes neste caso são bem próximos. Nota-se novamente que o poder é 0,05 quando $\lambda = 1$, que é a hipótese em teste.

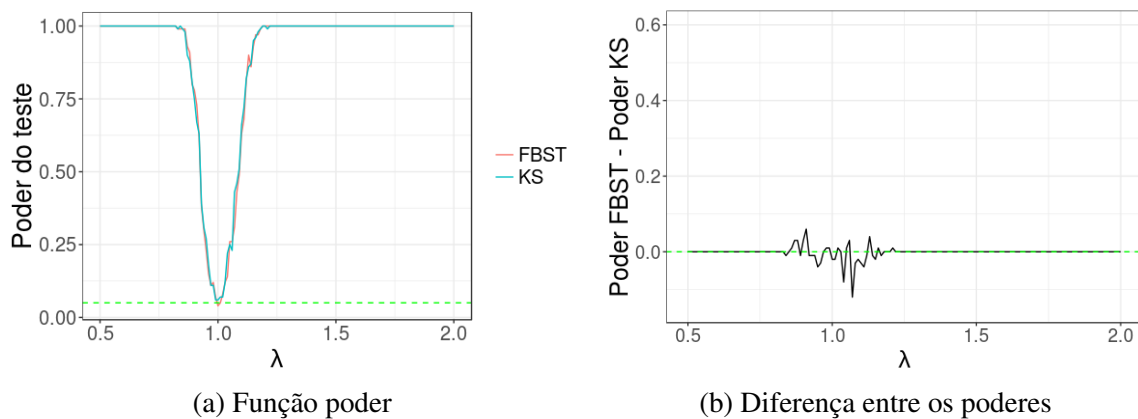


Figura 3 – Testando Exp(1)

3.3 Distribuição mistura de normais

Na Figura 4a, é mostrado a função poder quando a distribuição em teste é uma Normal padrão, no entanto os dados foram simulados de uma distribuição bimodal com média sempre centrada em 0 e modas equidistantes. Observa-se que a medida que as modas se distanciam de 0 aumenta a probabilidade de rejeitar H_0 , comportamento esperado pois fica mais fácil de identificar que não se trata de uma Normal(0,1). É possível ver ainda que o FBST consegue identificar que não se trata da distribuição em teste melhor que o KS (Figura 4b).

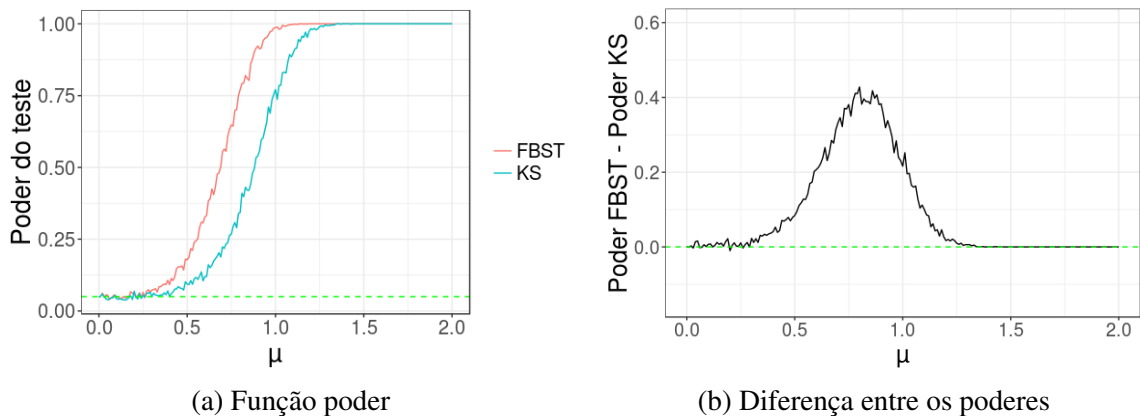


Figura 4 – Testando Normal(0,1) gerando de mistura de 2 normais

Quando foram gerados dados de uma distribuição que mistura 3 normais (uma Normal com moda em 0 e as duas outras se distanciando de 0 para mais e para menos conjuntamente) observa-se um comportamento semelhante a situação bimodal. Na Figura 5a e 5b nota-se que entre 0,5 e 1,5 o FBST tem poder maior que o KS, ainda destaca-se que a função poder do KS demora mais para atingir o 1 que no caso anterior.

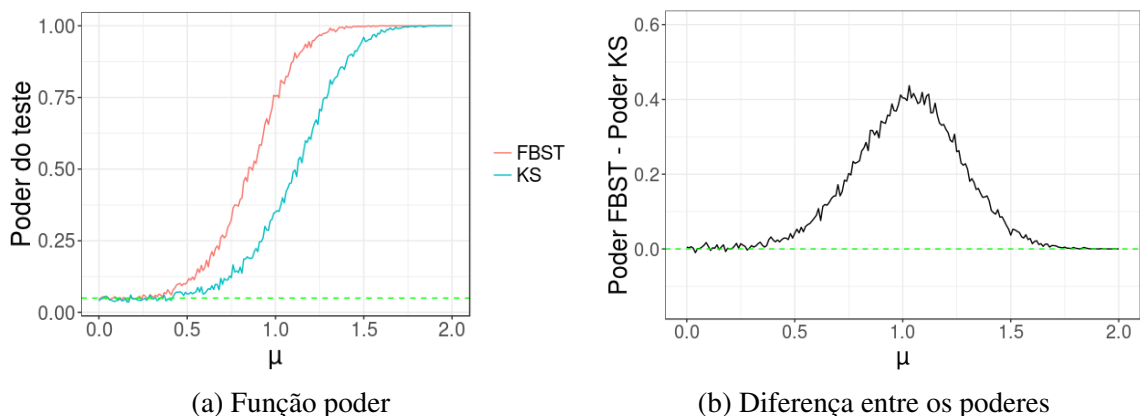


Figura 5 – Testando Normal(0,1) gerando de mistura de 3 normais

3.4 Distribuição beta

Para a distribuição Beta, foi escolhido testar se os dados são de uma $Beta(0,5;0,5)$, os parâmetros α e β variam igualmente de 0,01 a 3. A Figura 6a mostra que o nível de significância dos testes estão sendo controlados com o corte $c = 0,05$, nota-se ainda que em um espaço de variação do parâmetro antes e depois de 0,5 o poder do FBST é maior que do KS, a Figura 6b confirma isso.

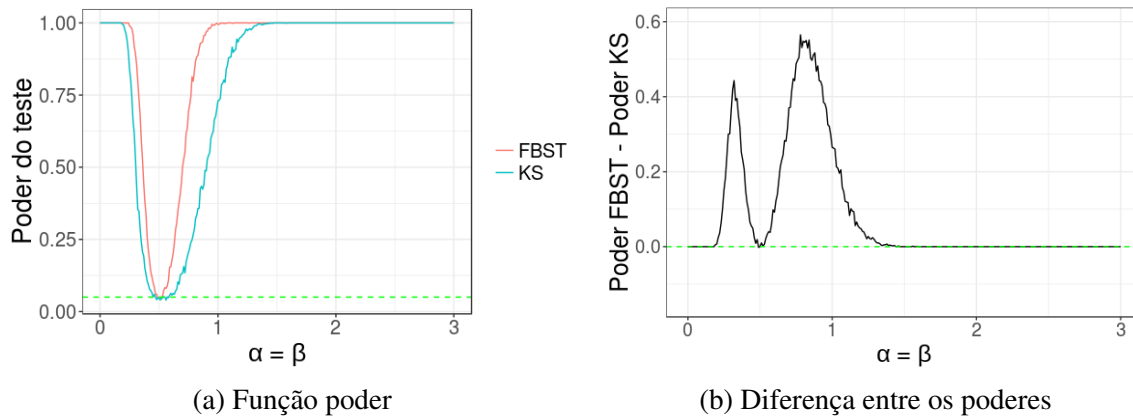


Figura 6 – Testando $Beta(0,5;0,5)$

Ao testar uma $Beta(1,1)$ foi observado um comportamento parecido com o caso anterior, novamente a função poder do FBST domina a função poder do KS (Figura 7a), o que é constatado ao observar que a diferença entre os poderes do FBST e KS nunca é negativa (Figura 7b).

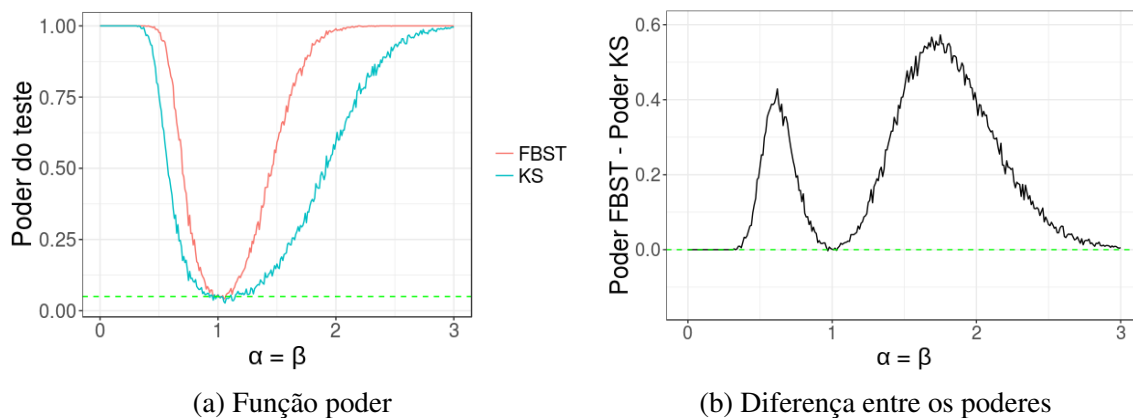


Figura 7 – Testando $Beta(1,1)$

3.5 Distribuição trigonométrica

Em seguida, foram gerados dados da distribuição $f(x|k) = \frac{2}{\pi} |\text{sen}(\pi kx)|$ e testar se são provenientes de uma distribuição uniforme padrão. O comportamento desta distribuição é como

na Figura 8a e 8b, quanto maior o valor de k menos suave é a densidade e “parece” mais com uma densidade da uniforme no intervalo $[0;1]$.

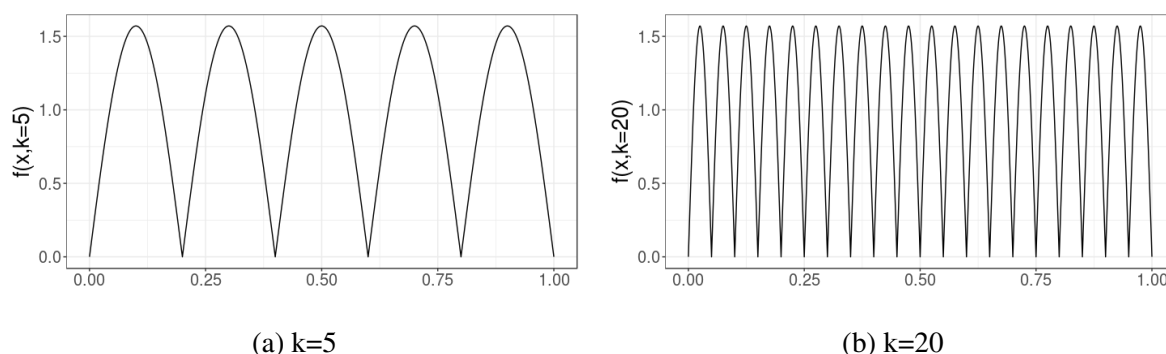


Figura 8 – Exemplo densidade trigonométrica

Para este teste, por se tratar de uma densidade trigonométrica foi feito a estimação da densidade com diferentes pesos para a série ortogonal. Foram considerados $\pi_r = 2^r$, $\pi_r = 5^r$, $\pi_r = r^2$ e $\pi_r = r^5$. Nos resultados da função poder pode ser observado que com um tamanho de amostra $n = 100$ o peso a priori que teve melhores resultados foi r^2 . A partir de $k = 5$ já é possível notar que os demais pesos e o teste KS já tem um poder bem baixo comparado com o peso r^2 . Ainda é possível ver o que nível de significância do teste é controlado em 5% todos os casos.

Com uma amostra de tamanho 1000 é possível ver que os testes tem maior poder que no caso anterior (obviamente devido ao tamanho da amostra), mas ainda com amostra de tamanho maior o peso r^2 é o que mantém maior poder do teste.

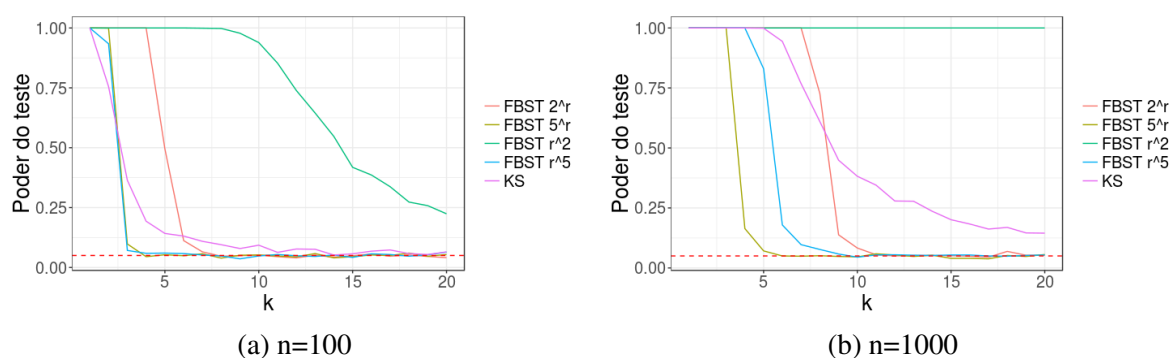


Figura 9 – Função poder testando $U(0,1)$ gerando de distribuição trigonométrica

3.6 Distribuição onda quadrada

Mais uma distribuição analisada foi que o pode ser chamada de onda quadrada, a ideia dessa distribuição é criar 2^k intervalos entre 0 e 1 e nesses intervalos alternar atribuir massa de

probabilidade positiva ou nula, essa densidade pode ser escrita como

$$f(x|k) = \begin{cases} 2, & \text{se } (10 \times \lfloor xk \rfloor) \bmod 2 = 0 \\ 0, & \text{caso contrário.} \end{cases} \quad x \in [0, 1], \quad k = 0, 1, \dots$$

A Figura 10a e 10b mostra a uniforme dividida em $2^1 = 2$ e $2^4 = 16$ intervalos respectivamente, a medida que aumenta o número de intervalos criados a distribuição mais se aproxima de uma uniforme, sendo assim, espera-se que quanto maior o número de k , menor o poder do teste já que fica mais complicado de diferenciar de uma uniforme.

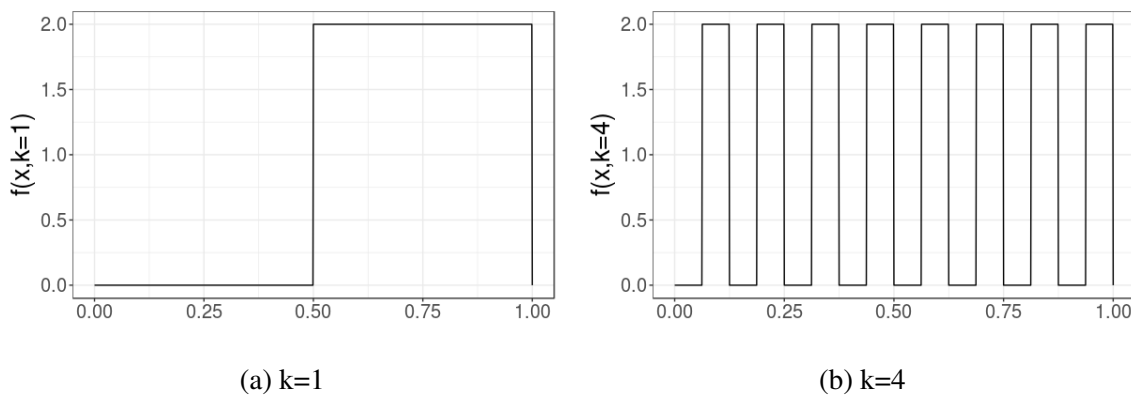


Figura 10 – Exemplo densidade onda quadrada

Novamente foi trabalhado com diferentes pesos a priori para os coeficientes da série ortonormal pelo fato de não se tratar de uma densidade suave. Nas Figuras 11a e 11b observa-se que para ambos tamanhos de amostra o peso que resulta em maior poder do teste é $\pi_r = r^2$, nota-se ainda que o teste KS e o FBST com os outros pesos a priori tem desempenho bem próximos tanto para amostra de tamanho 100 quanto 1000.

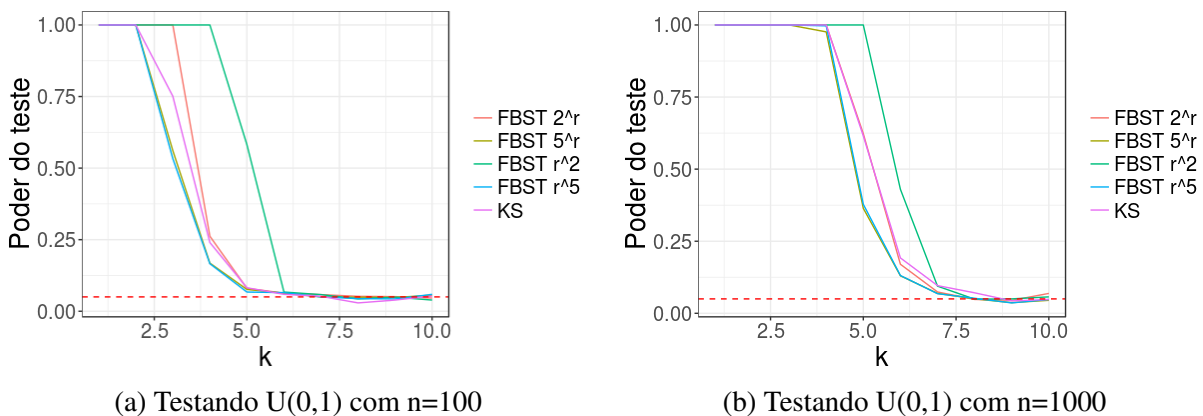


Figura 11 – Função poder testando $U(0,1)$ gerando de distribuição onda quadrada

SIMULAÇÕES COM DUAS POPULAÇÕES

Neste capítulo será mostrado os estudos de simulação feitos para o teste de igualdade de duas populações. Desta forma o que está sendo testado é

$$\begin{cases} H_0 : f_x(x|\theta) = f_y(x|\theta) \\ H_1 : f_x(x|\theta) \neq f_y(x|\theta). \end{cases}$$

A cargo de comparação foi traçada a função poder do KS para duas populações e a função poder do teste FBST proposto na Seção 2.4. Para isso foram replicadas 1000 amostras de tamanho 50, 100 e 500 de duas populações. Para o cálculo do e-valor via *bootstrap* foram feitas 500 reamostragens em todos os tamanhos de amostra. O corte definido foi $c = 0,05$ e a sequência de funções ortonormais usadas foi $\phi_r = \sqrt{2}\cos(\pi r x)$ e os pesos a priori são $\pi_r = 5^r$

4.1 Distribuição normal

Para a distribuição normal foi fixada uma população com distribuição $Normal(0, 1)$ e a outra população com distribuição $Normal(\mu, 1)$. Nesta situação também foi analisada a função poder do teste-T de igualdade de médias com variância conhecidas. Perceba nas figuras 12 e 13 que sob H_0 , isto é, quando $\mu = 0$ da população 2, fazendo assim com que as populações tenham mesma distribuição, o nível de significância do teste está controlado em 0,05. Ademais é possível ver que as funções poder dos testes FBST, KS e T apresentam comportamentos bem próximos e não é possível dizer que algum teste se sobressai em nenhum tamanho de amostra.

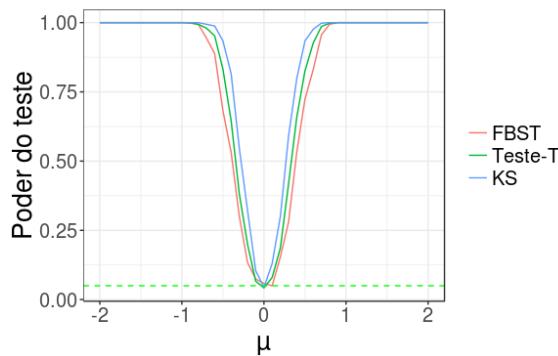


Figura 12 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição normal e $n_1 = n_2 = 100$

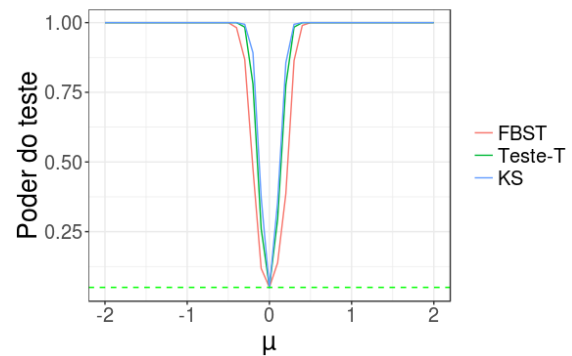


Figura 13 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição normal e $n_1 = n_2 = 500$

4.2 Distribuição exponencial

Para a distribuição exponencial, as amostras da população 1 foram geradas sempre da distribuição $Exp(2)$ e as da outra população de uma distribuição $Exp(\lambda)$ com $\lambda \in (1, 5)$.

Nota-se nas figuras 14 e 15 que o comportamento das funções poderes do FBST e do KS para igualdade de duas populações são bem semelhantes. Ressalta-se ainda conforme o tamanho da amostra é maior é possível ver que a função poder dos testes vai pra 1 mais rápido, confirmando a ideia de que quanto mais observações a decisão é mais precisa.

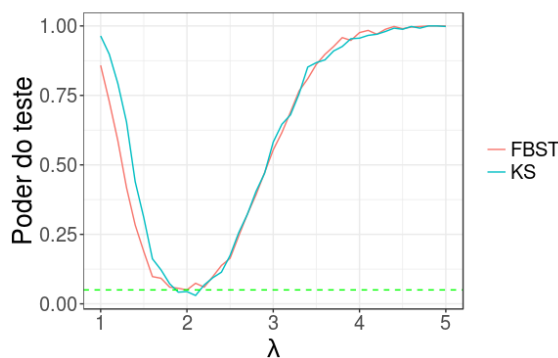


Figura 14 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição exponencial e $n_1 = n_2 = 100$

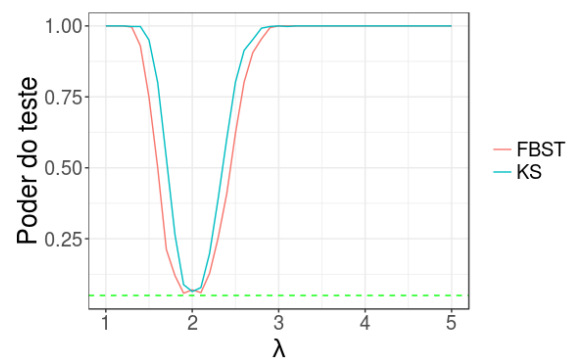


Figura 15 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição exponencial e $n_1 = n_2 = 500$

4.3 Distribuição mistura de normais

Considerando a distribuição de mistura de normais, foi considerado 2 cenários em cada uma das distribuições. Um primeiro caso, gerando duas amostras da mesma distribuição de mistura. Segundamente uma amostra da distribuição $Normal(0, 1)$ e outra de mistura de normais.

4.3.1 Mistura de 2 normais

Aqui, foi fixado uma população que mistura $Normal(-1, 1)$ e $Normal(1, 1)$, e outra população foi considerada mistura de $Normal(-\mu, 1)$ e $Normal(\mu, 1)$. Com isso, foi trabalhado com uma distribuição que começa com duas médias em 0 e depois as médias vão ficando equidistantes de 0.

É possível ver que para amostras de tamanho 100 (Figura 16) o FBST mostrou ter um poder maior; já na Figura 17 é possível ver que as funções poder estão próximas. Isto mostra que para amostras grandes os desempenhos dos testes são iguais.

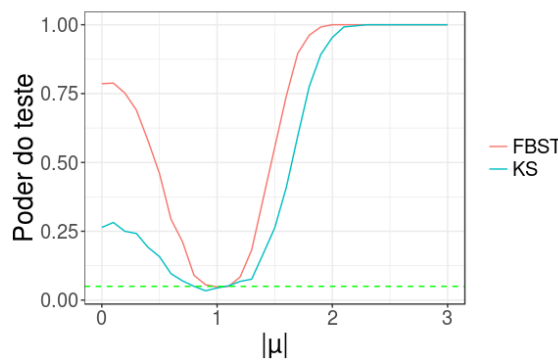


Figura 16 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 2 normais e $n_1 = n_2 = 100$

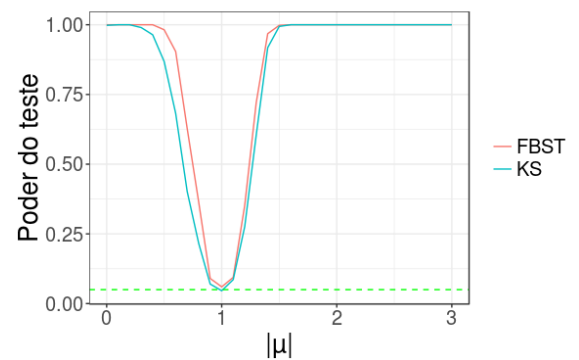


Figura 17 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 2 normais e $n_1 = n_2 = 500$

No próximo cenário, mostrado nas figuras 18 e 19, uma amostra foi gerada de uma distribuição $Normal(0, 1)$ a outra população foi considerada mistura de $Normal(-\mu, 1)$ e $Normal(\mu, 1)$. Nota-se aqui que o poder inicia em 0,05 quando H_0 é verdadeiro. Quanto mais distante de H_0 maior o poder para ambos os testes, mas o FBST chega antes que o KS.

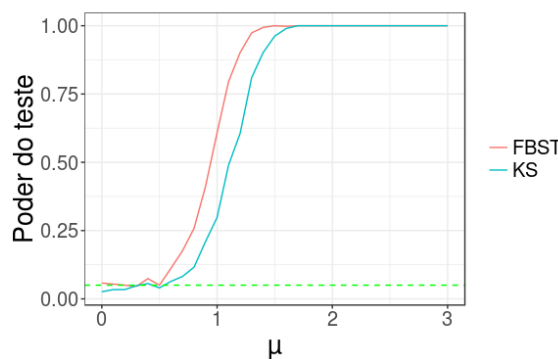


Figura 18 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 2 normais e outra amostra de uma $Normal(0, 1)$ com $n_1 = n_2 = 100$

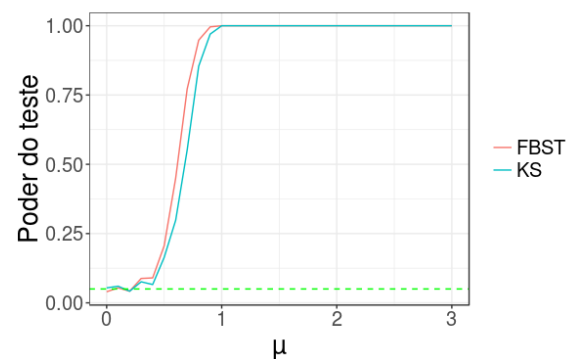


Figura 19 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 2 normais e outra amostra de uma $Normal(0, 1)$ com $n_1 = n_2 = 500$

4.3.2 Mistura de 3 normais

Nesta situação, com mistura de 3 normais uma das amostras é sempre provida de uma distribuição que mistura $Normal(-1, 1)$, $Normal(0, 1)$ e $Normal(1, 1)$ e a outra amostra é de uma mistura de $Normal(-\mu, 1)$, $Normal(0, 1)$ e $Normal(\mu, 1)$. Com isso, foi variado μ no intervalo de 0 a 3 e traçado a função poder dos testes.

Na Figura 20, é possível notar que o FBST para 2 populações tem um poder de teste maior que o KS para duas populações. A Figura 21 mostra que com amostra de tamanho 500 o poder dos testes são equivalentes.

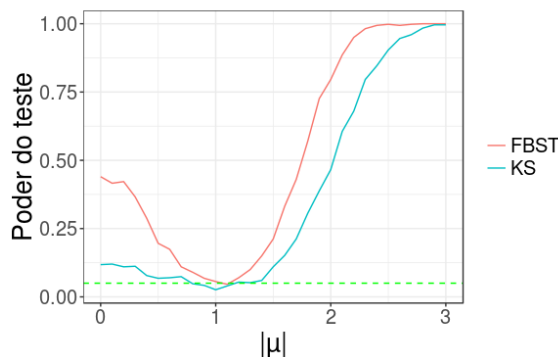


Figura 20 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 3 normais e $n_1 = n_2 = 100$

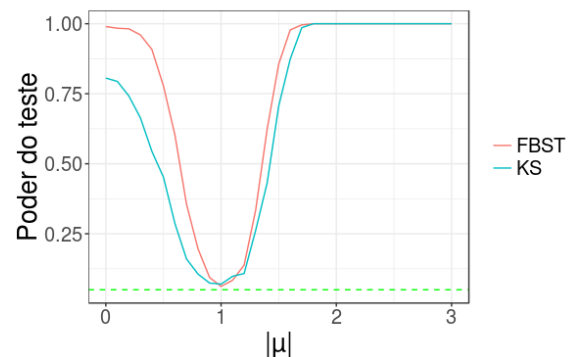


Figura 21 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição de mistura de 3 normais e $n_1 = n_2 = 500$

Testando igualdade de populações baseado em uma amostra de da distribuição $Normal(0, 1)$ e outra de mistura de 3 normais assim como anteriormente. Percebe-se que o nível de significância de ambos os testes está controlado mas que quando H_0 é falso tem um intervalo em que o FBST apresenta probabilidade de rejeitar H_0 maior que o KS (figuras 22 e 23).

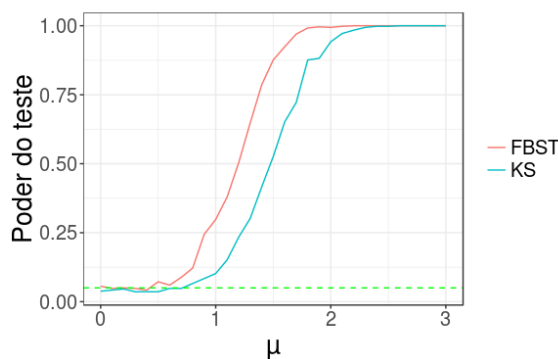


Figura 22 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 3 normais e outra amostra de uma $Normal(0,1)$ com $n_1 = n_2 = 100$

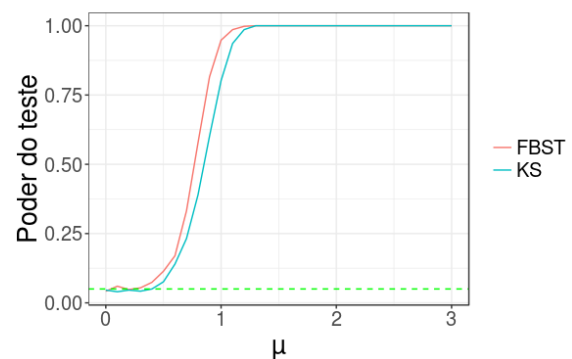


Figura 23 – Testando igualdade de 2 populações com uma amostra proveniente de uma distribuição de mistura de 3 normais e outra amostra de uma $Normal(0,1)$ com $n_1 = n_2 = 500$

4.4 Distribuição beta

Com a distribuição Beta, foi trabalhado os casos em que a distribuição não é unimodal que são na distribuição $Beta(0.5, 0.5)$ e $Beta(1, 1)$.

4.4.1 Beta (0.5,0.5)

O primeiro cenário é com simulações de uma amostras da distribuição $Beta(0.5, 0.5)$ e outra amostra da distribuição $Beta(\alpha, \beta)$ com $\alpha = \beta \in (0.5, 5)$ variando conjuntamente.

Nas figuras 24 e 25 nota-se que o poder do teste FBST está sendo controlado sob H_0 e nota-se ainda que o poder o FBST para igualdade de fuas populações cresce e chega a 1 mais rapidamente que o KS. Isto acontece em todos os tamanhos de amostra mas é mais forte quando se têm amostra pequena.

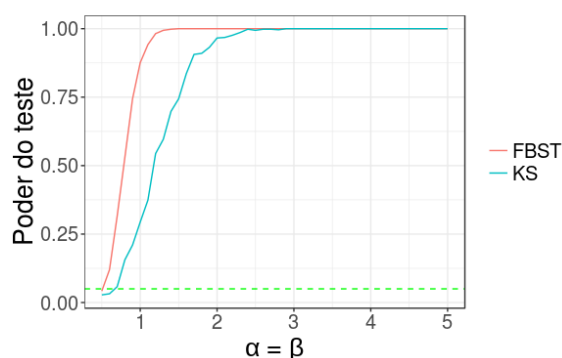


Figura 24 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (0.5,0.5) e $n_1 = n_2 = 100$

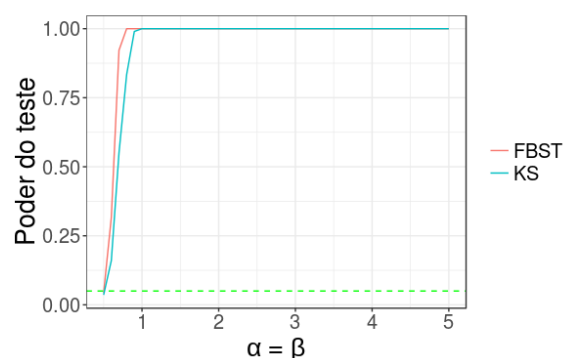


Figura 25 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (0.5,0.5) e $n_1 = n_2 = 500$

4.4.2 Beta(1,1)

Aqui uma das amostras geradas sempre vêm distribuição $Beta(1, 1)$ e outra amostra da distribuição $Beta(\alpha, \beta)$ com $\alpha = \beta \in (0.5, 5)$ variando ao mesmo tempo.

Nas figuras 26 e 27 é possível ver que a função poder do teste do FBST é superior do que a do KS, ainda que para amostras grandes a diferença entre os testes é menor. Confirma-se então que o teste proposto é tem poder maior que o usual KS e ainda é aplicável em amostras pequenas.

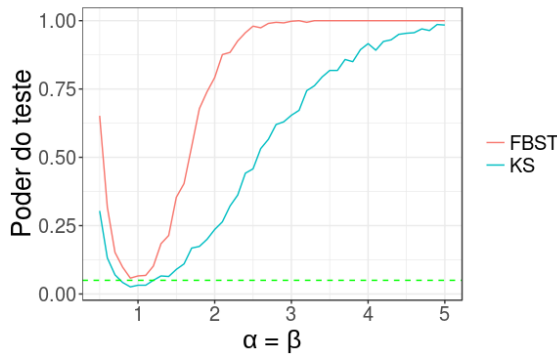


Figura 26 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (1,1) e $n_1 = n_2 = 100$

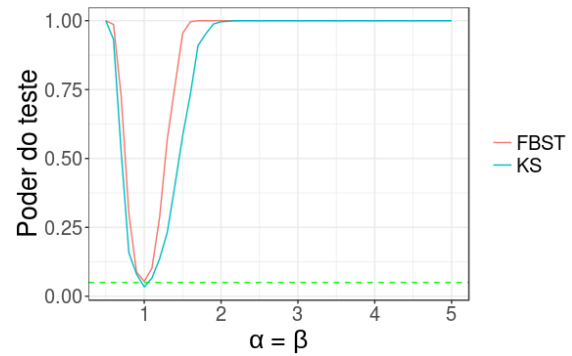


Figura 27 – Testando igualdade de 2 populações com amostras provenientes da distribuição beta, fixado uma população com distribuição Beta (1,1) e $n_1 = n_2 = 500$

4.5 Distribuição trigonométrica

A distribuição usada neste caso foi $f(x|k) = \frac{2}{\pi} |\text{sen}(\pi kx)|$. Para lembrar o comportamento da função veja a Figura 9a e a Figura 9b. Por se tratar de uma densidade menos suave foi feito a estimação da densidade com diferentes pesos a priori para a série ortogonal. Foram considerados $\pi_r = 2^r$, $\pi_r = 5^r$, $\pi_r = r^2$ e $\pi_r = r^5$.

Ao trabalhar com esta distribuição trigonométrica foi analisado dois cenários, um gerando as duas amostras de mesma distribuição trigonométrica e outro comparando esta distribuição com uma $Uniforme(0, 1)$.

4.5.1 Duas amostras de uma distribuição trigonométrica

Considerando um amostra da distribuição trigonométrica sempre com $K = 4$ e a outra amostra da mesma distribuição variando o parâmetro K de 1 a 30.

É possível ver que sob H_0 , ou seja, quando $K = 4$ o poder do teste é 0,05, mostrando assim que o poder do teste está sob controle. Observa-se ainda que o FBST com o peso a priori $\pi_r = 5^r$ foi aquele que mostrou ter poder maior que os demais testes para todos os tamanhos de amostra (figuras 28 e 29).

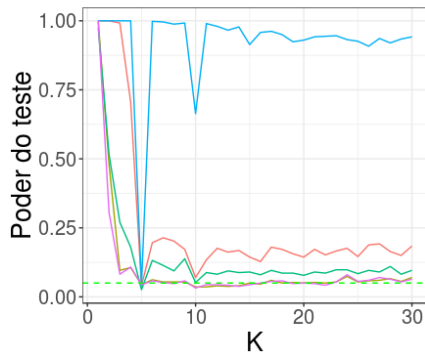


Figura 28 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição trigonométrica e $n_1 = n_2 = 100$

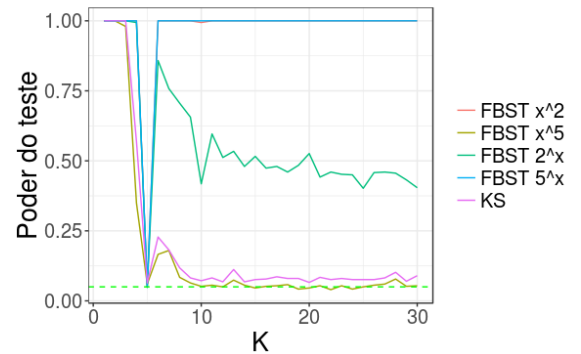


Figura 29 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição trigonométrica e $n_1 = n_2 = 500$

4.5.2 Uma amostra da distribuição uniforme e outra da distribuição trigonométrica

Fixando uma amostra sempre da distribuição $Uniforme(0,1)$ e a outra amostra da distribuição trigonométrica variando o valor de K . Vale lembrar que a medida que o valor de K aumenta, mais a distribuição fica parecida com uma uniforme, mas nunca igual, portanto o teste que mais manter um poder maior é o mais eficiente. Nas figuras [30](#) e [31](#) observa-se que o FBST com $\pi_r = 5^r$ é o teste que apresenta maior função poder. É possível ver ainda que quanto maior o tamanho da amostra o poder do teste demora mais até o decaimento para 0,05.

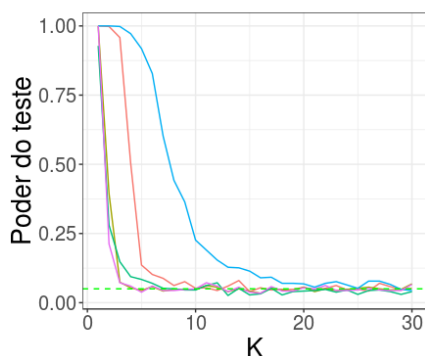


Figura 30 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição trigonométrica e $n_1 = n_2 = 100$

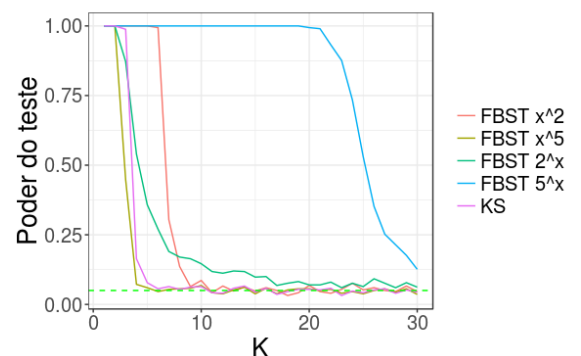


Figura 31 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição trigonométrica e $n_1 = n_2 = 500$

4.6 Distribuição onda quadrada

Para estudar o comportamento do teste na distribuição onda quadrada que é expressa por

$$f(x|k) = \begin{cases} 2, & \text{se } (10 \times \lfloor xk \rfloor) \bmod 2 = 0 \\ 0, & \text{caso contrário.} \end{cases} \quad x \in [0, 1], \quad k = 0, 1, \dots$$

Para relembrar o comportamento da função veja a Figura 10a e a Figura 10b. Também foi escolhido trabalhar com os diferentes pesos a priori $\pi_r = 2^r$, $\pi_r = 5^r$, $\pi_r = r^2$ e $\pi_r = r^5$ por se tratar de uma função pouco suave.

Foram feitas simulações considerando duas populações provenientes da mesma distribuição onda quadrada e outro caso com uma amostra da distribuição uniforme tradicional e outra da onda quadrada.

4.6.1 Duas amostras da distribuição onda quadrada

Ao simular dados da mesma distribuição, a amostra 1 foi fixada sempre de da distribuição com $K = 4$ e a outra amostra variando o K de 1 a 30. Percebe-se que o nível de significância do teste está sendo controlado em $H_0 = K = 4$. Ainda é notável que para amostras pequenas o teste resulta um poder bem baixo fora de H_0 mas que o teste com melhores resultados é o FBST com $\pi_r = 5^r$. Para amostras de tamanho 500 este foi o único teste que consagrou $P(\text{rejeitar } H_0 | H_0 \text{ falso}) = 1$.

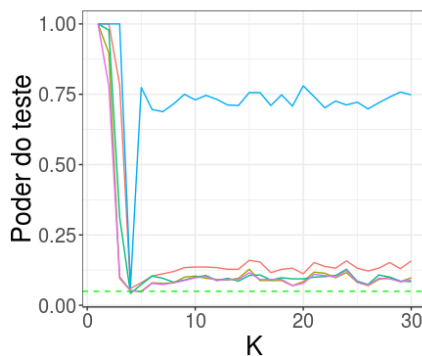


Figura 32 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição onda quadrada e $n_1 = n_2 = 100$

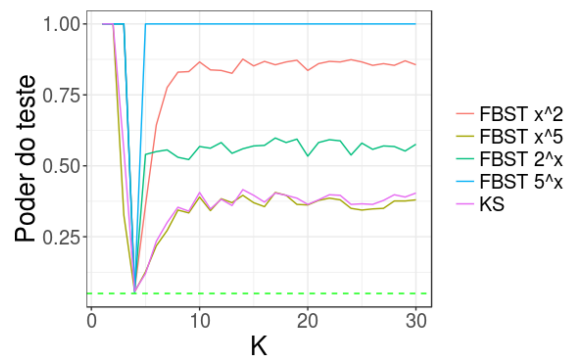


Figura 33 – Testando igualdade de 2 populações com amostras provenientes de uma distribuição onda quadrada e $n_1 = n_2 = 500$

4.6.2 Uma amostra da distribuição uniforme e outra da distribuição onda quadrada

Ao simular sempre dados de uma distribuição $Uniforme(0, 1)$ e testando a igualdade de populações com dados vindos de uma onda quadrada é possível ver que quase todos os testes tiveram desempenho iguais (figuras 34 e 35). Percebe-se que o poder do teste diminuiu rapidamente para 0,05 com todos os testes nos diferentes tamanhos de amostra. Contudo, ainda

que levemente, é factível dizer que o FBST com $\pi_r = 5^r$ teve melhor desempenho, justamente por sua curva da função poder estar substancialmente mais alta que os demais testes.

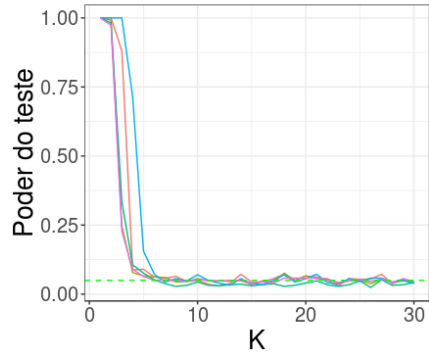


Figura 34 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição onda quadrada e $n_1 = n_2 = 100$

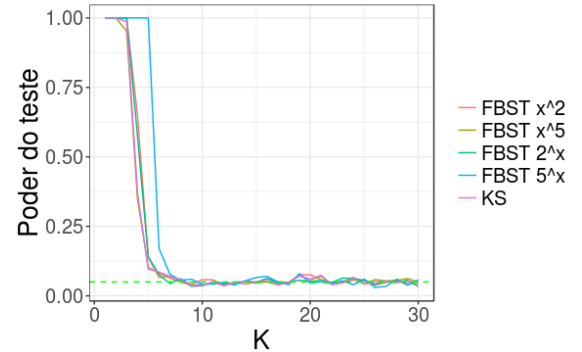


Figura 35 – Testando igualdade de 2 populações entre uma população uniforme e outra com distribuição onda quadrada e $n_1 = n_2 = 500$

CONSIDERAÇÕES FINAIS

Neste trabalho, foi debatido o FBST e o cálculo do e-valor em problemas de alta dimensionalidade. Foi apresentado o e-valor modificado, que é uma transformação no e-valor usual mais fácil de definir o corte para controlar o nível de significância do teste. A partir dessa estatística de teste apresentada para o teste de adequabilidade de uma população, foi estendido o conceito para um teste de igualdade de duas populações e feitas as devidas correções necessárias para apresentar um teste capaz de ter poder alto e o nível de significância definido.

Nas simulações feitas para o teste de uma população foi observado que em problemas com dados provenientes de uma distribuição unimodal, o desempenho do FBST modificado é semelhante ao teste KS, já consagrado na estatística. No entanto, ao trabalhar com uma densidade geradora dos dados com comportamento atípico, basicamente distribuições multimodais e/ou menos suaves o poder do FBST é maior em alguns intervalos e ademais os poderes são iguais entre o FBST e o KS, isto devido ao método utilizado para obter a posteriori da densidade usar séries ortonormais e assim ser mais preciso para estas distribuições.

Um comportamento semelhante foi visto nos resultados de simulações feitas com o teste para igualdade de duas populações. Observou-se que o desempenho em distribuições unimodais é semelhante entre o KS e o FBST. Ainda foi visto que neste caso a escolha do peso a priori não interfere no poder do teste FBST. Já no caso em que se trata de uma distribuição menos “comportada”, claramente o FBST apresenta um poder maior. Foi visto ainda que a escolha do peso a priori é importante e responsável pelo desempenho do FBST para duas populações.

Como proposta para trabalhos futuros, fica investigar as propriedades teóricas do problema, vistas aqui por meio de simulação, como, por exemplo, explicar porque o peso a priori 5^r é melhor que r^2 às vezes e outras não. Também pode ser encontrada uma propriedade que garanta que ao escolher um corte fixo, o nível de significância do teste tende a 0 e o poder 1 e quando o tamanho da amostra tende a infinito.

REFERÊNCIAS

BRUNK, H. Univariate density estimation by orthogonal series. **Biometrika**, Oxford University Press, v. 65, n. 3, p. 521–528, 1978. Citado na página [23](#).

JONSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. [S.l.]: Prentice Hall Inc, 1992. Citado na página [22](#).

JR, F. J. M. The kolmogorov-smirnov test for goodness of fit. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 46, n. 253, p. 68–78, 1951. Citado na página [24](#).

PEREIRA, C. A. d. B.; STERN, J. M. Evidence and credibility: full bayesian significance test for precise hypotheses. **Entropy**, Molecular Diversity Preservation International, v. 1, n. 4, p. 99–110, 1999. Citado na página [19](#).

PEREIRA, C. A. d. B.; STERN, J. M.; WECHSLER, S. Can a significance test be genuinely bayesian? **Bayesian Analysis**, International Society for Bayesian Analysis, v. 3, n. 1, p. 79–100, 2008. Citado na página [22](#).

WHITTLE, P. On the smoothing of probability density functions. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 334–343, 1958. Citado na página [23](#).