ANOVA: Objectivos

- Verificar as condições de aplicabilidade de testes de comparação de médias;
- Utilizar ANOVA a um factor, a dois factores e mais de dois factores e interpretar interacções;
- Utilizar testes não paramétricos ANOVA em ordens para testar diferenças de medianas;
- Utilizar e interpretar os resultados de técnicas ANOVA multivariadas (MANOVA);
- Utilizar ANOVA de medições repetidas ou emparelhadas e ANOVA mistos;
- Realizar testes ANOVA emparelhados não paramétricos.

Testes (Não) Paramétricos

considera-se quase sempre Normal e variâncias homogéneas

- Os testes paramétricos consideram que a distribuição da variável é conhecida e as subpopulações com variâncias semelhantes;
- Os testes não paramétricos não impõem essa condição ainda que possam impor outras;
  - as variáveis nomeais podem ser reduzidas a binárias e logo as frequências seguem distribuições binomiais, pelo que os testes de proporções são considerados não paramétricos;
  - os testes não paramétricos para variáveis contínuas reduzem normalmente as variáveis a escalas ordinais;
  - os testes paramétricos impõem quase sempre uma distribuição Normal da população para as variáveis contínuas;
  - os testes paramétricos são normalmente mais potentes mas exigem a verificação do ajuste à Normal.

em amostras pequenas não é possível verificar a normalidade

conseguem valores de β menores para o mesmo valor de α

MBA Armando B. Mendes, DM, UAç.

Distribuição Normal ou Gaussiana

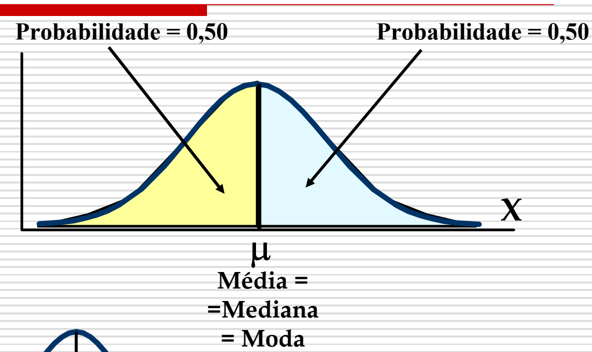
- A distribuição contínua mais comum, resultado da aplicação do Teorema do Limite Central:
  - utiliza dois parâmetros μ ou média e $\sigma^2 > 0$ ou variância e representa-se por $N(\mu,\sigma^2)$:
  - a função densidade de probabilidade é dada por: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$
  - a f.d.p. é simétrica em torno da média: $f(\mu - x) = f(\mu + x),\ \forall x > o$
  - como consequência da simetria é unimodal: Moda = Mediana = μ
  - as caudas são infinitas e assimptóticas relativamente ao eixo horizontal.

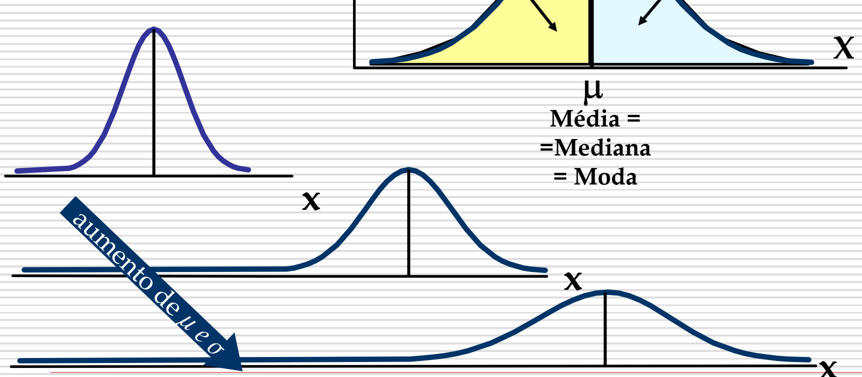
09-11-2006 Métodos Estatísticos

MBA Armando B. Mendes, DM, UAç.

Distribuição Normal

- simetria:
- aumento da variância e da média:





09-11-2006 Métodos Estatísticos

Verificação do Ajuste à Normal

- Teste não paramétrico de *Kolmogorov-Smirnov*:
  - teste mais potente e muito mais usado do que o teste de ajustamento do Qui-quadrado de Pearson;
  - apenas pode ser utilizado para testar o ajuste à Normal, pelo que a as hipóteses consideradas são:

    $H_0 : x \sim N(\mu,\sigma^2)$

    $H_1 : x \nsim N(\mu,\sigma^2)$

    semelhante aos gráficos de proporções normais

  - a estatística do teste é baseada na diferença máxima entre as frequências acumuladas nos valores da variável e as frequências da distribuição para os mesmos valores;

    na prática usa-se quase sempre porque normalmente não se conhecem os parâmetros da população

  - a correcção de Lilliefors deve ser usada quando os parâmetros da distribuição são estimados da amostra.

MBA Armando B. Mendes, DM, UAç.

Verificação do Ajuste à Normal

- Teste não paramétrico de *Shapiro-Wilk*:
  - apenas pode ser utilizado para testar o ajuste à Normal, pelo que a as hipóteses consideradas são:

    $H_0 : x \sim N(\mu,\sigma^2)$

    $H_1 : x \nsim N(\mu,\sigma^2)$

  - a estatística do teste é:

    $$W = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

    constantes conhecidas e calculadas segundo a distribuição

    valores pequenos indicam fraco ajuste à normal

  - este teste é mais potente do que o teste de *Kolmogorov-Smirnov* para amostras de dimensão inferior a 30.

    o SPSS produz resultados para este teste se n<51

09-11-2006 Métodos Estatísticos

Teste de Homogeneidade de Variâncias

- Teste de Levene

teste paramétrico pouco sensível à falta de Normalidade

- pode ser usado para duas ou mais amostras, pelo que as hipóteses a testar são:

  $H_0 : \sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2$

  $H_1 : \exists_{i,j}\ \sigma_i^2 \neq \sigma_j^2$

- assim, a estatística do teste é:

soma das dimensões das amostras para as várias variáveis

$$W = \frac{(N-k)\sum_{i=1}^{k}(\bar{Z}_i - \bar{Z})^2}{(k-1)\sum_{i=1}^{k}\sum_{i=1}^{n}(Z_{ij} - \bar{Z}_i)^2} \overset{a}{\sim} F(k-1, N\text{-k})$$

- existem versões para variáveis com distribuição aproximadamente normal (nesse caso Z é o módulo da variável original menos a média) e para não normais (usa-se a mediana) ou para a existência de observações atípicas (média aparada)

o SPSS produz resultados para média, mediana e média aparada

Igualdade de Médias de Populações Normais

Igualdade de Médias de Populações Normais

- Sejam $X_{1,1}, \ldots, X_{1,n1}$ e $X_{2,1}, \ldots, X_{2,n2}$ duas amostras aleatórias independentes de populações $N(\mu_1,\sigma_1^2)$ e $N(\mu_2, \sigma_2^2)$, respectivamente, então:
  - pelo que a estatística de teste para $\mu_1 = \mu_2$ quando as variâncias são conhecidas é:

$$Z_0 = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

  - no caso de variâncias desconhecidas mas consideradas iguais (homogéneas) a estatística do teste é:

$$t_0 = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim t_{(n_1+n_2-2)}$$

Igualdade de Médias de Populações Normais

- No caso mais geral em que as variâncias são desconhecidas e podem ser diferentes não é conhecida uma solução exacta.
  - se as amostras forem suficientemente grandes, pode-se usar uma variável fulcral com uma distribuição aproximada:

$$\frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \overset{a}{\sim} N(0,1)$$

usa-se com segurança se n_1 e n_2 >= 60

  - o erro da utilização desta expressão é tanto maior quanto menor a dimensão da amostra.

Igualdade de Médias de Amostras Emparelhadas

- Sejam $X_{1,1}, \dots, X_{1,n1}$ e $X_{2,1}, \dots, X_{2,n2}$ duas amostras aleatórias emparelhadas de populações $N(\mu_1,\sigma_1^2)$ e $N(\mu_2, \sigma_2^2)$, então para testar as hipóteses seguintes:

$H_0 : \mu_1 - \mu_2 = c$ (ou $\geq$ ou $\leq$)

$H_1 : \mu_1 - \mu_2 \neq c$ (ou $<$ ou $>$)

- a estatística do teste é dada em termos da variável diferença ($D_i = X_{1i} - X_{2i}$)

$$T = \frac{\overline{D} - \mu_D}{S_D / \sqrt{n}} \sim N(0,1)$$

Teste para uma mediana ou emparelhadas

- Teste de Wilcoxon:
  - alternativa ao teste de t-*student* para amostras pequenas, muito afastadas da Normal ou ordinais
  - não é imposta qualquer tipo de distribuição, mas seja qual for não se deve afastar muito da simetria
  - as hipóteses são em termos da mediana:

$H_0 : \theta = c$ *ou* $\theta_1 - \theta_2 = c$

$H_1 : \theta \neq c$ (ou $<$ ou $>$) ou $\theta_1 - \theta_2 \neq c$ (ou $<$ ou $>$)

neste caso é necessário construir uma variável com o valor *c* constante

  - a estatística do teste é baseada na variável ($D_i = X_i - k$ *ou* $X_{i1} - X_{i2}$) e na soma das ordens com sinal positivo (s^+) ou negativo (s^-) de D_i

$$T = \frac{\min(s^+, s^-) - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24 - \sum_{i=1}^{g} \frac{e_i^3 - e_i}{48}}} \sim N(0,1)$$

correcção usada quando existem vários grupos de empates (g)

Ficha de Trabalho 3

Pergunta a) e b) com dois factores, comparações binárias

Porquê a Análise de Variância?

- É incorrecto comparar médias de mais do que duas populações utilizando testes *t-student*, uma vez que o nível de significância deixa de medir a probabilidade de erro do tipo I;
- de facto verifica-se uma acumulação de erros para *k* pares que conduz a uma probabilidade de tipo I de $\approx 1-(1-\alpha)^k$ que é muito superior a α;
  - ex(1): usar α = 5% para 3 pares de comparações $\Rightarrow$ P (erro tipo I acumulada) ≈ 14,3%
  - ex(2): usar α = 5% para 5 pares de médias $\Rightarrow$ $\Rightarrow$ 10 comparações $\Rightarrow$ P (erro tipo I) ≈ 40%
- O método de análise de variância (ANOVA – - *ANalysis Of VAriance*) surgiu como resposta

a correcção de Bonferroni usa um $\alpha' \approx \alpha/k$ mais baixo para que o erro se aproxime do nível de signif. desejado

MBA Armando B. Mendes, DM, UAç.

Os Dados em Análise de Variância

- Um conjunto de factores ou tratamentos diferentes que correspondem a variáveis nominais (variáveis independentes);
  - 1 variável de factores: "one way" ANOVA
  - 2 variáveis de factores: "two way" ANOVA
  - 3 ou mais variáveis: ANOVA multifactorial
  - os factores podem ainda ser fixos ou aleatórios.
- Existe um (ANOVA) ou mais (MANOVA) variáveis quantitativas de medida da resposta aos factores (variáveis dependentes);
- O objectivo é testar o efeito que os diferentes tratamentos têm nas médias por factor das medidas de resposta.

os tratamentos resultam das combinações dos valores possíveis para os factores

nos factores aleatórios os níveis são escolhidos aleatoriam.

indica o sentido da causalidade

igual ao teste t mas para várias amostras

09-11-2006 Métodos Estatísticos

MBA Armando B. Mendes, DM, UAç.

Exemplo:

qual o melhor método de treino?

	Factor (método de treino)		
Níveis do Factor			
Indivíduos Escolhidos Aleatoria.			
Variável Dependente (resposta)	21 hrs	17 hrs	31 hrs
	27 hrs	25 hrs	28 hrs
	29 hrs	20 hrs	22 hrs

tempo de realização da tarefa após o treino

09-11-2006 Métodos Estatísticos

MBA

Armando B. Mendes, DM, UAç.

Modelo ANOVA "*one way*"

considerando observações amostrais

$$y_{it} = \bar{y} + (\bar{y}_t - \bar{y}) + (y_{it} - \bar{y}_t)$$

observações amostrais =

Média Global estimada como a média de todas as amostras em conjunto (ou média das médias)

\+ **efeito do tratamento estimado como a diferença entre a média do tratamento e a média global**

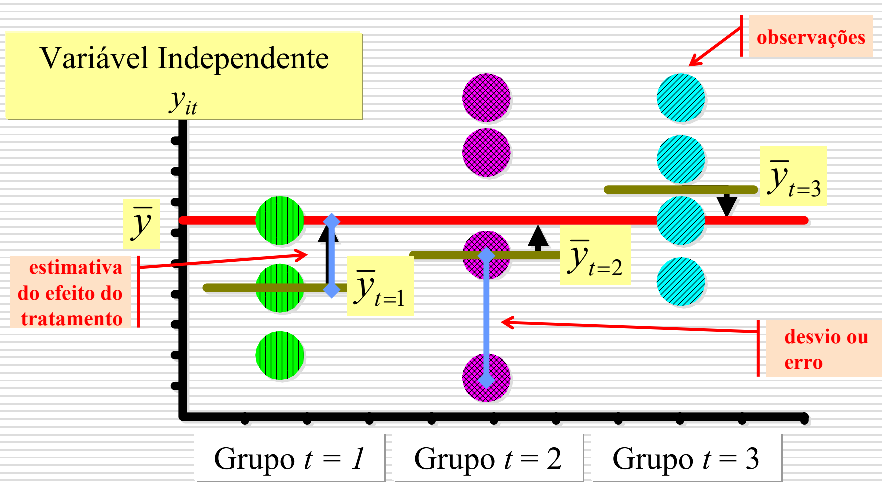
\+ **resíduos correspondentes à diferença entre as observações e a média por nível do factor**

09-11-2006 Métodos Estatísticos 20

MBA

Armando B. Mendes, DM, UAç.

Modelo ANOVA "*one way*"



09-11-2006 Métodos Estatísticos 21

MBA Armando B. Mendes, DM, UAç.

Hipóteses do ANOVA

- Pretende-se testar
  - se as médias populacionais são iguais;
  - se as amostras provêm da mesma população;
  - se os tratamentos têm algum efeito na variável dependente;
  - se o modelo considerado é válido;

expressões equivalentes

$$H_0 : \mu_{t=1} = \mu_{t=2} = \ldots = \mu_{t=k>2}$$

$$H_1 : \exists_{i,j} \; \mu_{t=i} \neq \mu_{t=j}$$

no caso de k=2 é preferível usar os testes t

pelo menos duas das médias são diferentes

- Não se testa:
  - se todas as médias são distintas entre si;
  - se uma média é superior ou inferior a outra.

apenas teste bilateral

09-11-2006 Métodos Estatísticos 22

MBA Armando B. Mendes, DM, UAç.

Decomposição da Variância

Soma dos Quadrados Total

TSS – *Total Sum of Squares*

Variância Total SQT

Soma dos Quadrados para os Factores

SST – *Sum of Square Treatment*

Soma dos Quadrados dos Erros

SSE – *Sum of Squares Error*

= **Variância devida aos Factores SQF** + **Variância devida à Aleatoriedade SQE**

- Variância entre os factores
- Variância entre as amostras
- Variância explicada pelo modelo ou tratamentos
- Variância entre os grupos

- Variância dentro das amostras ou grupos
- Variância do erro ou resíduos
- Variância não explicada pelo modelo \ tratamentos

09-11-2006 Métodos Estatísticos 23

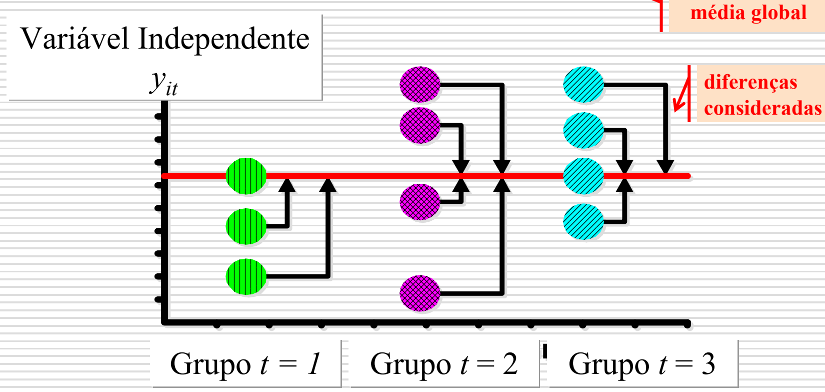
MBA Armando B. Mendes, DM, UAç.

Soma dos Quadrados Total

$$SQT = \sum_{t=1}^{k} \sum_{i=1}^{n_t} (y_{it} - \bar{y})^2$$

número de elementos do grupo com tratamento *t*

média global



09-11-2006 Métodos Estatísticos

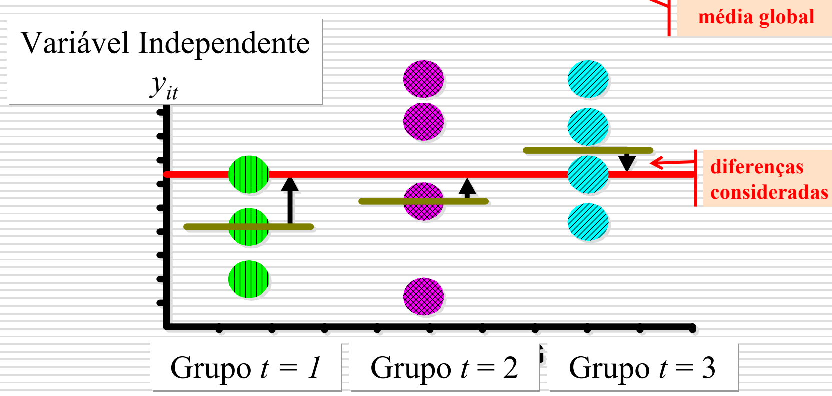
MBA Armando B. Mendes, DM, UAç.

Soma dos Quadrados dos Factores

$$SQF = \sum_{t=1}^{k} n_t (\bar{y}_t - \bar{y})^2$$

número de elementos do grupo com tratamento *t*

média global



09-11-2006 Métodos Estatísticos

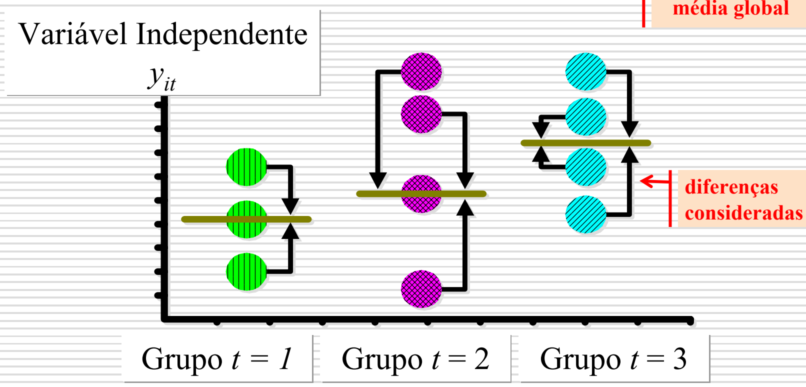
MBA Armando B. Mendes, DM, UAç.

Soma dos Quadrados do Erro

$$SQE = \sum_{t=1}^{k} \sum_{i=1}^{n_t} (y_{it} - \bar{y}_t)^2$$

número de elementos do grupo com tratamento *t*

média global



09-11-2006 Métodos Estatísticos

MBA Armando B. Mendes, DM, UAç.

Teste F de Snedecor (Fisher)

- Estatística do teste:

$$F = \frac{SQF/(k-1)}{SQE/(N-k)} \sim F_{(k-1,N-k)}$$

QMF - Quadrado Médio dos Factores

graus de liberdade

QME - Quadrado Médio dos Erros

número de indivíduos no total de todas as amostras

- Tabela ANOVA *one-way*

fonte de variação	soma de quadrados	graus de lib.	quadrados médios	estatística do teste
Factores (entre grupos)	SQF	k-1	SQF/(k-1)	QMF/QME
resíduos (dentro grupos)	SQE	N-k	SQE/(N-1)	
total	SQT	N-1		

forma como o SPSS arruma os cálculos

09-11-2006 Métodos Estatísticos

Pressupostos do Teste F

- As amostras são aleatórias para cada população (ou tratamento);
- As amostras são independentes;
- As populações seguem uma distribuição Normal:
  - testar o ajuste à normal para cada tratamento;
  - o teste mostra alguma robustez perante pequenos afastamentos da normal.
- As populações são homogéneas:
  - é necessário testar a igualdade de variâncias para cada tratamento;
  - maior robustez a este pressuposto se as amostras tiverem igual dimensão.

Testes *Post-Hoc a posteriori*

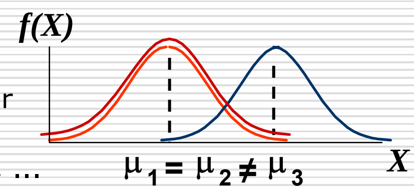
daí a denominação de testes *a posteriori*

- Quando o teste ANOVA rejeita a hipótese nula de igualdade de médias, os testes *Post-Hoc* permitem identificar quais as médias distintas;
- As hipóteses para todos os pares de tratamentos (*i*, *j*) são:

$$H_0 : \mu_{t=i} = \mu_{t=j}$$
$$H_1 : \mu_{t=i} \neq \mu_{t=j}$$

comparação múltipla de médias

  - O SPSS dispõe de uma grande variedade de testes deste tipo;
  - nenhum é claramente melhor do que os restantes;
  - ex: Turkey, Bonferroni, LSD, ...



Teste *Post-Hoc* de Tukey-Kramer

- Considerado um dos mais robustos a desvios à Normalidade e homogeneidade;
- Recomendado para amostras de grande dimensão;
- A estatística do teste envolve diferenças de médias:

$$Q = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{QME/2}(1/n_i + 1/n_j)} \sim DistribuiçãoTabelada$$

- Em caso de resultados contraditórios entre o ANOVA e os testes *Post-Hoc*, deve-se dar preferência ao ANOVA por ser mais potente.

Testes *Post-Hoc* para amostras menores

- O teste de Bonferroni é o mais recomendado para amostras pequenas

$$Q = \frac{\overline{X}_i - \overline{X}_j}{\sqrt{QME}(1/n_i + 1/n_j)} \sim F_{(1,N-k)}$$

usa um nível de significância modificado para compensar a acumulação de erros

- Outros testes são igualmente recomendados para amostras pequenas
  - *Least Significance Difference* (LSD): a potência do teste decresce fortemente com o número de comparações;
  - Teste de Scheffé: pode ser usado para qualquer número de comparações mas é demasiado restrito, *i.e.* é difícil rejeitar H_0

ANOVA de Kruskal-Wallis

- Também conhecido por ANOVA em ordens;
- Teste não paramétrico de comparação de *k* medianas que pode ser usado para variáveis dependentes ordinais;

  não impõe uma distribuição aos dados

- Pode ser entendido como uma extensão do teste de Wilcoxon-Mann-Whitney que é igual quando *k=2*;
- Apenas quando existem factores fixos *i.e.* não admite factores aleatórios;
- A estatística do teste segue uma distribuição do $\chi^2_{(k-1)}$ se cada nível do factor tiver pelo menos 5 observações.

ANOVA de Kruskal-Wallis

- Pressupostos:
  - As amostras são independentes
  - A variável dependente é pelo menos ordinal
  - As populações têm variâncias homogéneas;
  - As populações têm distribuições com forma semelhante;

  ao contrário dos testes ANOVA o de Kruskal-Wallis é robusto para estas duas condições.

- É possível usar este teste para ANOVA factorial mas o SPSS não tem o procedimento implementado.