Relatório Hugo Cobra (INTO 2016)

Felipe Figueiredo

June 21, 2016

Table of Contents

# Metodologia utilizada

A análise descritiva foi apresentada na forma de tabelas os dados observados, expressos pela média, desvio padrão, mediana, amplitude interquartílica (AIQ), mínimo e máximo para dados numéricos (quantitativos) e pela frequência (n) e percentual (%) para dados categóricos (qualitativos). Ao longo do texto, as variáveis normais foram sumarizadas como média +/- desvio padrão (DP), e as variáveis não normais como mediana +/- AIQ.

Foi aplicado o teste de normalidade de Shapiro-Wilks nas variáveis numéricas. A comparação das variáveis basais e finais foi feita com o teste t para variáveis normalmente distribuídas, e com o teste de Mann-Whitney para variáveis que tiveram a hipótese de normalidade rejeitada. Para variáveis normais, foi aplicada a correlação de Pearson, e para variáveis não normais, a correlação não paramétrica de Spearman.

A fim de investigar possíveis fatores de risco para o desfecho, foram ajustados modelos de regressão logística, tendo o desfecho como variável resposta. A significância global de cada um destes modelos foi avaliada pelo teste de Wald.

Em todas as análises, foi adotada a significância de 5%. As análises estatísticas foram feitas utilizando o software estatístico R, versão 3.3.0 (<https://www.R-project.org>).

# Resultados

## Tabelas descritivas

### Dados numéricos

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Média | DP | Mediana | AIQ |
| **Idade** | 66.19 | 9.25 | 67 | 12 |
| **IMC** | 30.05 | 5.45 | 29.68 | 7 |
| **Albumina** | 3.63 | 0.66 | 3.6 | 0.75 |
| **Globulina** | 2.67 | 0.86 | 2.73 | 1.08 |
| **Tempo** | 88.69 | 20.9 | 87 | 27 |

A média (DP) das idades é 66.19 (9.25)

### Dados categóricos

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ATB | Sexo | Diagnóstico | DM | ASA | Desfecho |
| N:156 | F:231 | AR : 50 | N:224 | I : 16 | N:279 |
| S:128 | M: 53 | OA :230 | S: 60 | II :260 | S: 5 |
| NA | NA | ON : 2 | NA | III : 6 | NA |
| NA | NA | Outros: 2 | NA | NA's: 2 | NA |

## Tabelas de Contingência

As tabelas de contingência a seguir, associam duas variáveis categóricas. Nas estão representadas as categorias do Desfecho (S = infeccção e N = não infeccão). OS p-valores correspondem ao teste de Fisher.

### Sexo por tratamento

|  |  |  |
| --- | --- | --- |
|  | N | S |
| **F** | 46.1 | 35.2 |
| **M** | 8.8 | 9.9 |

P-valor: 0.2235331

### Diabete por tratamento

|  |  |  |
| --- | --- | --- |
|  | N | S |
| **N** | 43 | 35.9 |
| **S** | 12 | 9.2 |

P-valor: 0.7725722

### Diagnostico por tratamento

|  |  |  |
| --- | --- | --- |
|  | N | S |
| **AR** | 8.8 | 8.8 |
| **OA** | 45.4 | 35.6 |
| **ON** | 0 | 0.7 |
| **Outros** | 0.7 | 0 |

P-valor: 0.210475

### ASA por Tratamento

|  |  |  |
| --- | --- | --- |
|  | N | S |
| **I** | 2.8 | 2.8 |
| **II** | 50.4 | 41.8 |
| **III** | 1.8 | 0.4 |

P-valor: 0.3976111

### Desfecho por Tratamento

|  |  |  |
| --- | --- | --- |
|  | N | S |
| **N** | 53.9 | 1.1 |
| **S** | 44.4 | 0.7 |

P-valor: 1

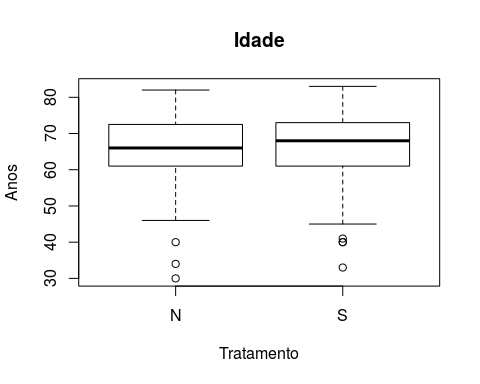
### Desfecho por ASA

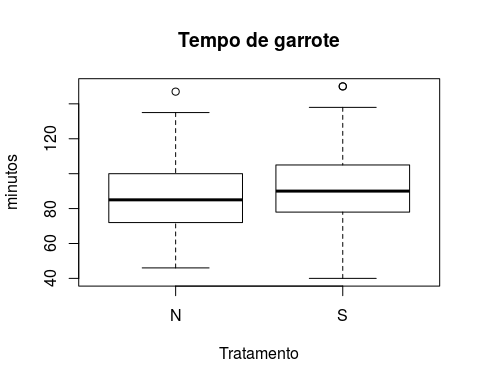
|  |  |  |
| --- | --- | --- |
|  | N | S |
| **I** | 5.7 | 0 |
| **II** | 90.8 | 1.4 |
| **III** | 2.1 | 0 |

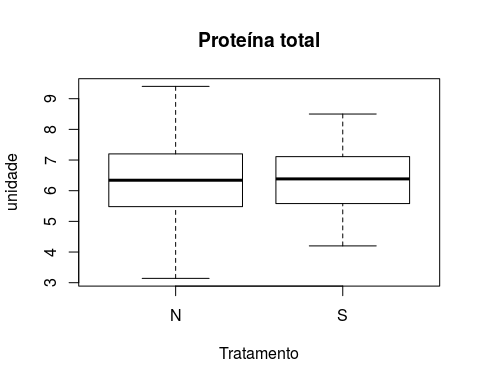
P-valor: 1

## Gráficos

### Dados numéricos

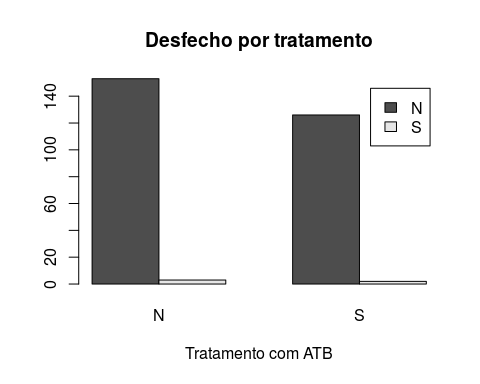




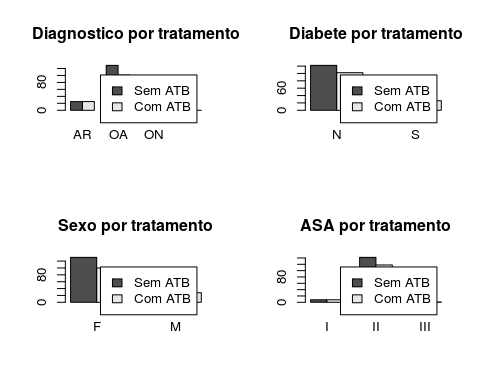


### Dados categóricos

#### Desfecho por tratamento



#### Outros dados, por tratamento



## Análise de Regressão

### Modelos de uma variável

Foi feita uma análise de regressão logística, tomando como variável resposta o Desfecho (binário). Cada modelo de regressão logística simples considerou como variável preditora cada um dos seguintes dados:

* Tratamento
* Diabete
* Sexo
* Idade
* IMC
* Proteína total
* Diagnóstico

A **única** análise que retornou diferença significativa foi levando em conta o sexo.

**~~Nenhum~~** ~~dos modelos ajustados foi considerado significativo ao nível de 5% para estes dados.~~ Por conta disso, apenas o resultado bruto de cada teste está sendo apresentado, sem formatação. A regressão logística foi avaliada pelo teste de Wald.

O log da razão de chances de se selecionar um homem ao invés de uma mulher na amostra é: 1.9271641. Usando a exponencial para cancelar o lograritmo, podemos interpretar este resultado como: os homens tem uma chance 6.87 maior que as mulheres de ter o desfecho [infecção na cirurgia, etc].

**Sugestão de redação**: *Dos fatores de risco analisados neste estudo, o único identificado como significativo foi o sexo. Nesta amostra, os homens têm uma chance 6.87 maior que as mulheres de desenvolver a condição...*

A seguir, as tabelas de resultados das regressões logísticas analisadas.

##   
## Call:  
## glm(formula = Desfecho ~ ATB, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.1971 -0.1971 -0.1971 -0.1775 2.8840   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.9318 0.5830 -6.744 1.54e-11 \*\*\*  
## ATBS -0.2113 0.9208 -0.229 0.818   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 50.307 on 283 degrees of freedom  
## Residual deviance: 50.254 on 282 degrees of freedom  
## AIC: 54.254  
##   
## Number of Fisher Scoring iterations: 7

##   
## Call:  
## glm(formula = Desfecho ~ DM, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.2125 -0.2125 -0.2125 -0.2125 2.7576   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.7796 0.4523 -8.357 <2e-16 \*\*\*  
## DMS -16.7864 2288.9810 -0.007 0.994   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 50.307 on 283 degrees of freedom  
## Residual deviance: 47.910 on 282 degrees of freedom  
## AIC: 51.91  
##   
## Number of Fisher Scoring iterations: 19

##   
## Call:  
## glm(formula = Desfecho ~ Sexo, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.3414 -0.1319 -0.1319 -0.1319 3.0820   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.7406 0.7102 -6.675 2.47e-11 \*\*\*  
## SexoM 1.9272 0.9261 2.081 0.0374 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 50.307 on 283 degrees of freedom  
## Residual deviance: 46.037 on 282 degrees of freedom  
## AIC: 50.037  
##   
## Number of Fisher Scoring iterations: 7

##   
## Call:  
## glm(formula = Desfecho ~ Idade, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.3160 -0.1998 -0.1827 -0.1671 2.9757   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -2.06966 2.78453 -0.743 0.457  
## Idade -0.03007 0.04329 -0.695 0.487  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 50.307 on 283 degrees of freedom  
## Residual deviance: 49.861 on 282 degrees of freedom  
## AIC: 53.861  
##   
## Number of Fisher Scoring iterations: 7

##   
## Call:  
## glm(formula = Desfecho ~ IMC, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.2587 -0.2010 -0.1863 -0.1724 2.9647   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.31048 2.55216 -2.081 0.0375 \*  
## IMC 0.04298 0.08024 0.536 0.5922   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 50.018 on 275 degrees of freedom  
## Residual deviance: 49.737 on 274 degrees of freedom  
## (8 observations deleted due to missingness)  
## AIC: 53.737  
##   
## Number of Fisher Scoring iterations: 7

##   
## Call:  
## glm(formula = Desfecho ~ ptn, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.26855 -0.14210 -0.11422 -0.09179 3.14577   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -1.6253 3.7617 -0.432 0.666  
## ptn -0.5348 0.6530 -0.819 0.413  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 23.297 on 249 degrees of freedom  
## Residual deviance: 22.609 on 248 degrees of freedom  
## (34 observations deleted due to missingness)  
## AIC: 26.609  
##   
## Number of Fisher Scoring iterations: 8

##   
## Call:  
## glm(formula = Desfecho ~ Diagnóstico, family = binomial(link = "logit"))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -0.2857 -0.1620 -0.1620 -0.1620 2.9460   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -3.1781 0.7217 -4.404 1.06e-05 \*\*\*  
## DiagnósticoOA -1.1483 0.9266 -1.239 0.215   
## DiagnósticoON -14.3880 2797.4420 -0.005 0.996   
## DiagnósticoOutros -14.3880 2797.4420 -0.005 0.996   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 50.307 on 283 degrees of freedom  
## Residual deviance: 48.792 on 280 degrees of freedom  
## AIC: 56.792  
##   
## Number of Fisher Scoring iterations: 16

### Modelo com todas as variáveis

Por fim, foi criado também um modelo logístico para o desfecho tendo como preditores **todas** as outras variáveis do estudo (resultado não mostrado). Neste modelo "saturado", as combinações não resultaram em significância estatística das chances de uma categoria para outra.

Sugestão: este resultado pode ser omitido da redação sem perda de generalidade, e foi incluído apenas para compleição da análise.