

SME0803 Análise Exploratória de Dados – Associação entre variáveis - Data 25/05/2012
Aula Prática 3 – Entregar Exercício 2 feito em SPSS dia 04/06/2012

Exercício 1: Suspeita-se que a idade está relacionada ao tempo de reação a um certo estímulo (Dados acuidade, disponíveis na CoteiaWiki). Utilize os comandos a seguir para avaliar a associação entre essas variáveis. Interprete os resultados.

```
> Acuidade <- read.table("http://wiki.icmc.usp.br/images/0/0f/Acuidade.txt", header=TRUE)
> X <- Acuidade$idade
> Y <- Acuidade$tempo
> plot(X, Y, pch=16)
> summary(lm(Y~X))
> abline(lm(Y~X))
```

Exercício 2: O conjunto de dados Banco.csv (Banco.sav) está disponível na Coteia Wiki. Esse banco de dados possui informações de clientes como Identificação, Sexo (0: M, 1: F), Idade (em anos), Cartão de crédito (0: não possui, 1: possui), Limite de cheque especial (0: não possui, 1: possui), Inadimplência (0: não, 1: sim) e Saldo devedor (em reais).

Utilize medidas descritivas e de associação para responder às perguntas:

- (a) O fato de possuir cartão de crédito está relacionado com a inadimplência?
- (b) O fato de possuir cheque especial está relacionado com a inadimplência?
- (c) A inadimplência está relacionada com o sexo dos clientes?
- (d) A inadimplência está relacionada com a idade?
- (e) Forneça análises complementares se achar necessário.

Utilize comandos do tipo

```
> dadosBanco<- read.csv("http://www.icmc.usp.br/~cibele/Dados/Banco.csv ", header=TRUE, sep=";")
> attach(dadosBanco)
> names(dadosBanco)
> boxplot(SaldoDevedor ~ Sexo)
> freq <- table(ChequeEspecial, Inadimplente)
> chisq.test(freq)
```

Exercício 3: Considere o conjunto de dados Câncer.csv (Câncer.sav) disponível na Coteia Wiki. Esse banco de dados possui informações de pacientes como Identificação, Sexo (0: M, 1: F), Idade (em anos), Doenças psiquiátricas (0: não possui, 1: possui), Pais com câncer (0: não possui, 1: possui), Prática atividades físicas (0: não, 1: sim), Dieta rica em gordura e Desenvolvimento de câncer (0: não, 1: sim)

Utilize medidas descritivas e de associação para identificar a associação de variáveis com o desenvolvimento de câncer.

```
> dadosCancer<- read.csv("http://www.icmc.usp.br/~cibele/Dados/Cancer.csv ", header=TRUE, sep=";")
> attach(dadosCancer)
> names(dadosCancer)
> boxplot(Idade ~ DesenvolveuCancer)
> freq <- table(DesenvolveuCancer, Sexo)
> chisq.test(freq)
```