

Exercício 3

Descrição

No arquivo **imoveis.dat** são apresentados dados relativos a uma amostra de 27 imóveis. Na ordem são apresentados os valores das seguintes variáveis:

- imposto do imóvel (em 100 USD)
- área do terreno (em 1000 pés quadrados)
- área construída (em 1000 pés quadrados)
- idade da residência (em anos)
- preço de venda do imóvel (em 1000 USD)

Enunciado

Ajuste um modelo normal linear do preço de venda contra as demais variáveis explicativas. Use o método AIC para selecionar as variáveis explicativas. Faça uma análise de diagnóstico com o modelo selecionado. Interprete os coeficientes estimados. Seja $y(z)$ o valor do preço de venda de um imóvel que não está na amostra com os valores das variáveis explicativas do modelo final representados por z . Como fica a estimativa intervalar de coeficiente $(1 - \alpha)$, $0 < \alpha < 1$, para $y(z)$? Alguma restrição para os valores de z ?

Leitura dos dados

Para ler os dados no R fazemos:

```
imoveis <- data.frame(  
  scan("dados/imoveis.dat", list(imposto=0, areat=0, areac=0, idade=0, preco=0)))
```

imposto	areat	areac	idade	preco
4.918	3.472	0.998	42	25.9
5.021	3.531	1.500	62	29.5
4.543	2.275	1.175	40	27.9
4.557	4.050	1.232	54	25.9
5.060	4.455	1.121	42	29.9
3.891	4.455	0.988	56	29.9

Análise descritiva

```
df <- gather(imoveis, key = var, value = value, -preco)  
ggplot(df, aes(x=value, y=preco)) + geom_point() + stat_smooth(method = "lm", se = F) +  
  facet_wrap(~var, scales = "free") + xlab("Valor da variável") + ylab("Preço")
```

