

**UNIVERSIDADE FEDERAL DO PARANÁ
DEPARTAMENTO DE ESTATÍSTICA**

AGATHA DE MELO

GRR: 20137524

ELIAS SANTIAGO DINIZ

GRR: 20135636

NATHALIE DO AMARAL PORTO MARTINS

GRR: 20137583

RELATÓRIO DE ANÁLISE DE DADOS DA LIGA PROFISSIONAL DE GOLFE

Base de dados – Professional Golf Association

**CURITIBA
2017**

AGATHA DE MELO
GRR: 20137524
ELIAS SANTIAGO DINIZ
GRR: 20135636
NATHALIE DO AMARAL PORTO MARTINS
GRR: 20137583

RELATÓRIO DE ANÁLISE DE DADOS DA LIGA PROFISSIONAL DE GOLFE

Base de dados – Professional Golf Association

Relatório apresentado à Disciplina de Modelos Lineares Generalizados da Universidade Federal do Paraná, como requisito parcial para aprovação na disciplina.

Professor: Dr. Cesar Augusto Taconeli

CURITIBA
2017

SUMÁRIO

| | |
|---------------------------------------|-----------|
| 1 INTRODUÇÃO | 5 |
| 2 MATERIAL E MÉTODOS | 6 |
| 2.1 TAMANHO DA AMOSTRA | 6 |
| 2.2 MÉTODO ESTATÍSTICO | 6 |
| 3 RESULTADOS E DISCUSSÃO | 7 |
| 3.1 PRECISÃO E SEXO | 7 |
| 3.2 ANÁLISE DE RESÍDUOS | 7 |
| 3.3 RELAÇÃO ENTRE AS VARIÁVEIS | 9 |
| 4 CONCLUSÕES | 10 |

RESUMO

Golfe é um dos poucos esportes com bola que não exige uma área de jogo normalizada. É praticado num campo, o qual geralmente consiste numa progressão de nove ou dezoito buracos. A Ladies Professional Golf Association (LPGA) e o Professional Golfers Association (PGA) são entidades profissionais que reúnem os jogadores de golfe em países como o Reino Unido e Estados Unidos. Da base de dados da PGA/LPGA, foi obtida uma amostra composta de 353 observações, divididas em três variáveis: Variável resposta – Distância da condução da bola (m); Variável x_1 – Precisão do Fairway (campo) em porcentagem; Variável x_2 – Sexo, sendo 1 – feminino e 2 – masculino. O método estatístico utilizado foi o de Regressão Linear. Fixando-se a variável sexo, observou-se que a condução de bola diminuiu em média 0,083107 m. Uma observação foi identificada como possível ponto de influência, mas após nova análise, optou-se por não retirá-la do modelo.

Palavras-chave: Regressão Linear. Condução. Precisão.

1 INTRODUÇÃO

Golfe, criado na Escócia (século XV), é um esporte no qual os jogadores usam diversos tipos de tacos para arremessar uma bola para uma série de buracos numa vasta extensão de terreno, usando o menor número possível de tacadas.

É um dos poucos esportes com bola que não exige uma área de jogo normalizada. Em vez disso, é praticado num campo, o qual geralmente consiste numa progressão de nove ou dezoito buracos. Cada buraco inclui uma área de terreno inicial (*tee*) e uma área final (*green*), na qual se encontra o buraco propriamente dito. Entre as duas áreas existem diversos tipos padronizados de terreno e obstáculos, e cada buraco possui uma configuração única.

Do inglês, a *Ladies Professional Golf Association* (LPGA) é uma organização norte-americana profissional para golfistas do sexo feminino. Além disso, o *Professional Golfers Association* (PGA) são entidades profissionais que reúnem os jogadores de golfe em países como o Reino Unido e Estados Unidos.

O objetivo deste trabalho foi analisar os dados de PGA e LPGA para verificar se a condução de bola do jogador está relacionada com a precisão da tacada, e o com o sexo do jogador.

2 MATERIAL E MÉTODOS

2.1 TAMANHO DA AMOSTRA

A amostra é composta de 353 observações, divididas em três variáveis:

- Variável resposta – Distância da condução da bola (m);
- Variável x_1 – Precisão do Fairway (campo) em porcentagem;
- Variável x_2 – Sexo, sendo 1 – feminino e 2 – masculino.

Os dados foram fornecidos pela Ladies Professional Golf Association (LPGA) e pelo Professional Golfers Association (PGA), em 2008.

2.2 MÉTODO ESTATÍSTICO

Para análise, utilizou-se o *software* estatístico R[®], e o método estatístico utilizado foi o de Regressão Linear.

3 RESULTADOS E DISCUSSÃO

3.1 PRECISÃO E SEXO

Após análise estatística, foi observado que as variáveis precisão e sexo influenciam na condução da bola (Tabela 1).

Tabela 1. ANOVA da influência das variáveis precisão e sexo sobre a condução da bola. Diferença estatística para $\alpha = 0,05$.

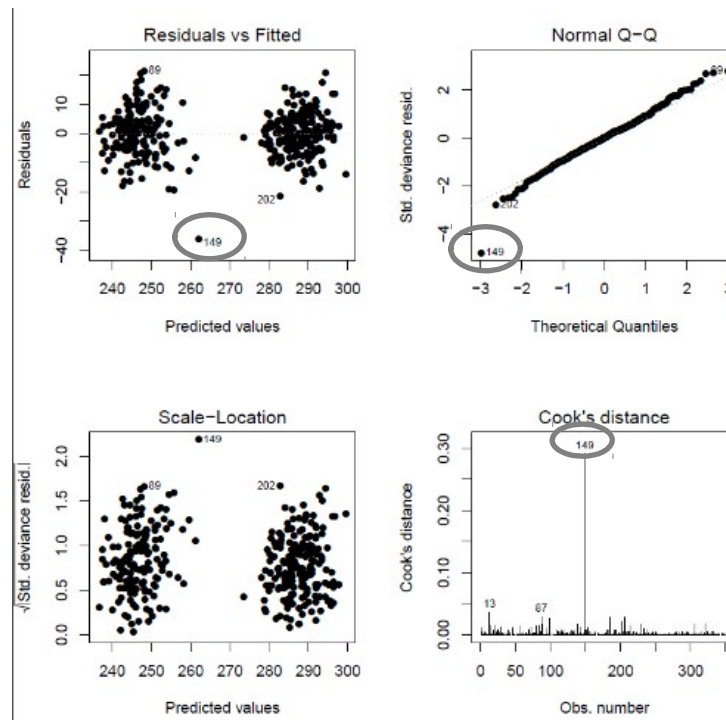
| | Estimativa | Desvio Padrão | Valor <i>t</i> | <i>p</i> Valor |
|-------------------|-------------------|----------------------|-----------------------|-----------------------|
| Intercepto | 303,00020 | 4,99458 | 60,67 | $< 2 \cdot 10^{-16}$ |
| Precisão | -0,83107 | 0,07332 | -11,33 | $< 2 \cdot 10^{-16}$ |
| Sexo | 37,27097 | 0,88191 | 42,26 | $< 2 \cdot 10^{-16}$ |

Fixando-se a variável sexo, pode-se notar que quando a precisão é aumentada em 1%, a condução de bola diminui em média 0,083107 m. Além disso, quando fixada a precisão, e mudando a variável sexo de feminino para masculino, a condução de bola aumenta em 37,27 m.

3.2 ANÁLISE DE RESÍDUOS

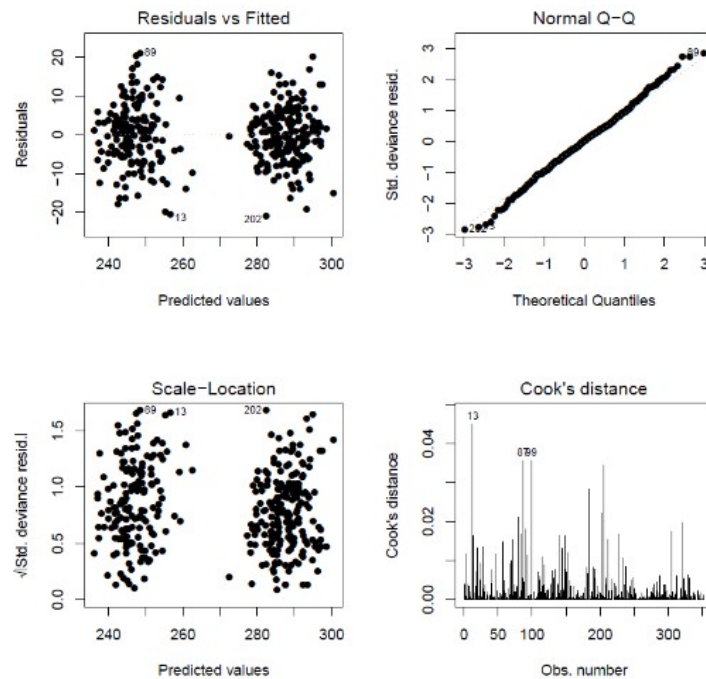
Através da análise resíduos dos dados originais, foi observado que o ponto 149 é um possível ponto de influência (Figura 1)

Figura 1. Análise de resíduos dos dados originais.



Após a remoção deste ponto, foi observado que não houve grande diferença no modelo (Figura 2).

Figura 2. Análise com remoção da observação 149.



Com isso, foi gerado um novo modelo, com a seguinte ANOVA (Tabela 2).

ANOVA da influência das variáveis precisão e sexo sobre a condução da bola.
Diferença estatística para $\alpha = 0,05$.

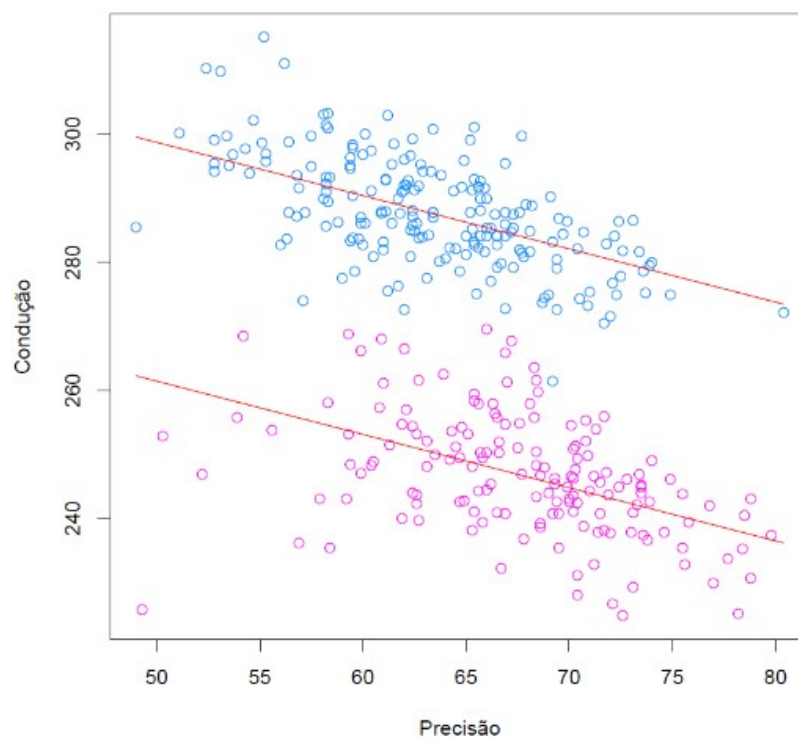
| | Estimativa | Desvio Padrão | Valor t | p Valor |
|-------------------|-------------------|----------------------|-----------------------------|-----------------------------|
| Intercepto | 307,45554 | 4,91792 | 62,52 | $< 2 \cdot 10^{-16}$ |
| Precisão | -0,89341 | 0,07209 | -12,39 | $< 2 \cdot 10^{-16}$ |
| Sexo | 36,76617 | 0,85978 | 42,76 | $< 2 \cdot 10^{-16}$ |

Por falta de evidências, não foi considerada a observação 149 como *outlier* e optou-se por mantê-la no modelo.

3.3 RELAÇÃO ENTRE AS VARIÁVEIS

Após análise, verificou-se que a precisão dos homens será sempre maior do que a das mulheres, independentemente da posição (Figura 3). Também observou-se, que quanto maior a condução, menor é a precisão.

Figura 3. Precisão de homens e mulheres.



Azul – Homens. Rosa – Mulheres

4 CONCLUSÕES

Fixando-se a variável sexo, pode-se notar que quando a precisão é aumentada em 1%, a condução de bola diminui em média 0,083107 m. Além disso, quando fixada a precisão, e mudando a variável sexo de feminino para masculino, a condução de bola aumenta em 37,27 m.

Por falta de evidências, não foi considerada a observação 149 como *outlier* e optou-se por mantê-la no modelo.

A precisão dos homens será sempre maior do que a das mulheres, independentemente da posição