



Ex. 06 - Regressão Polinomial

Exercício 01. Planeje um experimento no delineamento inteiramente casualizado com um fator quantitativo.

Exercício 02. Para um conjunto de dados obtidos de um experimento instalado no delineamento inteiramente casualizado com um fator quantitativo, faça a análise dos dados e interprete os resultados.

EXERCÍCIO 01

- **Áreas de atuação:**

1. Genética Quantitativa;
2. Melhoramento de Plantas;
3. Hortaliças;

- **Título do experimento:**

- Uso de radiação gama na obtenção de mutantes de alho

- **Hipóteses testadas:**

- H0: O desenvolvimento vegetativo de plantas de alho não é afetado pela mutagênese;
- Ha: O desenvolvimento vegetativo de plantas de alho são afetados em, pelo menos, um dos tratamentos mutantes;

- **Objetivos:**

- Verificar o potencial uso de mutagênese por radiação gama sobre dentes de alho na obtenção de novas cultivares.

- **Fatores e níveis:**

- Intensidade de radiação gama (Gy);
- Níveis do tratamento:
 - a. 2,5 Gy
 - b. 5,0 Gy
 - c. 7,5 Gy
 - d. 10,0 Gy
 - e. 12,5 Gy
 - f. 15,0 Gy
 - g. Testemunha (0,0 Gy)

- **Variável resposta:**

- Altura de plantas aos 42 dias após o plantio;

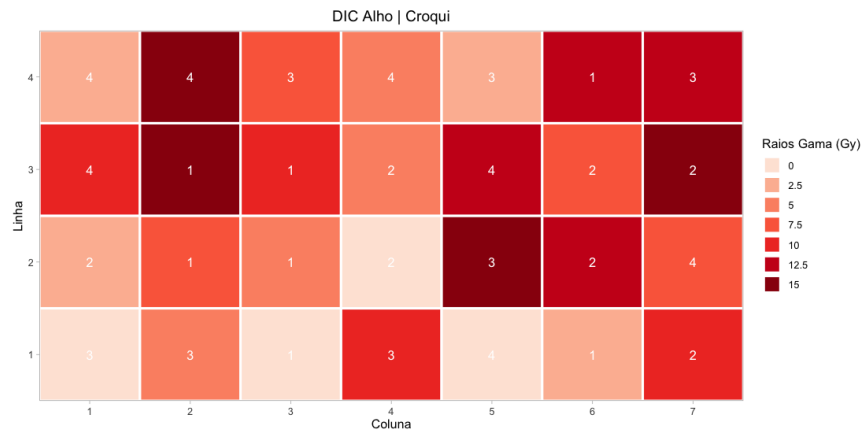
- **Design Experimental:**

```
# Plota o croqui da área
croquiDIC <- ggplot(SampleDIC, aes(x = LINHA, y = COLUNA, fill = DOSE)) +
  geom_tile(color = "white", lwd = 1) +
  geom_text(aes(label = REPETIÇÃO), color = "white", size = 4) +
  scale_color_continuous() +
  scale_x_continuous(breaks = unique(SampleDIC$LINHA), labels = unique(SampleDIC$LINHA),
                    expand = c(0, 0)) +
  scale_y_continuous(breaks = unique(SampleDIC$COLUNA), labels = unique(SampleDIC$COLUNA),
                    expand = c(0, 0)) +
  labs(
    x = "Coluna",
    y = "Linha",
    title = "DIC Alho | Croqui",
    fill = "Raios Gama (Gy)" +
  theme_light() +
```

```

theme(
  axis.text.x = element_text(angle = 0, vjust = 0.5, hjust = 0.5),
  axis.text.y = element_text(angle = 0, vjust = 0.5, hjust = 0.5),
  panel.grid = element_blank(),
  plot.title = element_text(hjust = 0.5)
)
print(croquiDIC)

```



A coloração foi realizada em função da exposição à radiação gama em dentes de alho da variedade Ito.

EXERCÍCIO 02

Para esse exercício, será utilizada apenas a variável **altura de plantas**.

▼ Análise exploratória:

1. Gráfico de pontos:

```

ggplot(DICgarlic, aes(x = DOSE, y = ALTURA)) +
  geom_point() +
  expand_limits(y = 0) +
  labs(
    x = "Dose (Gy)",
    y = "Altura (cm)",
    title = "DIC Alho | Diagrama de Dispersão") +
  theme(
    plot.title = element_text(hjust = 0.5)
  )

```

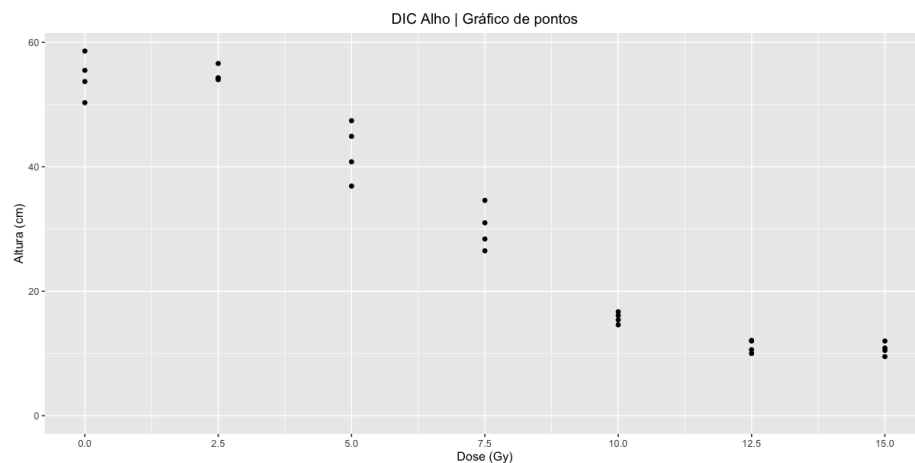


Gráfico de pontos para dados de altura de plantas de alho, aos 42 DAP, sob diferentes doses de irradiação gama.

2. Gráfico BoxPlot:

```
ggplot(DICgarlic, aes(x = DOSE, y = ALTURA)) +  
  geom_boxplot() +  
  expand_limits(y = 0) +  
  labs(  
    x = "Dose (Gy)",  
    y = "Altura (cm)",  
    title = "DIC Alho | BoxPlot") +  
  theme(  
    plot.title = element_text(hjust = 0.5))
```

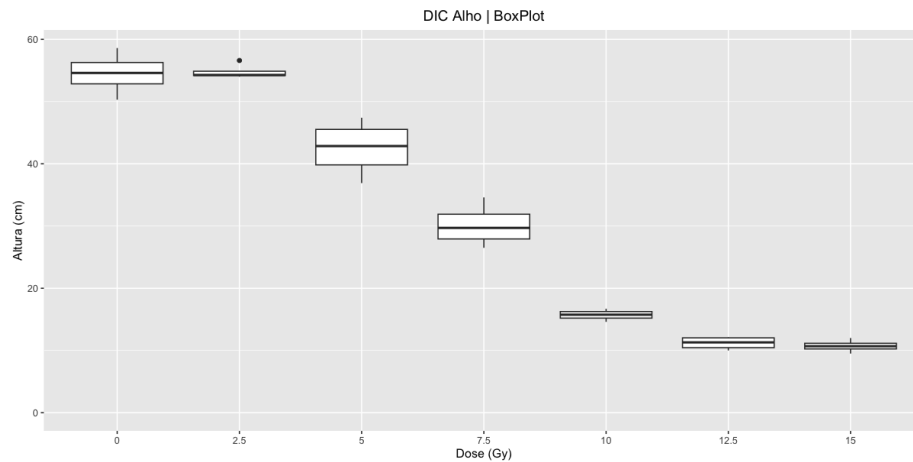


Gráfico BoxPlot para dados de altura de plantas de alho, aos 42 DAP, sob diferentes doses de irradiação gama.

▼ Validação das pressuposições da ANOVA:

- **Levene** | Teste de homogeneidade de variâncias

```
with(DICgarlic,  
  levene.test(ALTURA, DOSE, location = "mean"))
```

Resultado:

```
data: ALTURA  
Test Statistic = 3.742, p-value = 0.0109
```

Portanto → De acordo com o teste de Levene a 5% de probabilidade de erro, **as variâncias não podem ser consideradas homogêneas**.

- **Shapiro-Wilk** | Teste de normalidade dos resíduos

```
# Modelo linear  
lmDIC = lm(ALTURA~DOSE, DICgarlic)  
resDIC <- residuals(lmDIC)          # Resíduos  
resStudDIC <- rstandard(lmDIC)      # Resíduos studentizados  
  
shapiro.test(resStudDIC)
```

Resultado:

```
data: resStudDIC  
W = 0.95789, p-value = 0.3104
```

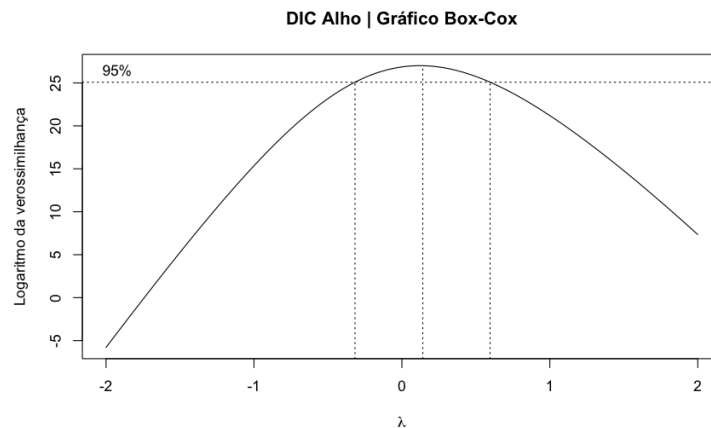
Portanto → De acordo com o teste de Shapiro-Wilk a 5% de probabilidade de erro, **os resíduos podem ser considerados normais**.



Visto que o conjunto de dados não atende as pressuposições da ANOVA, deverá ser realizada uma transformação de dados. Para isso, será utilizada a análise do gráfico Box-Cox para a obtenção de um lambda aproximado que defina a transformação adequada.

- **Gráfico Box-Cox**

```
with(DICgarlic, {
  plot_boxcox <- boxcox(ALTURA ~ DOSE,
                        ylab = "Logaritmo da verossimilhança")
  title(main = "DIC Alho | Gráfico Box-Cox")
  print(plot_boxcox)
})
```



- **Transformação de dados**

Portanto → Com λ próximo a 0, aplica-se a transformação $\log(y + 0,5)$ para a variável altura de planta.

```
DICgarlic$ALTURAt <- log(DICgarlic$ALTURA)
```

- **Levene** | Teste de homogeneidade de variâncias

```
with(DICgarlic,
  levene.test(ALTURAt, DOSE, location = "mean"))
```

Resultado:

```
data: ALTURAt
Test Statistic = 1.8169, p-value = 0.1443
```

Portanto → De acordo com o teste de Levene a 5% de probabilidade de erro, as variâncias podem ser consideradas homogêneas.

- **Shapiro-Wilk** | Teste de normalidade dos resíduos

```
# Modelo linear após transformação
lmDICT = lm(ALTURAt~DOSE, DICgarlic)
resDICT <- residuals(lmDICT)          # Resíduos
resStudDICT <- rstandard(lmDICT)      # Resíduos studentizados

shapiro.test(resStudDICT)
```

Resultado:

```
data: resStudDICT
W = 0.97689, p-value = 0.7706
```

Portanto → De acordo com o teste de Shapiro-Wilk a 5% de probabilidade de erro, os resíduos podem ser considerados normais.

▼ ANOVA - Análise de variância

- Anova utilizando funções do pacote [ExpDes.pt](#):

```
# Code:
DICgarlic$DOSE <- as.numeric(DICgarlic$DOSE)
with(DICgarlic,
  dic(DOSE, ALTURAt, hvar = "levene", quali = F, mcomp = "tukey", sigF = 0.05, sigT = 0.05))
```

```
# ANOVA por meio do ExpDes.pt:
```

```
-----  
Quadro da análise de variancia
```

```
-----  
          GL      SQ      QM  Fc      Pr>Fc  
Tratamento  6 12.4883 2.08138 283 5.752e-19  
Residuo     21  0.1545 0.00735  
Total       27 12.6427  
-----
```

```
CV = 2.65 % [...]
```

- Interpretações:
 - A hipótese nula é rejeitada, aceitando-se a hipótese alternativa de que há diferenças significativas a 5% de probabilidade de erro entre, pelo menos, dois tratamentos.
 - Logo, é possível dizer que a radiação gama promoveu influência sobre a variável altura de plantas de alho.

▼ Regressão Polinomial

- Para realizar a regressão, os regressores foram estimados por meio da função “dic” presente no pacote já mencionado, o [ExpDes.pt](#):

```
with(DICgarlic,  
      dic(DOSE, ALTURAt, hvar = "levene", quali = F, mcomp = "tukey", sigF = 0.05, sigT = 0.05))
```

Resultado:

```
[...]
```

```
Ajuste de modelos polinomiais de regressao
```

```
-----  
Modelo Linear
```

```
=====
```

	Estimativa	Erro.padrao	tc	valor.p
b0	4.2107	0.0292	144.1153	0
b1	-0.1295	0.0032	-39.9481	0

```
-----
```

```
R2 do modelo linear
```

```
-----  
0.939857  
-----
```

```
Análise de variancia do modelo linear
```

```
=====
```

	GL	SQ	QM	Fc	valor.p
Efeito linear	1	11.7372	11.7372	1595.85	0
Desvios de Regressao	5	0.7511	0.1502	20.42	0
Residuos	21	0.1544	0.0073		

```
-----  
-----
```

```
Modelo quadratico
```

```
=====
```

	Estimativa	Erro.padrao	tc	valor.p
b0	4.1354	0.0374	110.4877	0
b1	-0.0933	0.0117	-7.9863	0
b2	-0.0024	0.0008	-3.2197	0.0041

```
-----
```

```
R2 do modelo quadratico
```

```
-----  
0.945962
```

```

-----

Analise de variancia do modelo quadratico
=====
              GL    SQ      QM      Fc      valor.p
-----
Efeito linear      1 11.7372 11.7372 1595.85      0
Efeito quadratico  1  0.0762  0.0762   10.37  0.00411
Desvios de Regressao 4  0.6748  0.1687   22.94      0
Residuos           21 0.1544  0.0073
-----

Modelo cubico
=====
      Estimativa Erro.padrao    tc    valor.p
-----
b0    3.9760      0.0413    96.2227      0
b1    0.1193      0.0261    4.5702  0.0002
b2   -0.0407      0.0043   -9.5334      0
b3    0.0017      0.0002    9.1099      0
-----

R2 do modelo cubico
-----
0.994839
-----

Analise de variancia do modelo cubico
=====
              GL    SQ      QM      Fc      valor.p
-----
Efeito linear      1 11.7372 11.7372 1595.85      0
Efeito quadratico  1  0.0762  0.0762   10.37  0.00411
Efeito cubico      1  0.6104  0.6104   82.99      0
Desvios de Regressao 3  0.0645  0.0215    2.92  0.05781
Residuos           21 0.1544  0.0073
-----

```

Com base nos resultados obtidos, é possível concluir que:

- O modelo com melhor ajuste é o cúbico, com R^2 de 99.48%;
- No modelo cúbico:

- Efeito linear foi significativo;
- Efeito quadrático foi significativo;
- Efeito cúbico foi significativo;

∴ Todos os regressores estimados foram significativos. Portanto, pode-se utilizar esse polinômio de maior grau para ajustar aos dados do experimento avaliado.

- A equação geral:

$$f(x) = 3.9760 + 0.1193 \times x - 0.0407 \times x^2 + 0.0017 \times x^3$$

- A equação encontrada reflete a altura de plantas **transformada por meio do log natural** dos dados observados. Portanto a equação final encontrada seria:

$$f(x) = \exp(3.9760 + 0.1193 \times x - 0.0407 \times x^2 + 0.0017 \times x^3)$$

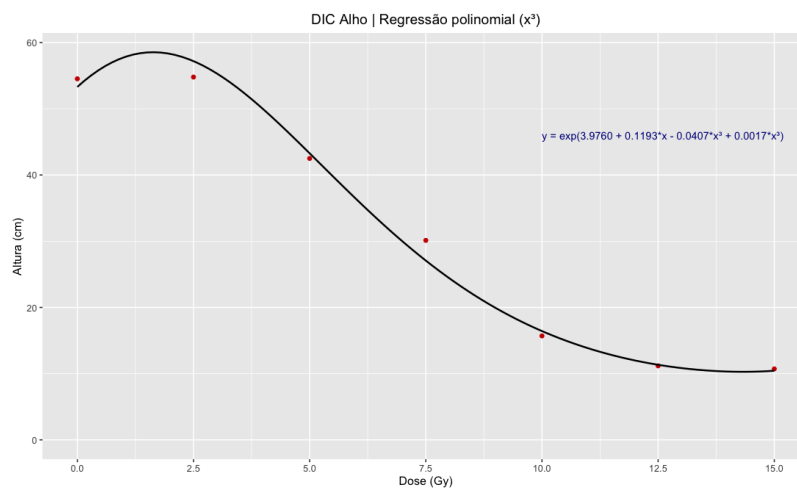
- Logo, podemos obter um gráfico com a média dos valores de altura e a regressão encontrada:

```

fun2 <- function(x){
  exp(3.9760 + 0.1193*x - 0.0407*x^2 + 0.0017*x^3)
}

```

```
ggplot(DICgarlic, aes(x=DOSE, y = ALTURA)) +
  stat_summary(fun = mean, geom = "point", color = "red3") +
  stat_function(fun = fun2, color = "black", size = 0.8) +
  expand_limits(y = 0) +
  scale_x_continuous(breaks = seq(min(DICgarlic$DOSE), max(DICgarlic$DOSE), by = 2.5)) +
  labs(
    x = "Dose (Gy)",
    y = "Altura (cm)",
    title = "DIC Alho | Regressão polinomial (x³)" +
  theme(
    plot.title = element_text(hjust = 0.5)) +
  annotate("text", x = 10, y = 50, label = "y = exp(3.9760 + 0.1193*x - 0.0407*x² + 0.0017*x³)",
    color = "blue4", size = 3.5, hjust = 0, vjust = 4)
```



Regressão cúbica aplicada aos dados de altura de plantas de alho, aos 42 DAP, sob diferentes doses de irradiação gama.

CONCLUSÕES

- A regressão polinomial foi bem sucedida e conseguiu determinar regressores com bom ajuste ao conjunto de dados.
- O possível melhor ajuste de um modelo cúbico advém da natureza esperada dos dados obtidos. A mutagênese por meio de radiação gama promove um incremento leve sob doses menores, seguido por uma queda em doses seguintes, efeito conhecido como **Hormese**. Logo, a dose imediatamente superior a testemunha, de 2,5 Gy promoveu ganhos para a variável altura de plantas, no entanto, doses maiores são seguidas por quedas progressivas no vigor do alho.