

Laboratório 2 - Moscas e Longevidade (Continuação)

Contexto

No laboratório anterior, investigamos o custo reprodutivo em moscas-das-frutas machos utilizando uma ANOVA de um fator, considerando apenas uma variável por vez (Número de parceiras, Tipo de parceiras ou Tratamento). Essa abordagem nos permitiu identificar diferenças de longevidade entre os diferentes grupos do experimento.

Para relembrar:

- Qual era o propósito geral do estudo?
- Quais hipóteses estatísticas foram formuladas?
- Como a Longevidade se comportava nos diferentes níveis de cada fator:
 - Número de parceiras (0, 1, ou 8),
 - Tipo de parceiras (nenhuma, virgem, prenha),
 - Tratamento (none, 1 pregnant, 1 virgin, 8 pregnant, 8 virgin)?
- Quais foram as conclusões principais do teste ANOVA realizado no Lab 1?

O seguinte código pode ser utilizado para carregar os dados.

```
#Pacotes
if (!require("pacman")) install.packages("pacman")
```

Carregando pacotes exigidos: pacman

```
pacman::p_load(dplyr)
# Carregar os dados
base_url <- "https://raw.githubusercontent.com/vincentarelbundock/Rdatasets/"</pre>
file path <- "ba3186b8413f4d40d8f38120c32bcd31797e8d5e/csv/Stat2Data/FruitFlies.csv"
dat <- read.csv(paste0(base_url, file_path)) |>
  select(ID, Longevity, everything(), -rownames) |>
  mutate(
     Treatment =
       factor(Treatment,
              levels = c(
                "none",
                "1 pregnant", "1 virgin", "8 pregnant", "8 virgin")
              ),
     Type = factor(
       case_when(
          Type == 0 ~ "prenha",
          Type == 1 ~ "virgem",
          Type == 9 ~ "nenhuma",
          TRUE ~ NA_character_
      levels = c("nenhuma", "prenha", "virgem")
    )
```



#Partners em fator dat\$Partners <- as.factor(dat\$Partners)</pre> head(dat)

ID	Longevity	Partners	Type	Thorax	Sleep	Treatment
1	35	8	prenha	0.64	22	8 pregnant
2	37	8	prenha	0.68	9	8 pregnant
3	49	8	prenha	0.68	49	8 pregnant
4	46	8	prenha	0.72	1	8 pregnant
5	63	8	prenha	0.72	23	8 pregnant
6	39	8	prenha	0.76	83	8 pregnant

🛕 Atenção: Ordem padrão dos níveis de fatores no R

Por padrão, o R ordena os níveis de fatores alfabeticamente, não na ordem em que aparecem pela primeira vez nos dados. Por exemplo, se você criar um fator com os valores c("medium", "large", "small"), os níveis serão automaticamente ordenados como "large", "medium", "small". Esta ordenação é lexicográfica e pode depender da configuração local (locale) do sistema, o que pode gerar resultados inconsistentes entre diferentes plataformas.

Para controlar explicitamente a ordem dos níveis, sempre utilize o argumento levels na função factor() para especificar a ordem desejada, especialmente quando a ordem alfabética não reflete a ordem lógica ou hierárquica dos seus dados (como em escalas ordinais: "baixo", "médio", "alto").



1. Considere o banco sem o tratamento "none", utilizando, por exemplo, a seguinte transformação.

dat2 <- subset(dat, Treatment != 'none')</pre>

- a. Ajuste: (i) modelo de um fator usando treatment; (ii) modelo de dois fatores (type e partners) com interação.
- b. Apresente a tabela ANOVA e os parâmetros estimados para cada modelo. Quais semelhanças/diferenças existem?
- c. Considerando um nível de significância de $\alpha=0.05$, quais são as conclusões em relação ao modelo com dois fatores?
- d. Cheque os pressupostos do modelo com dois fatores (QQ-plot, teste Shapiro-Wilk, resíduos vs. ajustados, boxplot resíduos por níveis dos fatores, teste Levene, teste Bartlett). Quais são as suas conclusões?
- e. Interprete os parâmetros ajustados no modelo com dois fatores.
- f. Ajuste um modelo de dois fatores (type e partners) **sem** interação. Apresente a tabela ANOVA e as conclusões para $\alpha = 0.05$.
- g. Verifique pressupostos.
- h. Apresente e interprete os parâmetros estimados.
- i. Compare os modelos de dois fatores com e sem interação. Qual deles você escolheria do ponto de vista estatístico e do ponto de vista biológico? Justifique sua escolha.

2. Considere agora o banco original (com "none").

- a. Discuta a parametrização mais apropriada para um modelo de dois fatores com interação.
- b. Ajuste o modelo acima e apresente a tabela ANOVA. Quais são as suas conclusões? (Considere $\alpha = 0.05$).
- c. Verifique os pressupostos.
- d. Se ambos os fatores (Type e Partners) tivessem 3 níveis completos cada, um modelo de dois fatores com interação total teria um certo número esperado de parâmetros.
 - i. Quantos parâmetros você esperaria nesse caso ideal (explique como obteve esse número)?
 - ii. Quantos parâmetros foram realmente estimados ao ajustar o modelo com o banco original (incluindo o tratamento none)?
 - iii. Por que há diferença entre os dois números? Relacione sua resposta à presença de células ausentes ou restrições do delineamento.
- e. Apresente e interprete parâmetros.
- f. Repita os itens (b-e) para o modelo de dois fatores **sem** interação.
- g. Entre os modelos com e sem interação, qual escolheria e por quê?



Funções do R

Os estudantes são encorajados a consultar a documentação das funções listadas abaixo (?nome_da_funcao no R) e descobrir como elas funcionam através da experimentação e prática.

read.csv()	tidyr::pivot_longer()		
<pre>dplyr::select()</pre>	<pre>dplyr::bind_rows()</pre>		
<pre>dplyr::mutate()</pre>	<pre>broom::tidy()</pre>		
factor()	<pre>dplyr::arrange()</pre>		
<pre>dplyr::case_when()</pre>	ggplot2::ggplot()		
head()	<pre>ggplot2::geom_boxplot()</pre>		
<pre>subset()</pre>	<pre>ggplot2::geom_point()</pre>		
aov()	ggplot2::geom_col()		
lm()	<pre>ggplot2::geom_text()</pre>		
<pre>summary()</pre>	<pre>ggplot2::facet_wrap()</pre>		
TukeyHSD()	ggplot2::labs()		
<pre>car::linearHypothesis()</pre>	ggplot2::theme()		
<pre>shapiro.test()</pre>	<pre>ggplot2::theme_minimal()</pre>		
<pre>bartlett.test()</pre>	<pre>ggplot2::scale_y_continuous()</pre>		
<pre>car::leveneTest()</pre>	plot()		
<pre>dplyr::group_by()</pre>	abline()		
<pre>dplyr::summarise()</pre>	par()		
cor()	fitted()		
mean()	residuals()		
round()	car::qqPlot()		
sum()	coef()		
rowSums()	names()		
dplyr::n()	<pre>knitr::kable()</pre>		
	<pre>kableExtra::kable_styling()</pre>		
	library()		
	c()		
	rbind()		
-			