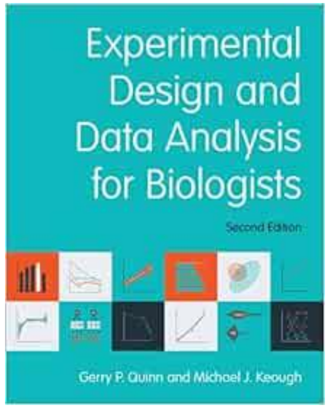
Aula 11 – Modelos Lineares

Análise de Variância (ANOVA)

Ao final da aula você será capaz de:

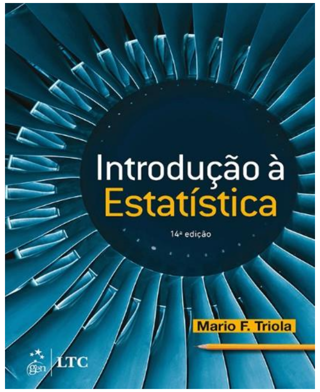
1. Definir a Análise de Variância (ANOVA) e seu propósito na comparação das médias de três ou mais grupos.
2. Explicar as principais premissas (pressupostos) da ANOVA.
3. Descrever o conceito fundamental de partição da variância.
4. Ler e interpretar os componentes de uma tabela de ANOVA de um fator.
5. Formular uma conclusão com base nos resultados da ANOVA de um estudo biológico.
6. Entender o propósito dos testes a posteriori (*post-hoc*) após um resultado significativo na ANOVA.

Aula baseada em:

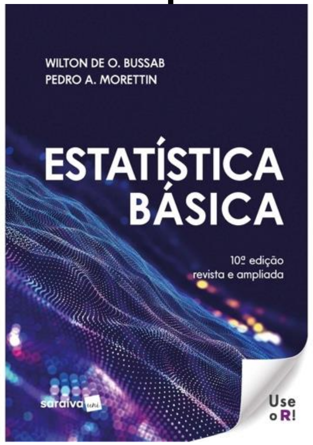


Cap 6

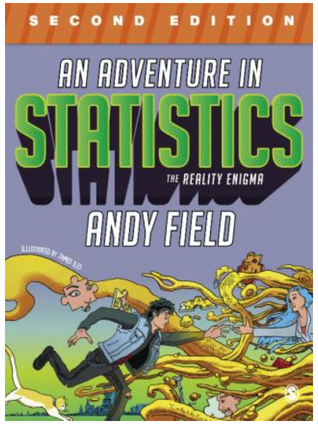
Cap 13



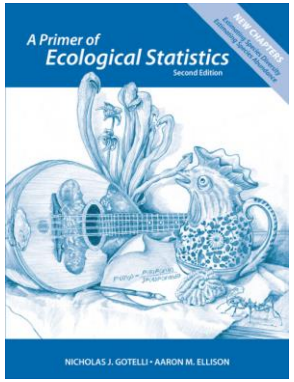
Cap 12



Cap 15



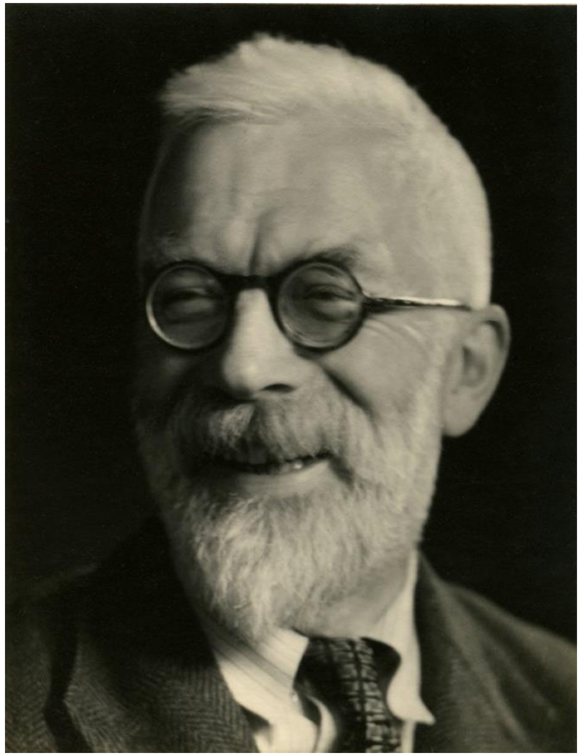
Cap 16



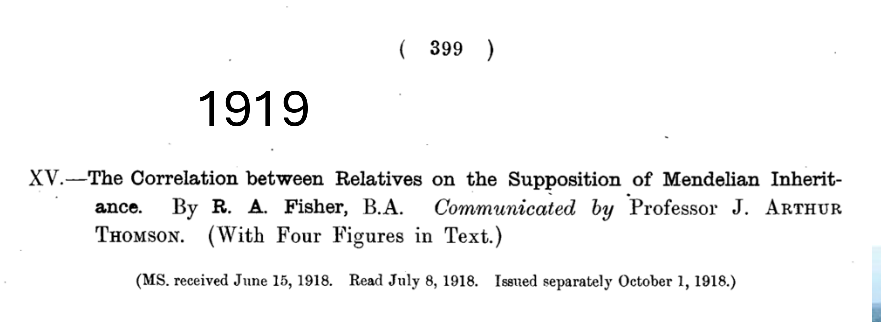
Cap 10

Antes, um pouco de história...

ROTHAMSTED RESEARCH



Ronald Aylmer Fisher

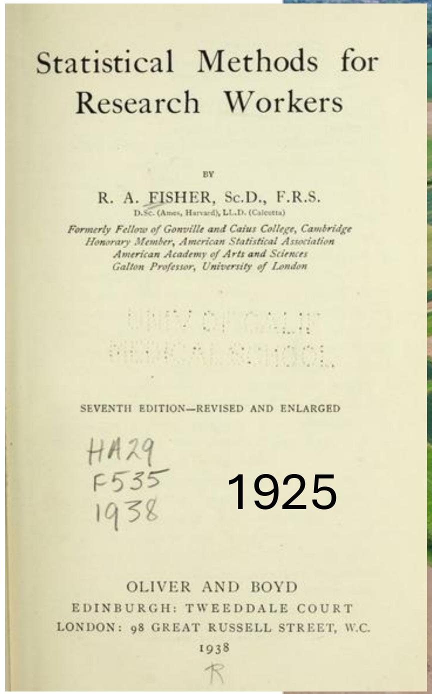


(399)

1919

XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. By R. A. Fisher, B.A. *Communicated by* Professor J. Arthur Thomson. (With Four Figures in Text.)

(MS. received June 15, 1918. Read July 8, 1918. Issued separately October 1, 1918.)



Statistical Methods for Research Workers

BY

R. A. FISHER, Sc.D., F.R.S.

D.Sc. (Ames, Harvard), LL.D. (Calcutta)

Formerly Fellow of Gonville and Caius College, Cambridge
Honorary Member, American Statistical Association
American Academy of Arts and Sciences
Galton Professor, University of London

SEVENTH EDITION—REVISED AND ENLARGED

HA29
F535
1938

OLIVER AND BOYD
EDINBURGH: TWEEDDALE COURT
LONDON: 98 GREAT RUSSELL STREET, W.C.
1938

1925

Questão biológica

- Imaginem que um ecólogo queira testar se a taxa média de crescimento das raízes de mangue é diferente quando expostas a duas espécies distintas de esponjas, um controle com uma esponja falsa e um grupo sem esponja.
- Com quatro grupos, como podemos determinar estatisticamente se existe uma diferença real no crescimento médio entre eles?

Problema de +2 categorias

Realizar múltiplos testes *t* entre todos os pares de grupos é problemático

Inflaciona a probabilidade de cometer um erro do Tipo I (encontrar incorretamente uma diferença significativa).

Nestes casos devemos usar uma Análise de Variância (ANOVA)

Modelo linear

- A melhor pergunta que um cientista pode fazer é "Qual modelo matemático descreve a relação entre minhas variáveis?”

$$Dados = Modelo + Erro$$

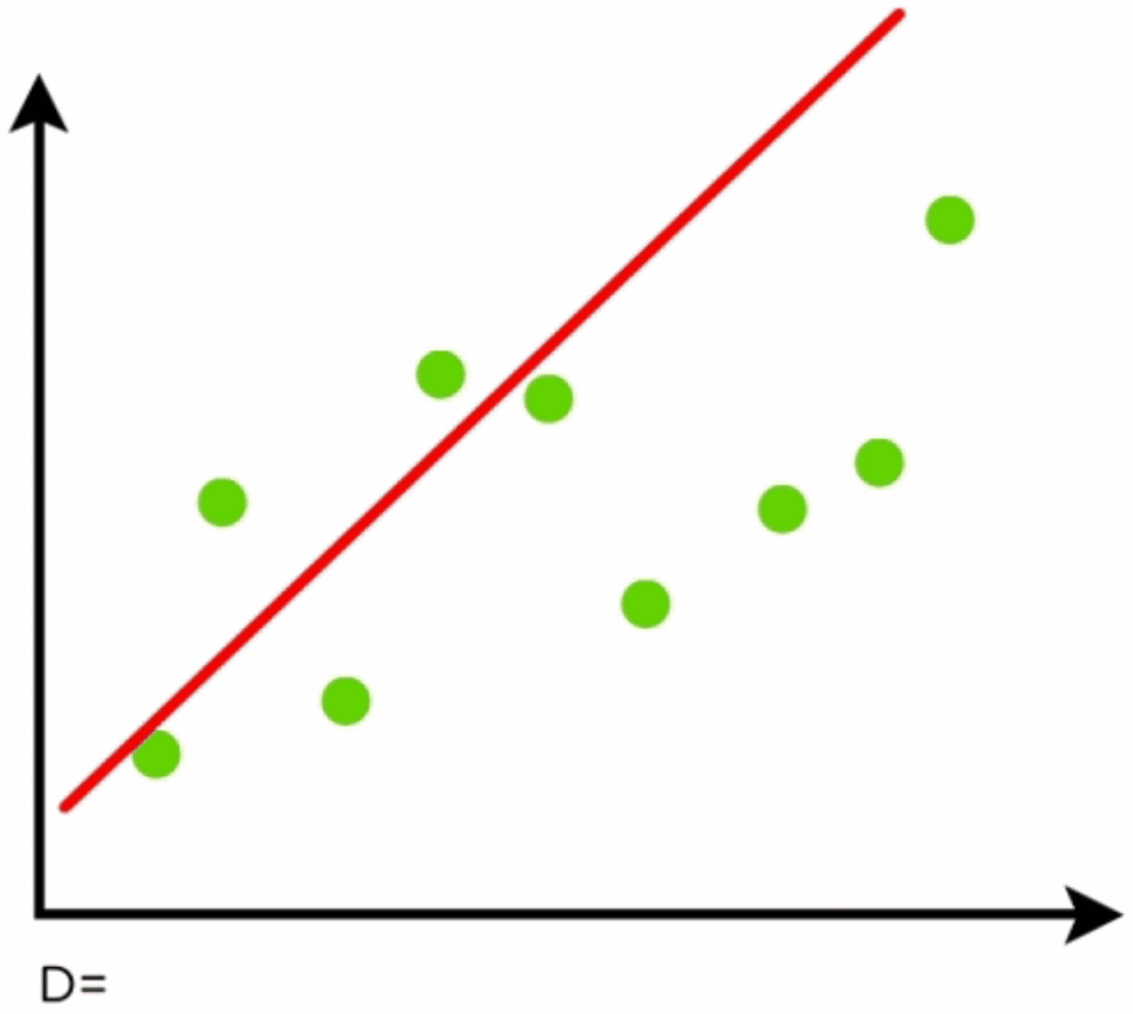
- **DADOS**: Sua variável de resposta contínua (ex: altura da planta, diversidade de espécies).
- **MODELO**: A sua hipótese científica expressa matematicamente. É a nossa melhor tentativa de explicar a variação nos dados.
- **ERRO**: A variação que nosso modelo não consegue explicar (variação aleatória ou residual)

Modelo linear

- Nosso trabalho como cientistas é construir o melhor modelo possível para minimizar o erro e explicar a maior parte da variação nos dados.

- A ANOVA é simplesmente o nome que damos quando construímos um modelo linear usando **<u>um preditor categórico</u>**

- Já sabemos como lidar com este tipo de variável (fator) usando variáveis *dummy* para traduzir



https://medium.com/@aulakhkaran5a/linear-regression-b712b88bb940

$$\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-2}$$

Quadrado Médio dos Resíduos

Critério de otimização para ajuste da reta de regressão (Mínimos Quadrados)

Construindo o modelo de ANOVA

Equação do modelo linear passo-a-passo

Dados = modelo + erro

Equação do modelo linear passo-a-passo

$$y_i = \alpha + \beta x_i + \varepsilon$$

Equação do modelo linear passo-a-passo

$$\boxed{y_i} = \boxed{\alpha + \beta x_i} + \boxed{\varepsilon}$$

Equação do modelo linear passo-a-passo

$$y_i = \alpha + \beta x_i + \varepsilon$$

Variável resposta

Variável preditora

Equação do modelo linear passo-a-passo

$$y_i = \alpha + \beta x_i + \varepsilon$$

intercepto

Inclinação (slope)

ANOVA: modelo linear com preditor categórico

- Agora, e se nosso preditor X não for contínuo? E se ele for categórico, como "Tipo de Fertilizante" (A, B, C) ou "Nível de Zinco" (Baixo, Médio, Alto)?

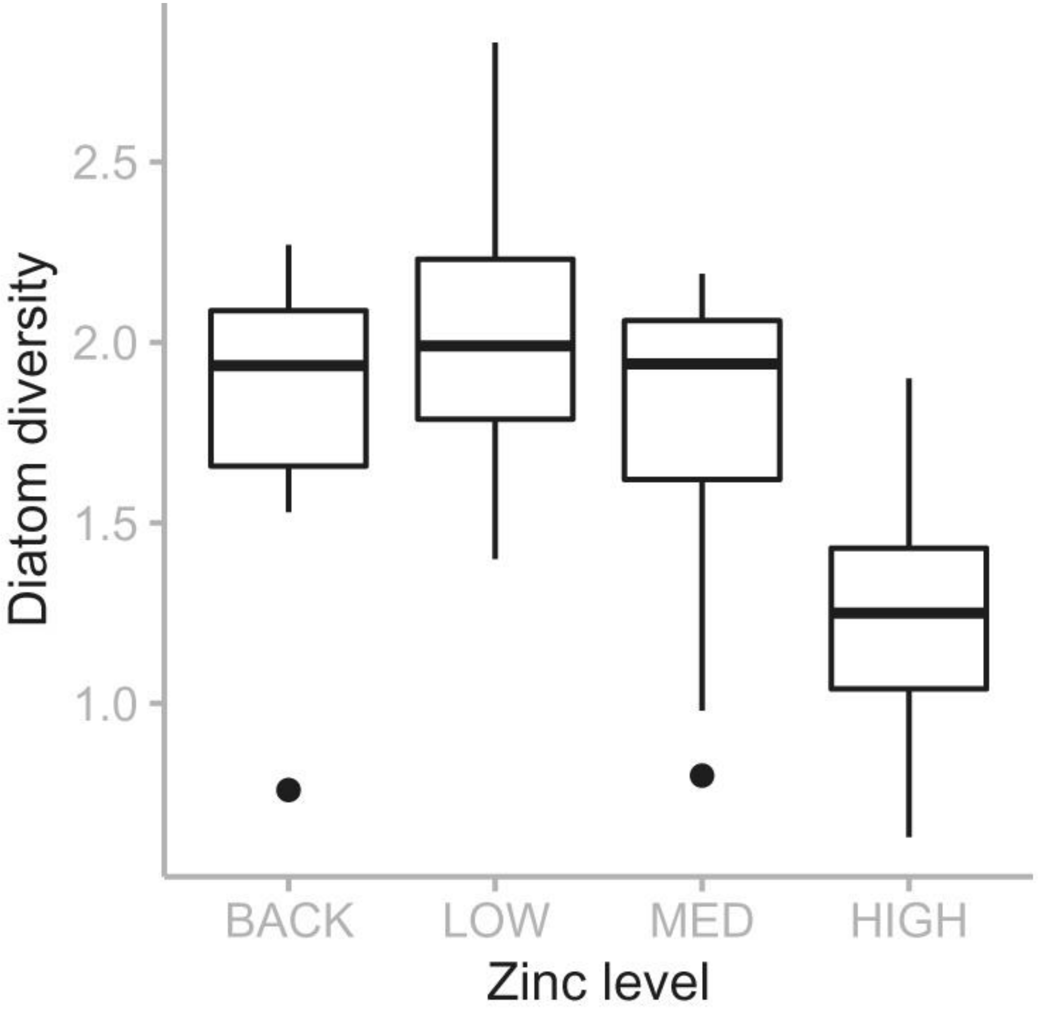


Fig. 6.6 Quinn & Keough 2023

(a) Zinc level	B	L	M	H
	0.8	0.7	1.8	2.6
	0.9	1.7	2.1	0.6
	2.4	1.0	0.6	1.2
	1.4	1.4	1.1	1.3
	1.3	1.2	2.4	2.2
	1.8	2.4	1.2	0.9
	2.1	1.1	0.9	1.9
	1.0	2.2	1.6	1.0
Means	1.4625	1.4625	1.4625	1.4625

Alta var. dentro de grupos

Baixa var. entre grupos

Média gr. é =

(b) Zinc level	B	L	M	H
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
	1.2	2.3	1.8	0.7
Means	1.2	2.3	1.8	0.7

Baixa var. dentro de grupos (obs=μ)

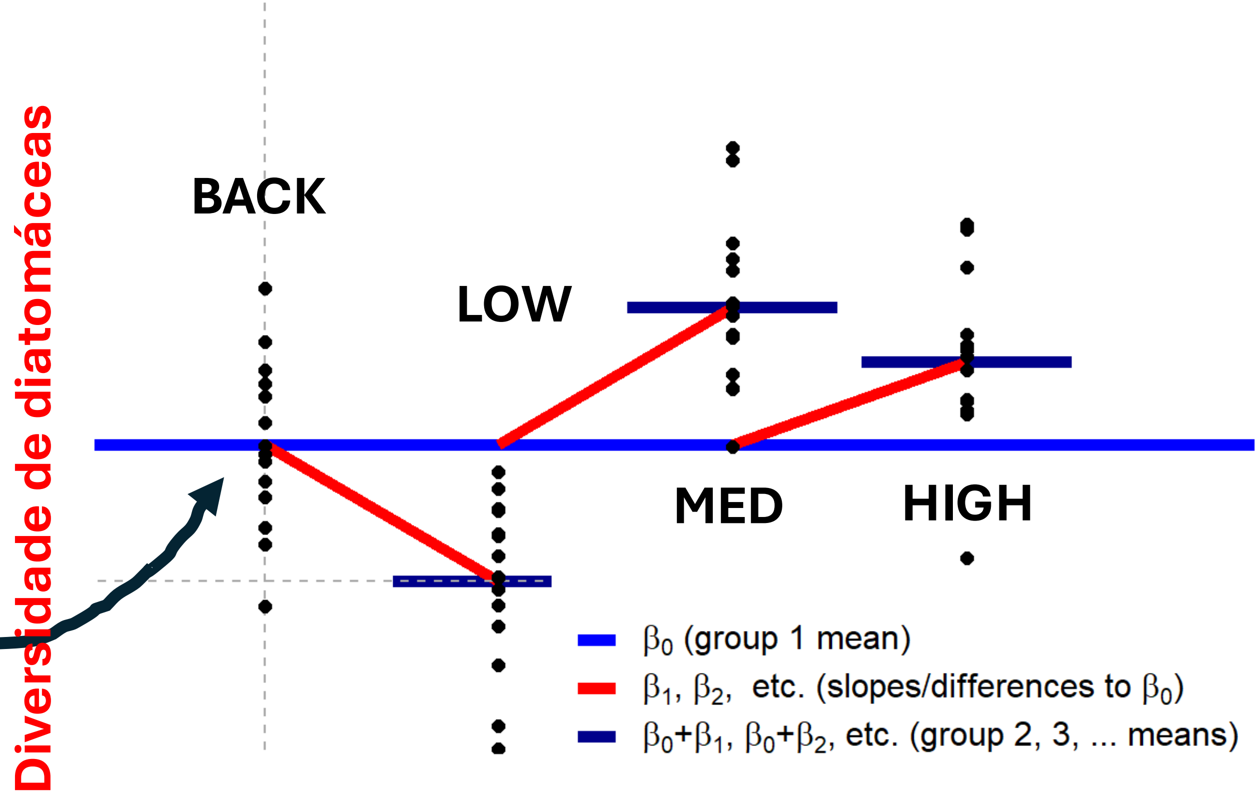
Alta var. entre grupos

Média gr. é ≠

ANOVA: modelo linear com preditor categórico

- Não podemos calcular uma única “inclinação” para categorias. Em vez disso, nosso modelo precisa descrever os dados usando as médias de cada grupo.

Grupo (BACK) é usado como referência



https://lindeloev.github.io/tests-as-linear/

ANOVA: modelo linear com preditor categórico

- Existem duas formas equivalentes de escrever este modelo:
  - O modelo de medidas por célula
  - Modelo de efeitos (formato de regressão)

O modelo de medidas por célula

$$y_{ij} = \mu_i + \varepsilon_{ij}$$

- Isso significa que o valor de uma observação individual (y_{ij}) é igual à média do seu grupo (μ_i) mais algum erro (ε_{ij}).
- Por exemplo, a diversidade de diatomáceas numa unidade amostral específica é o valor médio para aquele nível de zinco, mais a variação residual daquela u.a.

O modelo de efeitos

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

- Isso diz que o valor de uma observação individual (y_{ij}) é igual à média geral de todos os dados (μ), mais o efeito específico de pertencer ao grupo *i* (α_i), mais o erro aleatório (ε_{ij}).
- **O "efeito" (α_i) é simplesmente a diferença entre a média do grupo *i* e a média geral**
- Ambos são matematicamente equivalentes e descrevem o mesmo fenômeno.
- O Modelo de Efeitos é útil porque se parece mais com a estrutura intercepto + efeito da regressão.

Como a ANOVA funciona

- Para uma ANOVA de um fator, a variação total (Soma de Quadrados) é dividida em duas partes :
- **Variação Entre Grupos**: variação entre as médias amostrais de cada grupo. Reflete a influência do fator experimental ou do tratamento.
  - Por exemplo, se diferentes fertilizantes tiverem um efeito forte, a altura média das plantas para cada grupo será muito diferente, e a variação "entre grupos" será grande.
- **Variação Dentro dos Grupos (ou Residual/Erro)**: Mede a variação aleatória e não explicada entre as observações individuais dentro de cada grupo. Isso é considerado o "ruído" ou erro experimental nos dados.

Como a ANOVA funciona

- O princípio fundamental da ANOVA é que esses componentes são aditivos:

$$SQ_{Total} = SQ_{Entre\ grupos} + SQ_{Dentro\ grupos}$$

- Soma de quadrados

$$\sum (x - \bar{x})^2$$

- No entanto, grupos com mais amostras podem ter variação maior
- Temos de dividir a soma de quadrados pelos graus de liberdade, este é o Quadrado Médio (QM)

A estatística F

- A ANOVA compara a variação entre os grupos com a variação dentro dos grupos usando a estatística *F* (também chamada de razão F)

$$F = \frac{Variância\ \boldsymbol{entre}\ as\ amostras}{Variância\ \boldsymbol{dentro}\ das\ amostras} = \frac{QM_{grupos}}{QM_{Resíduos}}$$

A estatística F

- Se a estatística F for grande, significa que a variação entre os grupos é muito maior do que a variação aleatória dentro dos grupos. Isso sugere que o tratamento teve um efeito real.

- Se a estatística F for próxima de 1, significa que a variação entre os grupos é semelhante à variação aleatória, sugerindo que quaisquer diferenças entre as médias dos grupos são provavelmente devidas ao acaso

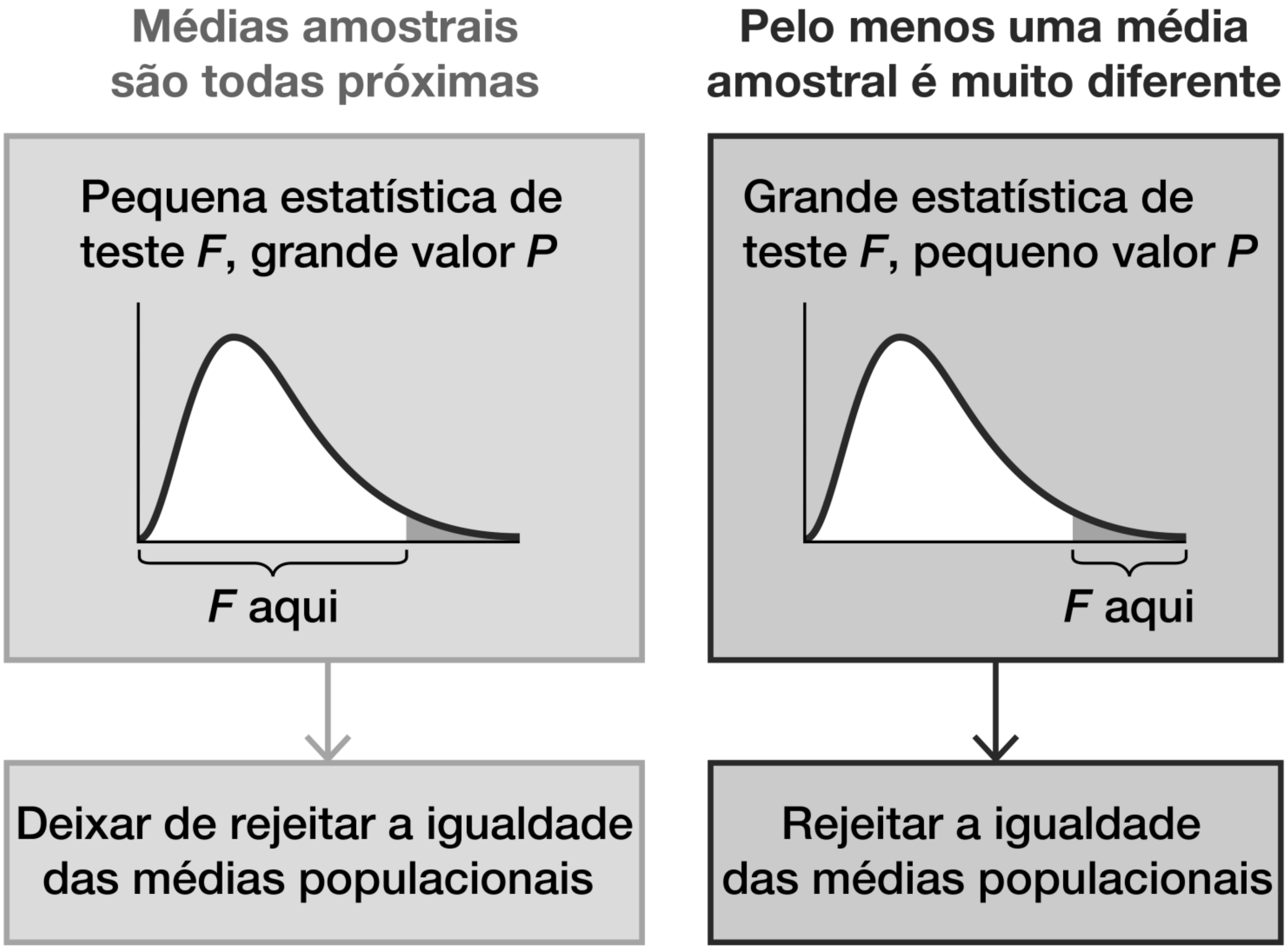


FIGURA 12.2 Relação Entre a Estatística de Teste *F* e o Valor *P*.

Triola

Tabela 13.4 Análise de variância de um experimento completamente ao acaso

Causa de variação	**GL**	**SQ**	**QM**	***F***
Grupos	$(t - 1)$	*SQG*	*QMG*	*F*
Resíduo	$(n - t)$	*SQR*	*QMR*	
Total	$(n - 1)$	*SQT*		

Tabela 15.5 Tabela de Análise de Variância (ANOVA).

F.V.	g.l.	SQ	QM	F
Entre	1	SQEnt	QMEnt	QMEnt/S^2_e
Dentro	n - 2	SQDen	QMDen (ou S^2_e)	
Total	n - 1	SQTot	QMTot (ou S^2)	

Bussab & Morettin

Tabela ANOVA

$$C = \frac{\left(\sum_i^g \sum_j^{n_i} x_{ij}\right)^2}{N}$$

Fonte de Variação	G.L.	Soma de Quadrados	Variância	*F*
Entre (efeito)	$g-1$	$SQ_E = \sum_i^g \frac{\left(\sum_j^{n_i} x_{ij}\right)^2}{n_i} - C$	$s_E^2 = \frac{SQ_E}{G.L._E}$	$F = \frac{s_E^2}{s_D^2}$
Dentro (erro)	$N-g$	$SQ_D = SQ_T - SQ_E$	$s_D^2 = \frac{SQ_D}{G.L._D}$	
Total	$N-1$	$SQ_T = \sum_i^g \sum_j^{n_i} x_{ij}^2 - C$		

O teste F como uma comparação de modelos

- A hipótese nula na ANOVA (H_0: μ_1=μ_2 =μ_3 =…) pode ser vista como uma comparação entre dois modelos:
- O **Modelo Completo** (Nossa Hipótese): Inclui o preditor categórico. Ele assume que as médias dos grupos são diferentes.
  - $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$
  - A variação não explicada por este modelo é a Soma dos Quadrados do Resíduo (SQ_{Res})
- O **Modelo Reduzido** (ou Nulo): assume que o preditor categórico não tem efeito (ou seja, todos α_i são zero). Essencialmente, ele assume que todas as observações vêm de uma única população com uma única média.
  - $y_{ij} = \mu + \varepsilon_{ij}$
  - A variação não explicada por este modelo é a Soma dos Quadrados Total (SQ_{Total}), pois a melhor previsão para qualquer ponto é simplesmente a média geral

O teste F como uma comparação de modelos

- O teste F avalia se o Modelo Completo é significativamente melhor em explicar a variação dos dados do que o Modelo Reduzido. A melhora na explicação, ou seja, a redução no erro, é a Soma dos Quadrados Entre Grupos (SQ_{Grupos}).

$$SQ_{grupos} = SQ_{Total}(erro\ modelo\ nulo) - SQ_{Res}(erro\ modelo\ completo)$$

- Portanto, a tabela ANOVA é, na verdade, um resumo da comparação entre esses dois modelos

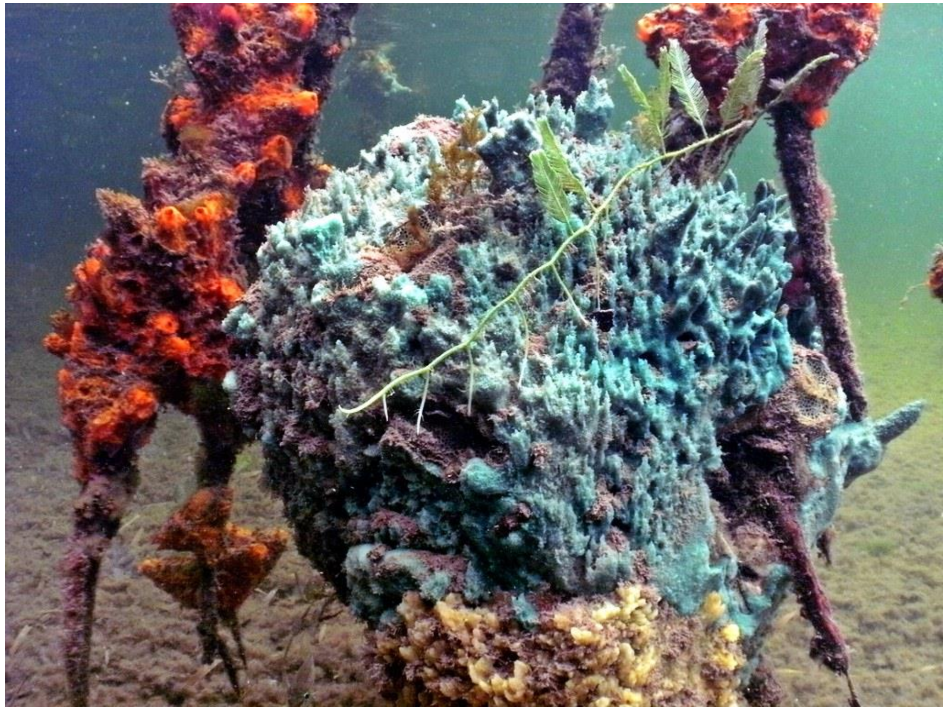
Premissas da ANOVA

- Para que o teste F forneça um valor-p confiável, os dados devem atender a três premissas :
  - **Independência**: As amostras são aleatórias e as observações dentro e entre os grupos são independentes umas das outras.
  - **Normalidade**: As populações das quais as amostras são retiradas são aproximadamente normalmente distribuídas. A ANOVA é geralmente robusta a violações moderadas desta premissa, especialmente com tamanhos de amostra iguais ou quase iguais.
  - **Homogeneidade de Variâncias** (Homocedasticidade): As variâncias de cada grupo são aproximadamente iguais. Esta é uma premissa crítica, e violações podem afetar seriamente a confiabilidade do teste F, particularmente se os tamanhos das amostras forem desiguais.

Exemplo biológico

Efeito de esponjas no crescimento de manguezal

- Um ecólogo marinho está estudando as raízes de mangue em um estuário. Ele percebe que diferentes espécies de esponjas se fixam nessas raízes e suspeita que a presença de certas esponjas pode afetar a taxa de crescimento das raízes.



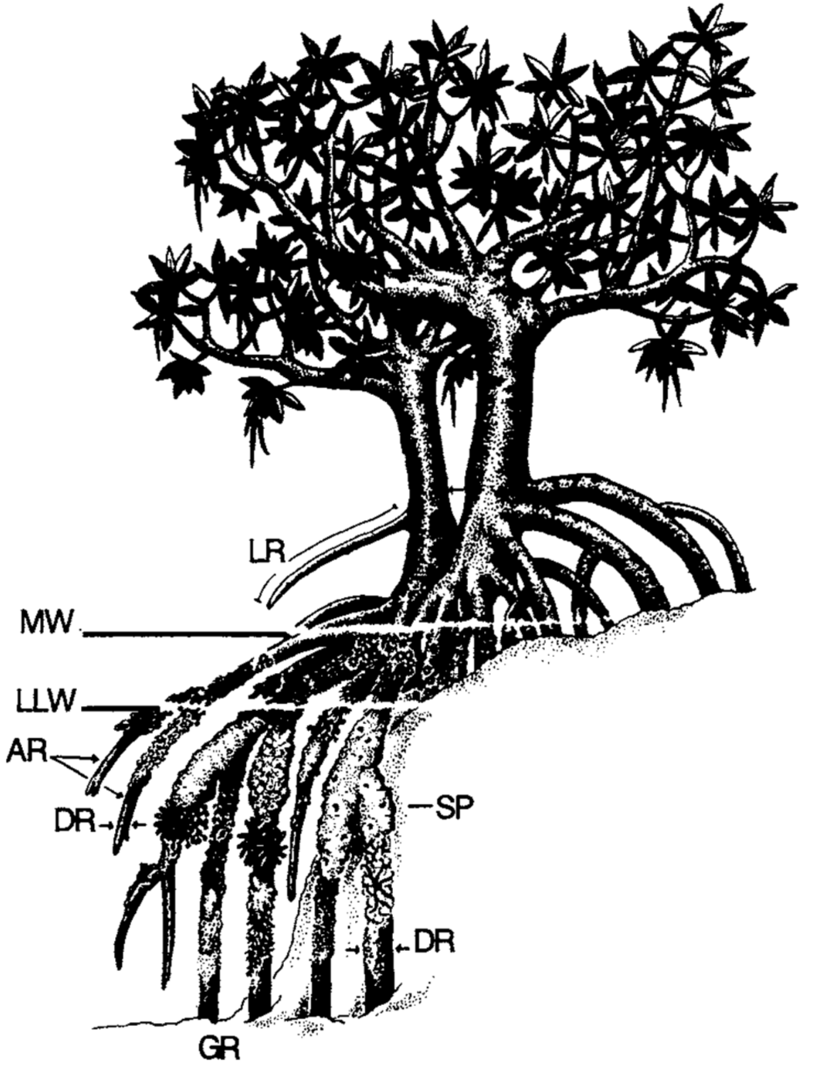
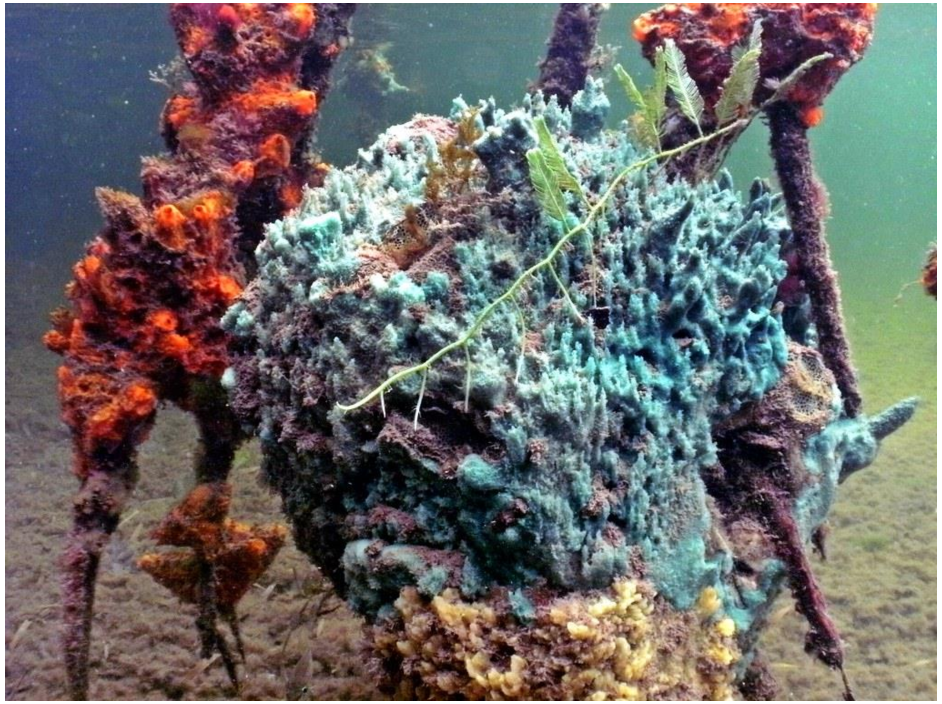
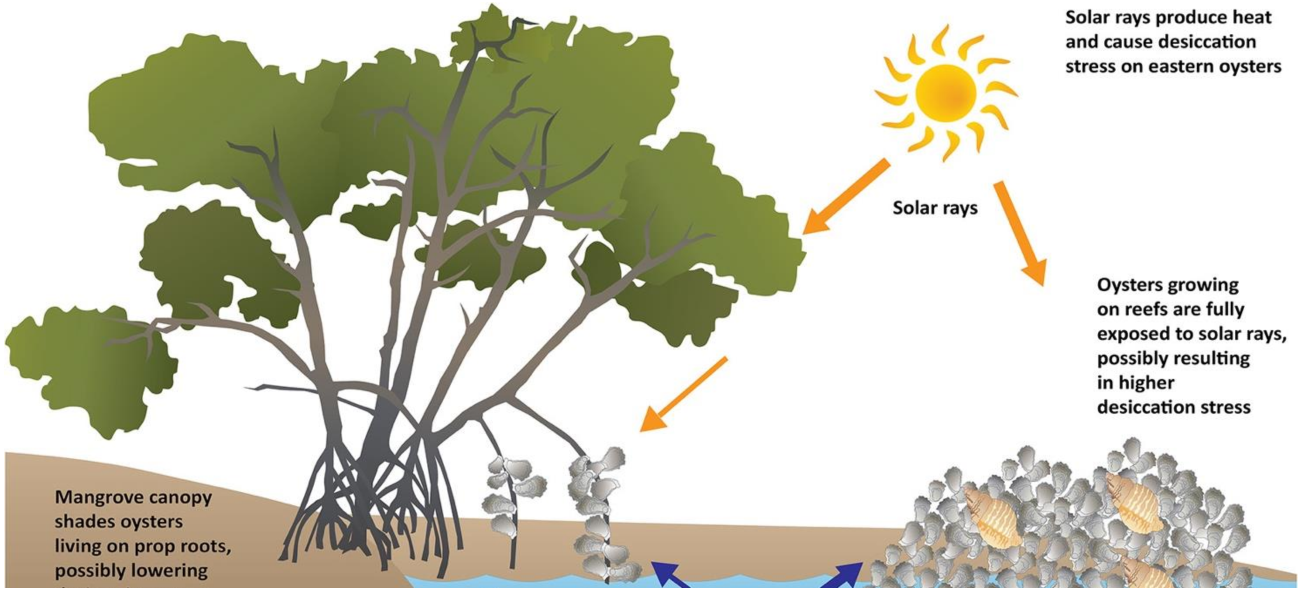


FIG. 2. Illustration of a fringing red mangrove tree at Twin Cays, showing relationship of cable roots to mean water (MW) and lowest low water (LLW) tidal levels. Aerial roots





Ellison et al. 1996 Ecology

Chacin et al. 2025 JEMBE

Efeito de esponjas no crescimento de manguezal

- Para investigar isso, ele estabelece um experimento com quatro condições nas raízes:
  - **Grupo Controle (Sem Esponja)**: Raízes sem nenhuma esponja.
  - **Esponja Falsa (Controle de Manipulação)**: Raízes com uma esponja artificial de espuma (para controlar o efeito físico da presença da esponja, mas sem a biologia da esponja viva).
  - **Esponja Tipo A**: Raízes com a espécie de esponja *Haliclona* sp.
  - **Esponja Tipo B**: Raízes com a espécie de esponja *Tedania* sp.

Efeito de esponjas no crescimento de manguezal

- O ecólogo mede a **taxa de crescimento das raízes** em cada uma dessas condições.

- A pergunta é: O ‘tipo de esponja’ (ou a ausência dela) tem um efeito significativo na taxa média de crescimento das raízes de mangue?

TABLE 10.12 Analysis of variance table for experiment testing effects of sponges on growth of red mangrove roots

Source	Degrees of freedom (df)	Sum of squares (*SS*)	Mean square (*MS*)	F-ratio	*P*-value
Treatments	3	2.602	0.867	5.286	0.003
Residual	52	8.551	0.164		
Total	55	11.153			

TABLE 10.12 **Analysis of variance table for experiment testing effects of sponges on growth of red mangrove roots**

Source	Degrees of freedom (*df*)	Sum of squares (*SS*)	Mean square (*MS*)	F-ratio	*P*-value
Treatments	3	2.602	0.867	5.286	0.003
Residual	52	8.551	0.164		
Total	55	11.153			

TABLE 10.12 **Analysis of variance table for experiment testing effects of sponges on growth of red mangrove roots**

Source	Degrees of freedom (df)	Sum of squares (*SS*)	Mean square (*MS*)	F-ratio	*P*-value
Treatments	3	2.602	0.867	5.286	0.003
Residual	52	8.551	0.164		
Total	55	11.153			

Atividade: Decodificando a tabela de ANOVA

- Instruções: Em pequenos grupos, vocês realizarão um cálculo simplificado de ANOVA usando um conjunto de dados fictício que compara quatro grupos (A, B, C, D) com cinco observações cada

Grupo A	Grupo B	Grupo C	Grupo D
25	31	22	33
26	25	26	29
20	28	28	31
23	27	25	34
21	24	29	28
Média=23	Média=27	Média=26	Média=31

Atividade: Decodificando a tabela de ANOVA

- Tarefa: Preencham a seguinte tabela de ANOVA calculando cada componente. Usem as médias fornecidas e o guia de cálculo passo a passo do livro da Sonia Vieira.

Fonte de Variação	Soma de Quadrados (SQ)	Graus de Liberdade (GL)	Quadrado Médio (QM)	Estatística F
Entre Grupos				
Resíduo (Dentro dos Grupos)				
Total				

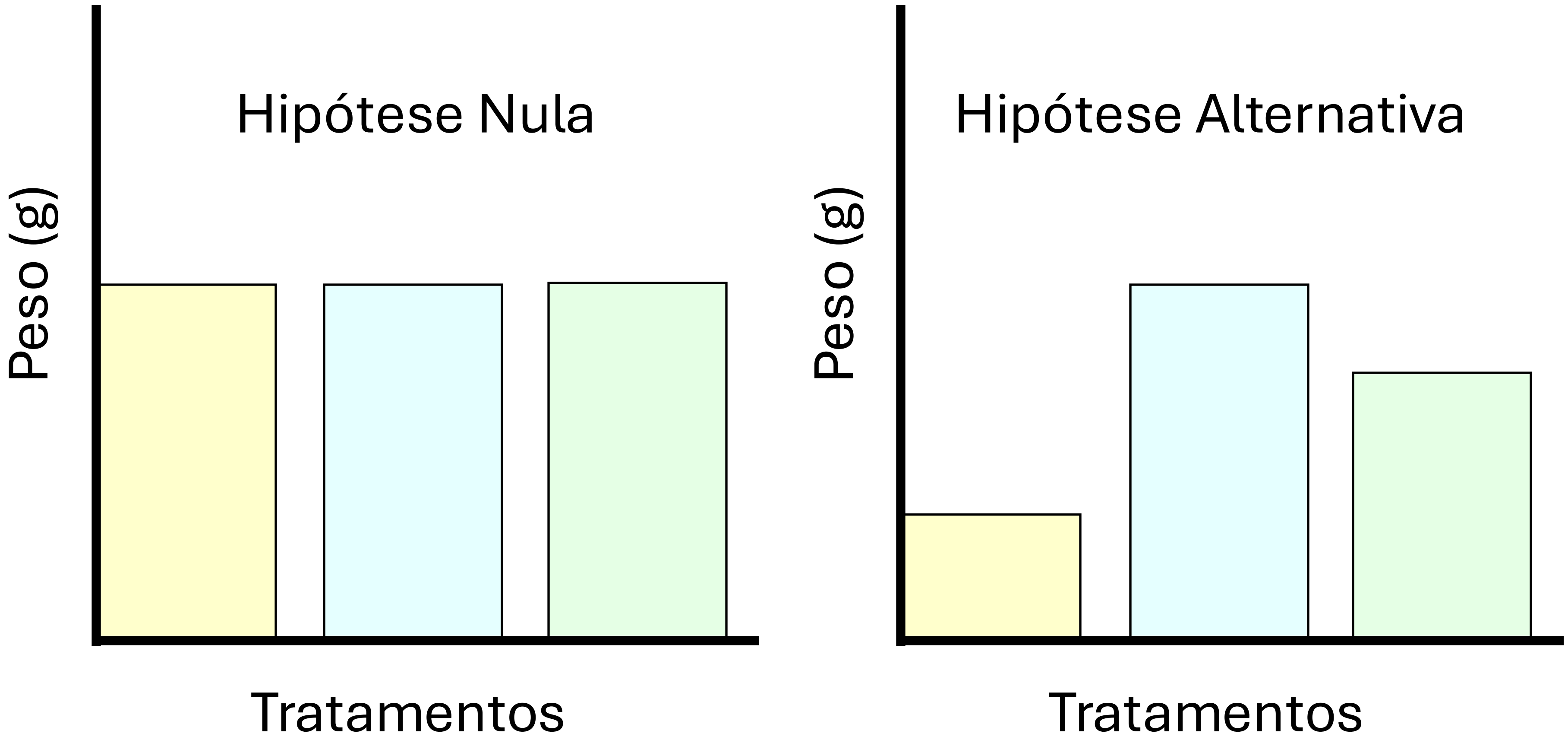
Atividade: Decodificando a tabela de ANOVA

- Graus de liberdade
  - De grupos: t-1
  - Do total: n-1
  - Do resíduo: (n-1)-(t-1)
- Soma dos Quadrados total
- Soma dos Quadrados grupos
- Soma dos Quadrados resíduos
- Quadrado médio dos grupos
- Quadrado médio dos resíduos
- F

Interpretando os resultados: teste de hipótese

- **Hipótese Nula** (H_0): As médias de todos os grupos são iguais ($\mu_A = \mu_B = \mu_C = \mu_D$).
- **Hipótese Alternativa** (H_1): *Pelo menos uma* média de grupo é diferente das outras.
- **O valor-p**: Assim como em outros testes, o valor-p informa a probabilidade de observar uma estatística F tão grande quanto a que você obteve, se a hipótese nula fosse verdadeira. Um valor-p pequeno (tipicamente ≤ 0,05) fornece evidências para rejeitar a hipótese nula.
  - No exemplo da atividade, a estatística F de 7,80 resulta em um valor-p de 0,002. Como 0,002 < 0,05, nós rejeitamos a hipótese nula e concluímos que há uma diferença entre as médias dos quatro grupos.
- **Uma Limitação Crucial**: O teste F da ANOVA nos diz que as médias dos grupos não são todas iguais, mas não nos diz *quais* grupos específicos são diferentes uns dos outros. O Grupo D é diferente do A? Existe diferença entre B e C? Para descobrir, precisamos realizar testes post-hoc ("após o fato").

Hipóteses


Hipótese Nula
Peso (g)
Tratamentos
Hipótese Alternativa
Peso (g)
Tratamentos

ANOVA: modelo linear com preditor categórico

Resultados da ANOVA podem ser idênticos, mas a interpretação depende do padrão das médias dos grupos em particular

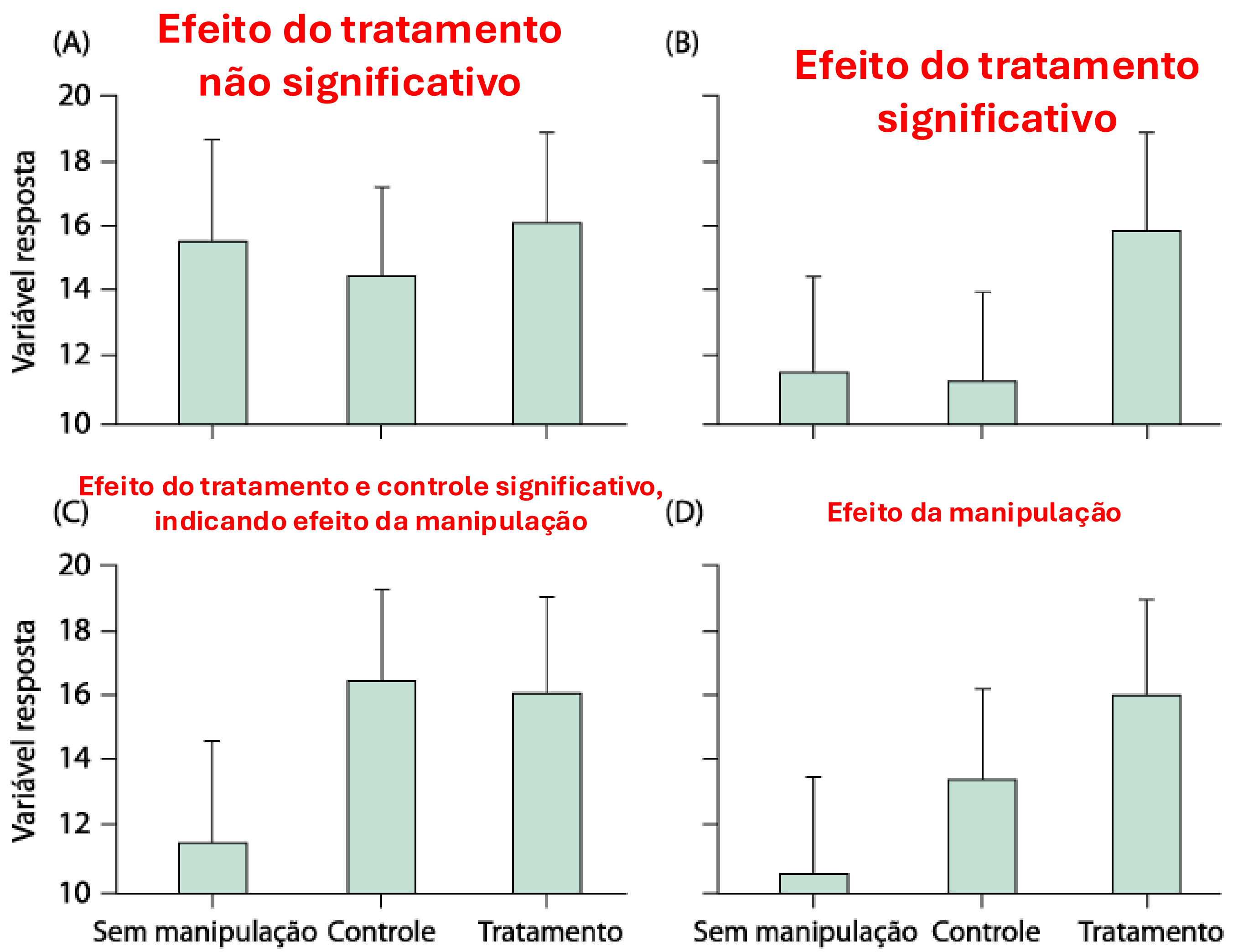


Figura 10.2 Possíveis resultados de um experimento hipotético com três tratamentos e uma

Gotelli & Ellison

Atividade: Interpretando uma ANOVA do Mundo Real (20 minutos)

- Vamos retornar ao estudo de crescimento de raízes de mangue por Ellison et al. (1996). Eles testaram os efeitos de um controle não manipulado, uma esponja falsa de espuma e duas espécies de esponjas vivas (*Haliclona* e *Tedania*) no crescimento das raízes.

- Os resultados da ANOVA estão na seguinte tabela

TABLE 10.12 Analysis of variance table for experiment testing effects of sponges on growth of red mangrove roots

Source	Degrees of freedom (df)	Sum of squares (*SS*)	Mean square (*MS*)	F-ratio	*P*-value
Treatments	3	2.602	0.867	5.286	0.003
Residual	52	8.551	0.164		
Total	55	11.153			

Atividade: Interpretando uma ANOVA do Mundo Real (20 minutos)

- As raízes não manipuladas tiveram um crescimento significativamente mais lento do que as raízes com qualquer uma das duas espécies de esponjas vivas (*Haliclona* e *Tedania*).
- Não houve diferença no crescimento entre as duas espécies de esponjas vivas, ou entre as esponjas vivas e o controle de espuma.

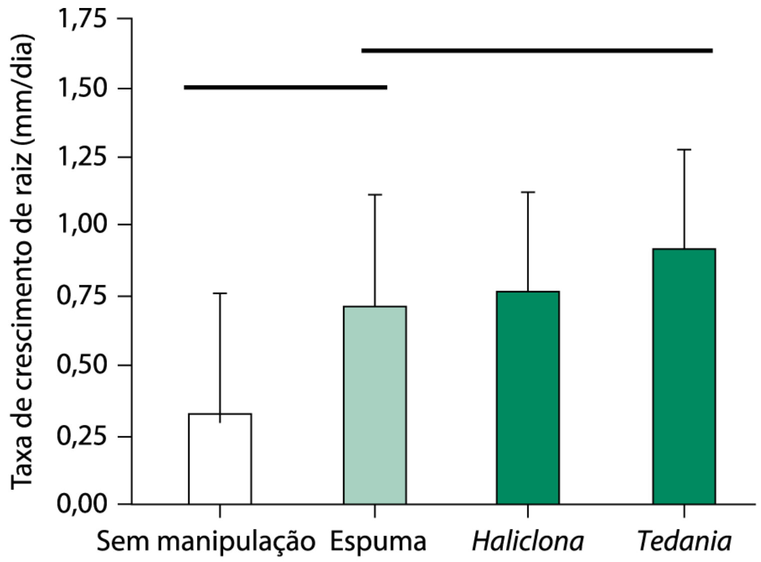


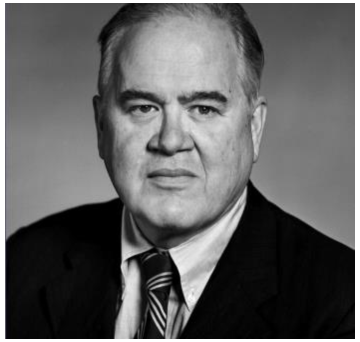
Figura 10.5 Taxa de crescimento médio (mm/dia) de raízes de mangue em quatro tratamentos experimentais. Os tratamentos foram estabelecidos (sem manipulação, espuma e dois tratamentos com esponjas vivas), com 14 réplicas em cada um. A altura da barra é a média da taxa de crescimento para o tratamento e as barras de erro verticais são de um desvio-padrão ao redor da média. As barras em verde-escuro indicam esponjas vivas, a verde-clara é a espuma inerte e a barra não sombreada é a média para as raízes sem manipulação. As linhas horizontais juntam os grupos tratamentos que não diferem significativamente pelo teste de DSH de Tukey (*ver* Tabela 10.13). (Dados de Ellison et al., 1996. As análises estatísticas desses dados estão presentes nas Tabelas 10.12 a 10.15.)

Atividade: Interpretando uma ANOVA do Mundo Real (20 minutos)

- Perguntas para Discussão em Grupo:
  - Olhando para a linha "Tipo de Esponja" na tabela ANOVA, qual é a sua conclusão geral sobre o efeito do tipo de esponja no crescimento das raízes de mangue? O que o valor-p significa neste contexto?
  - O que os resultados do teste post-hoc nos dizem sobre quais tipos de esponja são os mais importantes? Eles detalham a explicação fornecida pelo seu modelo geral.
  - Com base em todos os resultados, qual é a conclusão biológica que o ecólogo marinho pode tirar sobre o impacto da *Haliclona* sp. e da *Tedania* sp. no crescimento das raízes de mangue? Qual é o papel da esponja falsa neste estudo?

Testes Post-Hoc

- São testes que comparam cada par de grupos para descobrir onde estão as diferenças. Exemplos comuns incluem:
  - Teste HSD de Tukey (Diferença Honesta Significativa): Um dos testes mais utilizados para comparações par a par.



$$HSD = q\sqrt{\left(\frac{1}{n_i}+\frac{1}{n_j}\right)MS_{residual}}$$

q is the value from a statistical table of the studentized range distribution

  - Teste de Bonferroni: Outro método que ajusta o valor-p para cada comparação para controlar a taxa de erro geral do experimento

Testes Post-Hoc como Análise dos Parâmetros do Modelo

- Se a comparação de modelos (o teste F) nos diz que o nosso preditor categórico é importante ($P \leq 0{,}05$), o próximo passo é investigar como ele é importante.
- Os testes post-hoc não são um passo separado; eles examinam os parâmetros do nosso modelo ajustado.
- No Modelo de Efeitos, testamos se os α_i são diferentes uns dos outros.
- De forma mais direta, testamos hipóteses sobre as diferenças entre as médias dos grupos, como $H_0{:}\mu_A - \mu_B = 0$.

Testes Post-Hoc como Análise dos Parâmetros do Modelo

- Testes como o HSD de Tukey fazem exatamente isso: eles calculam intervalos de confiança para as diferenças entre cada par de médias.

- A diferença crucial é que eles usam o QM_{Res} do modelo geral (uma estimativa de variância mais robusta, pois usa todos os dados) e ajustam os níveis de significância para o problema de comparações múltiplas

Comentário sobre ANOVA Fatorial

Breve Introdução à ANOVA Fatorial

- Às vezes, os biólogos querem investigar o efeito de dois fatores ao mesmo tempo. Por exemplo:
  - Fator 1 (Linha): Sexo (Feminino vs. Masculino)
  - Fator 2 (Coluna): Faixa Etária (18-29, 30-49, 50-80)
- Isso requer uma ANOVA de Dois Fatores (Two-Way ANOVA). Uma característica fundamental desta análise é que ela nos permite testar um efeito de interação.
- **Efeito de Interação**: Uma interação ocorre se o efeito de um fator muda dependendo do nível do outro fator.

Breve Introdução à ANOVA Fatorial

- Efeito de **Interação**: Uma interação ocorre se o efeito de um fator muda dependendo do nível do outro fator.
- Exemplo: Se um fertilizante aumenta o crescimento em condições de sol, mas diminui o crescimento em condições de sombra, há uma interação entre fertilizante e luz.
- **O procedimento para uma ANOVA de dois fatores é testar primeiro a interação. Se ela for significativa, você não pode interpretar os efeitos principais de cada fator de forma independente**

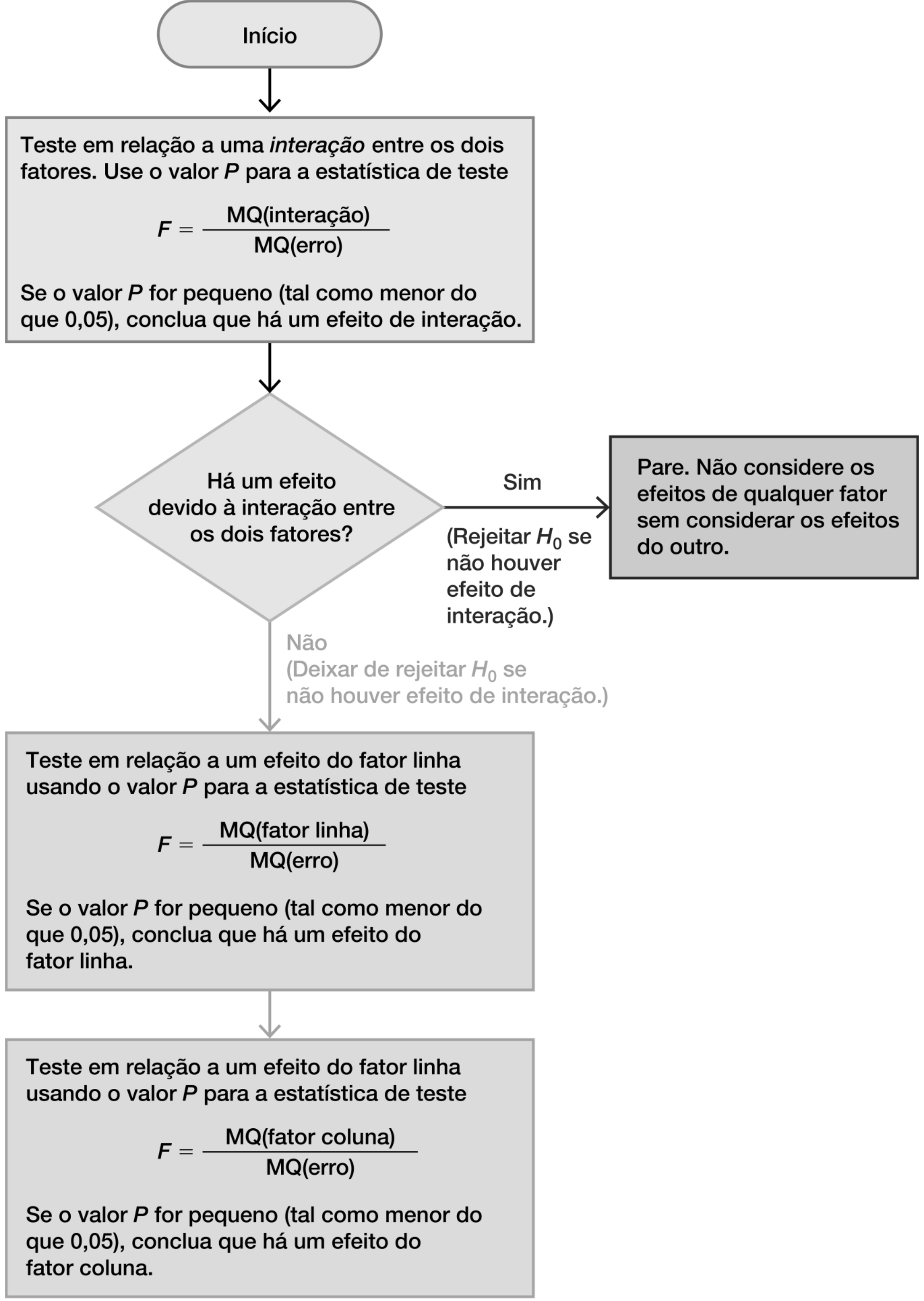


FIGURA 12.4 **Procedimento para a Análise de Variância de Dois Fatores.**

Triola

Breve Introdução à ANOVA Fatorial

- Visualizando Interações: Gráficos de interação são usados para visualizar isso.
  - Linhas paralelas sugerem que não há interação.
  - Linhas não paralelas ou que se cruzam sugerem um possível efeito de interação.

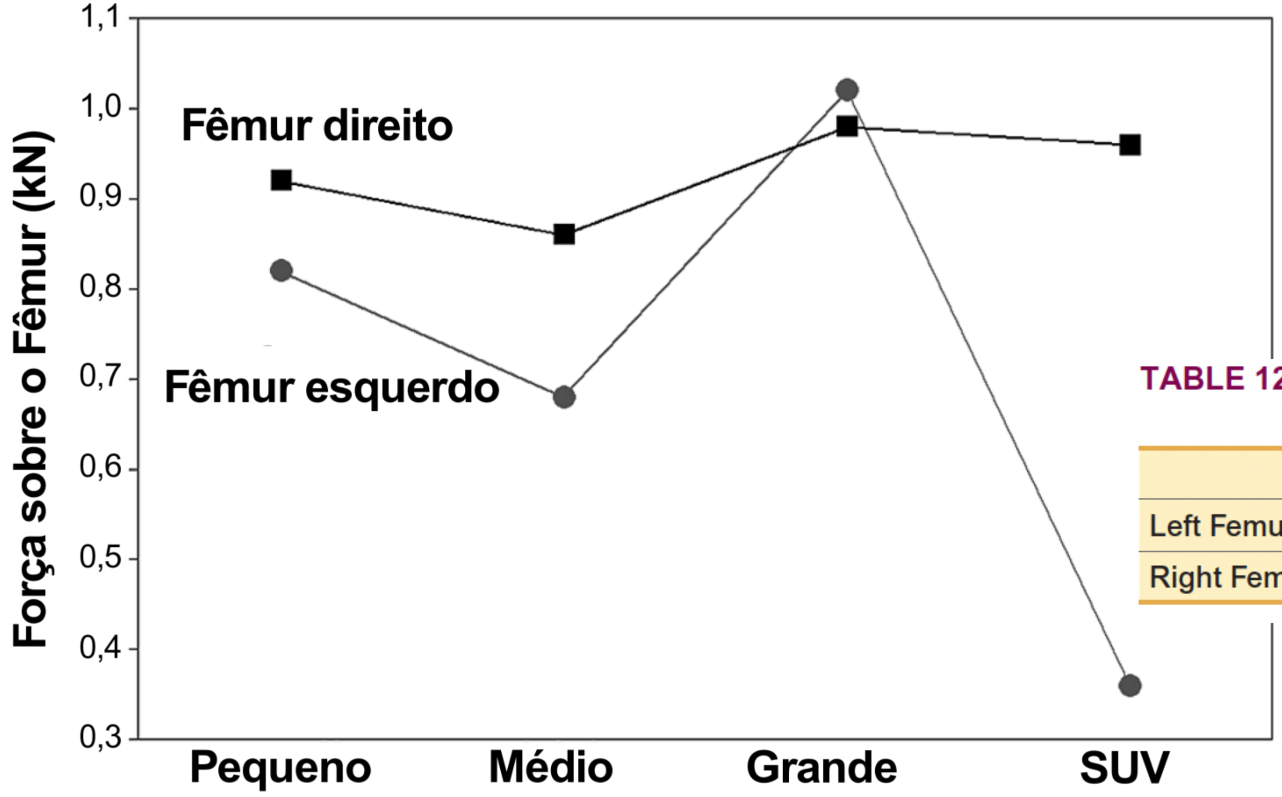


FIGURA 12.3 Gráfico de Interação de Lado do Fêmur e Categoria de Tamanho de Veículo da Tabela 12.4.

TABLE 12-3 Crash Test Force on Femur with Two Factors: Femur Side and Vehicle Size Category

	Small	Midsize	Large	SUV
Left Femur	1.6 1.4 0.5 0.2 0.4	0.4 0.7 1.1 0.7 0.5	0.6 1.8 0.3 1.3 1.1	0.4 0.4 0.6 0.2 0.2
Right Femur	2.8 1.0 0.3 0.3 0.2	0.6 0.8 1.3 0.5 1.1	1.5 1.7 0.2 0.6 0.9	0.7 0.7 3.0 0.2 0.2

TABLE 12-4 Means of Cells from Table 12-3

	Small	Midsize	Large	SUV
Left Femur	0.82	0.68	1.02	0.36
Right Femur	0.92	0.86	0.98	0.96

Triola

Efeito do tratamento e da covariável não significativo

Efeito da covariável significativo, mas do tratamento não significativo

Efeito do tratamento significativo, mas da covariável não significativo

Efeito do tratamento e covariável significativos

Interação tratamento e covariável em que a ordem dos níveis do tratamento difere

Interação tratamento e covariável em que a ordem não difere

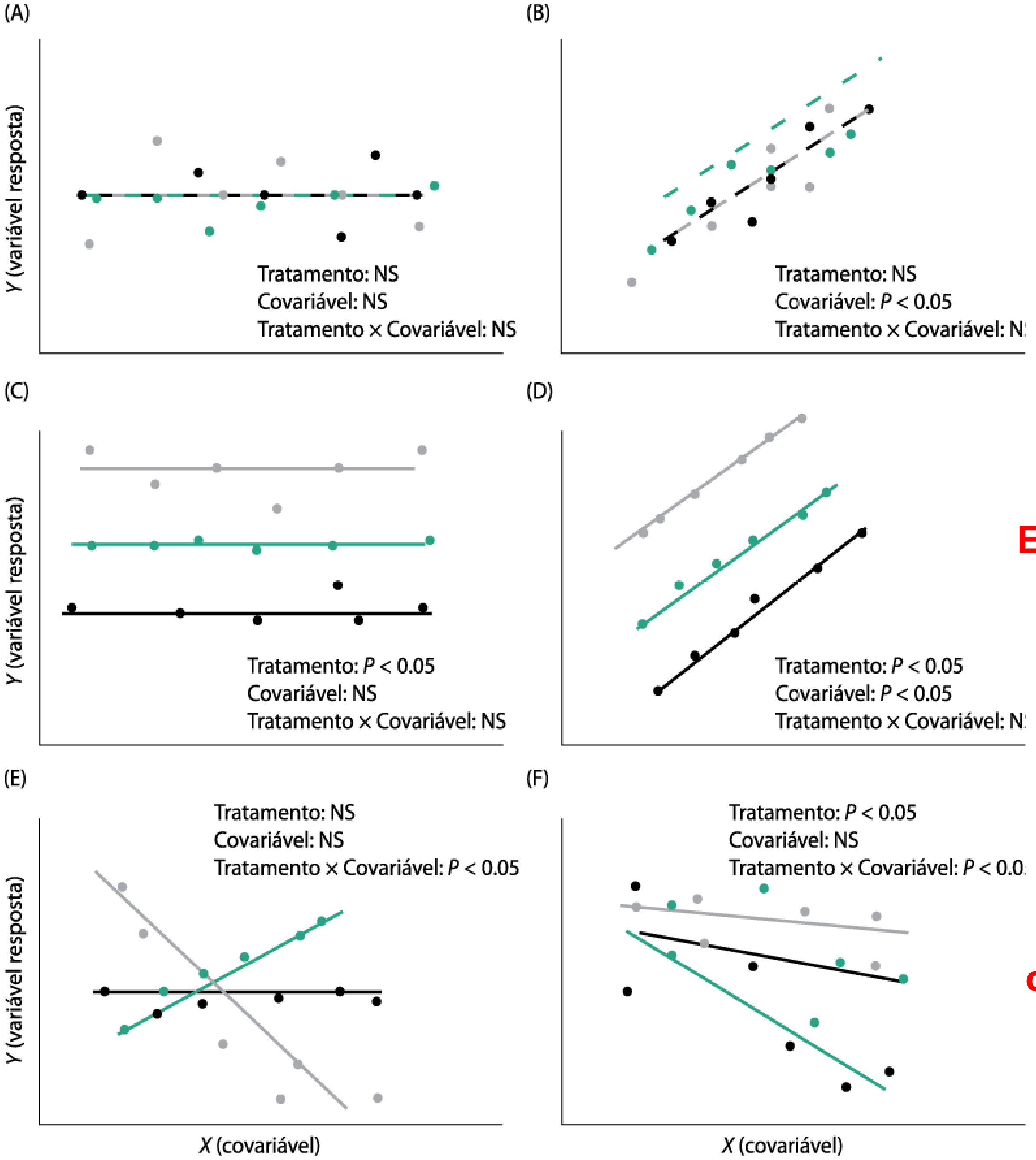


Fig 10.4 Gotelli

ANOVA de Dois Fatores como Modelo Linear

- A mesma lógica se aplica quando temos mais de um preditor categórico. Um modelo com dois fatores pode ser escrito como:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

- Onde:
  - α_ié o efeito do fator A (ex: Sexo).
  - β_j é o efeito do fator B (ex: Faixa Etária).
  - $(\alpha\beta)_{ij}$ é o termo de interação. Ele representa o efeito que é único à combinação específica dos níveis dos fatores *i* e *j*

ANOVA de Dois Fatores como Modelo Linear

- O teste de hipóteses na ANOVA de dois fatores é uma série de comparações de modelos:
  1. Primeiro, comparamos o modelo completo (com interação) a um modelo sem a interação para ver se a interação é significativa.
  2. Se não for, comparamos o modelo sem interação a modelos que removem cada fator principal para testar seus efeitos.

Recaptulação

- A ANOVA é uma ferramenta fundamental para comparar as médias de 3 ou mais grupos.
- Ela funciona *particionando a variância* em componentes sistemáticos (entre grupos) e aleatórios (dentro dos grupos).
- Um resultado F significativo indica que pelo menos uma média de grupo é diferente e deve ser seguido por testes post-hoc para identificar quais grupos diferem.
- Lembre-se sempre de verificar as premissas do teste e visualizar seus dados para interpretar corretamente seus resultados.

Recaptulação

- A ANOVA nada mais é do que ajustar um modelo linear com um preditor categórico.
- O teste F é uma comparação que pergunta: "O meu modelo, que inclui as informações dos grupos, explica mais variação do que um modelo que usa apenas a média geral?"
- Testes post-hoc são uma maneira de investigar os parâmetros do nosso modelo ajustado para entender onde estão as diferenças significativas.
- Essa abordagem unifica a ANOVA com a regressão, fornecendo uma base conceitual mais forte para entender a análise de dados em biologia.