Planejamento de Experimentos
Experimento com um fator

Enrico A. Colosimo/UFMG

Depto. Estatística - ICEx - UFMG

- **Pergunta Clínica Típica**: comparar **dois ou mais grupos** experimentais com relação a uma resposta quantitativa ou, em outras palavras, comparar níveis (> 2) de um fator.
- **Exemplos:**
  - Comparar três drogas para o tratamento de fibrose cística em que o desfecho é uma medida quantitativa de capacidade pulmonar.
  - Comparar a **idade** de pacientes entre três grupos de risco (baixo, médio, alto).
- Comparação de dois grupos: testes t ou Mann-Whitney.

Exemplo

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Prednisone, Azathioprine, and N-Acetylcysteine for Pulmonary Fibrosis

The Idiopathic Pulmonary Fibrosis Clinical Research Network*

ABSTRACT

BACKGROUND

A combination of prednisone, azathioprine, and N-acetylcysteine (NAC) has been widely used as a treatment for idiopathic pulmonary fibrosis. The safety and efficacy of this three-drug regimen is unknown.

METHODS

In this randomized, double-blind, placebo-controlled trial, we assigned patients with idiopathic pulmonary fibrosis who had mild-to-moderate lung-function impairment to one of three groups — receiving a combination of prednisone, azathioprine, and NAC (combination therapy), NAC alone, or placebo — in a 1:1:1 ratio. The primary outcome was the change in longitudinal measurements of forced vital capacity during a 60-week treatment period.

N ENGL J MED 366;21 NEJM.ORG MAY 24, 2012

STATISTICAL ANALYSIS

The trial was designed with a two-step procedure to control the experiment-wise error rate at the 0.05 level. The first step was based on an overall test of 2 degrees of freedom. If any difference between study groups was statistically significant at the 0.05 level, then each of the three pairwise comparisons would be tested at the 0.05 level.

Introdução

- A primeira vista, pode parecer correto realizar vários testes t entre os grupos, comparando-os *dois a dois*.
- No caso da comparação de três grupos (grupo A, grupo B e grupo C), temos **três testes** t de comparação entre médias: μ_A *vs* μ_B, μ_A *vs* μ_C e μ_B *vs* μ_C.
- Na comparação de quatro grupos, temos **seis testes** t de comparação entre médias.
- Se o número de grupos é igual a 10, precisaríamos de **45 testes** t dois a dois.

- **Observação 1:** O número de testes aumenta conforme o número de grupos aumenta. Para k grupos temos $\binom{k}{2}$ comparações.
- **Observação 2:** Tal procedimento (a realização de todas as comparações dois a dois) é estatisticamente incorreto.
  - O teste t foi proposto para, **em um mesmo experimento**, comparar-se uma média A com apenas outra, B, com probabilidade fixa do erro tipo I (α=0,05).
  - Se forem feitas mais de uma comparação envolvendo a média A, a probabilidade do erro tipo I (ou erro experimental conjunto) **passa a ser maior do que** α.

- O procedimento mais indicado para se evitar esse aumento no nível global de significância do experimento consiste em utilizar a técnica da **Análise de Variância** (ANOVA) que consiste nos seguintes passos:

    - PASSO 1 - Comparar todas as médias em um único teste. O objetivo inicial é identificar a existência de **ao menos uma diferença** entre grupos.

    - PASSO 2 - Caso o resultado anterior for significativo, aplica-se um ou mais métodos de **comparações múltiplas**. O objetivo é identificar quais as médias são diferentes, controlando o nível global de significância.

Exemplo típico: Resistência de uma nova fibra sintética para camisas de homens. (Montgomery, 1997, p.63)

- Um engenheiro de produto está interessado em investigar a resistência de uma nova fibra sintética.
- A resistência é afetada pela porcentagem de algodão usada na produção do tecido.
- O engenheiro decide testar o novo tecido com diferentes porcentagens de algodão: 15, 20, 25, 30 e 35 %.
- A resposta é a resistência da fibra medida em libra/polegada^2.
- O experimento usou 5 réplicas para cada nível do fator.

Exemplo típico: Resistência de uma nova fibra sintética para camisas de homens (Montgomery, 1997, p.63)

- Como realizar a aleatorização se temos somente um processo para produzir a fibra?
- Necessário aleatorizar a ordem/sequência de produção/medida da fibra.
- A aleatorização evita possíveis efeitos de variáveis de perturbação. Por exemplo: aquecimento, desgaste, aprendizado, etc.

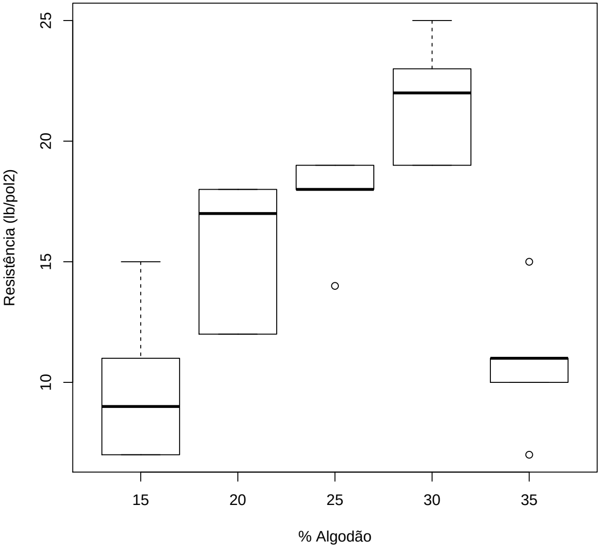
Exemplo: Resistência da fibra sintética

Tabela: Resistência da fibra sintética.(em libra/polegada^2)

Réplicas	**Porcentagem de Algodão**				
	15	20	25	30	35
1	7	12	14	19	7
2	7	17	18	25	10
3	15	12	18	22	11
4	11	18	19	19	15
5	9	18	19	23	11
	7	12	14	19	7
Média	9,8	15,4	17,6	21,6	10,8
DP	3,3	3,1	2,1	2,6	2,9

Média e desvio-padrão para o conjunto de dados são: 15,0 e 5,2.

Boxplot



ANOVA - Fontes de variação

- A ANOVA é baseada em estimativas de dispersão/variância.
- Neste caso, existem duas diferentes fontes de variação:
  - Variação natural ou intra-grupo: valores indivíduais em torno das médias populacionais (desvio-padrão intra-grupo ou do erro);
  - Variação entre-grupos: médias populacionais em torno da média global (desvio-padrão entre-grupos ou entre-tratamentos).
- Se a variabilidade intra populações é menor que a variabilidade entre grupos, sugere que as médias populacionais são de fato diferentes.

Desvio-padrão Intra-Grupo

- Objetivo: testar a seguinte hipótese:

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_k \text{ ou } H_0 : \tau_1 = \tau_2 = \ldots = \tau_k = 0$$

para as médias de k populações.

- Estimativa do desvio-padrão Intra-Grupo (σ):

$$s_E^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_k - 1)s_k^2}{N - k},$$

em que, s_i e n_i são o desvio-padrão e tamanho de amostra do i-ésimo grupo e $N = n_1 + \ldots + n_k$ é o tamanho amostra total.

- Esta quantidade é, simplesmente, a média ponderada das k variâncias amostrais. O subscrito E se refere a variabilidade "dentro de grupos" ou equivalente a s_c^2.

Desvio-padrão Entre-Grupos

- Necessitamos de uma estimativa da variação das médias em torno da média global.

- Estimativa do desvio-padrão Entre-Grupos (σ_T):

$$s_T^2 = \frac{n_1(\bar{y}_{1.} - \bar{y}_{..})^2 + n_2(\bar{y}_{2.} - \bar{y}_{..})^2 + \ldots + n_k(\bar{y}_{k.} - \bar{y}_{..})^2}{k-1},$$

  em que $\bar{y}_{i.}$ é a média amostral do i-ésimo grupo e $\bar{y}_{..}$ é a média global das N observações

- Se a hipótese nula é verdadeira, esta quantidade também estima a variância intra-grupo, σ^2.

Desvio-padrão Entre-Grupos e o Teste t ($k = 2$)

- O mesmo raciocínio vale para o teste t, em que $k = 2$.
- Estimativa do desvio-padrão Entre-Grupos (σ_T):

$$s_T^2 = \frac{n_1(\bar{y}_{1.} - \bar{y}_{..})^2 + n_2(\bar{y}_{2.} - \bar{y}_{..})^2}{2-1} = \frac{(\bar{y}_{1.} - \bar{y}_{2.})^2}{1/n_1 + 1/n_2}.$$

- Ou seja, podemos escrever que:

$$t = s_T / s_E.$$

Testa F

- A pergunta clínica pode ser traduzida pela seguinte questão: as médias amostrais variam em torno da média global mais do que as observações individuais variam em torno das médias amostrais?

- Em caso positivo, isto inidica que existe alguma diferença entre as médias populacionais.

- Precisamos de um estatístca para avaliar o tamanho desta diferença. A estatística F é usada para este propósito:

$$F = \frac{s_T^2}{s_E^2}.$$

Teste F

- Sob a hipótese nula, que as médias são iguais, tanto s_T^2 quanto s_E^2 estimam a variância comum σ^2, e F deve estar próximo de 1.
- Se existe uma diferença entre as populações, então a variância entre os grupos é maior que a variância dentro dos grupos, e F é maior que 1.
- Sob H_0, a razão F tem uma distribuição F com $k-1$ e $N-k$ graus de liberdade.
- Obs. Este resultado mostra que $t_r^2 = F_{1,r}$.

Distribuição F

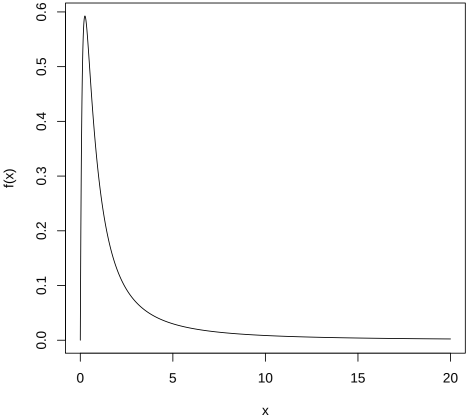
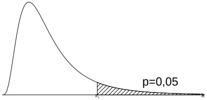


Figura: Distribuição F com 4 e 2 graus de liberdade.

Distribuição de Snedecor a 5% (p=0,05)



	1	2	3	4	5	6	7	8	9	10	12	14	15	16	18	20	30	40	60	120
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,41	19,42	19,43	19,43	19,44	19,45	19,46	19,47	19,48	19,49
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,71	8,70	8,69	8,67	8,66	8,62	8,59	8,57	8,55
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,87	5,86	5,84	5,82	5,80	5,75	5,72	5,69	5,66
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,64	4,62	4,60	4,58	4,56	4,50	4,46	4,43	4,40
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,96	3,94	3,92	3,90	3,87	3,81	3,77	3,74	3,70
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,53	3,51	3,49	3,47	3,44	3,38	3,34	3,30	3,27
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,24	3,22	3,20	3,17	3,15	3,08	3,04	3,01	2,97
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,03	3,01	2,99	2,96	2,94	2,86	2,83	2,79	2,75
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,86	2,85	2,83	2,80	2,77	2,70	2,66	2,62	2,58
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,74	2,72	2,70	2,67	2,65	2,57	2,53	2,49	2,45
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,64	2,62	2,60	2,57	2,54	2,47	2,43	2,38	2,34
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,55	2,53	2,51	2,48	2,46	2,38	2,34	2,30	2,25
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,48	2,46	2,44	2,41	2,39	2,31	2,27	2,22	2,18
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,42	2,40	2,38	2,35	2,33	2,25	2,20	2,16	2,11
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,37	2,35	2,33	2,30	2,28	2,19	2,15	2,11	2,06
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,33	2,31	2,29	2,26	2,23	2,15	2,10	2,06	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,29	2,27	2,25	2,22	2,19	2,11	2,06	2,02	1,97
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,26	2,23	2,21	2,18	2,16	2,07	2,03	1,98	1,93
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,22	2,20	2,18	2,15	2,12	2,04	1,99	1,95	1,90
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,20	2,18	2,16	2,12	2,10	2,01	1,96	1,92	1,87
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,17	2,15	2,13	2,10	2,07	1,98	1,94	1,89	1,84
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,15	2,13	2,11	2,08	2,05	1,96	1,91	1,86	1,81
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,13	2,11	2,09	2,05	2,03	1,94	1,89	1,84	1,79
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,11	2,09	2,07	2,04	2,01	1,92	1,87	1,82	1,77
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,09	2,07	2,05	2,02	1,99	1,90	1,85	1,80	1,75
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,08	2,06	2,04	2,00	1,97	1,88	1,84	1,79	1,73
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,06	2,04	2,02	1,99	1,96	1,87	1,82	1,77	1,71
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,05	2,03	2,01	1,97	1,94	1,85	1,81	1,75	1,70
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,04	2,01	1,99	1,96	1,93	1,84	1,79	1,74	1,68
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,95	1,92	1,90	1,87	1,84	1,74	1,69	1,64	1,58
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,86	1,84	1,82	1,78	1,75	1,65	1,59	1,53	1,47
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,78	1,75	1,73	1,69	1,66	1,55	1,50	1,43	1,35

Tabela 5: Quantis da Distribuição para probabilidade . Graus de liberdade do numerador dado no topo e do denominador na margem esquerda.

Fontes de variação - Tabela ANOVA

- **Observação:** Podemos organizar o teste F na seguinte tabela, chamada de tabela de Análise de Variância:

Tabela: Tabela da Análise de Variância (ANOVA).

Fontes de variação	SQ	GL	QM	F
Entre os grupos	SQ_T	$k-1$	$QM_T = SQ_T\ /k-1$	$QM_T\ /\ QM_E$
Dentro dos grupos	SQ_E	$N-k$	$QM_E = SQ_E\ /N-k$	-
Total	SQ_{Total}	$N-1$	-	-

em que:

- SQ_T é a "soma de quadrados" entre os grupos e é o numerador de s_T^2, ou seja, $SQ_T = n_1(\bar{y}_{1.} - \bar{y}_{..})^2 + \ldots + n_k(\bar{y}_{k.} - \bar{y}_{..})^2$.
- SQ_E é a "soma de quadrados" dentro dos grupos e é o numerador de s_E^2, ou seja, $SQ_E = (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \ldots + (n_k - 1)s_k^2$.
- Note que $QM_T = s_T^2$ e $QM_E = s_E^2$.

Exemplo Resistência Fibra - Fontes de variação

- Retomando ao exemplo, estamos interessados em testar se a média da variável resistência é igual para os cinco diferentes porcentagens de algodão. Ou seja,

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \quad H_0: \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = 0.$$

- Para começar, calculamos a estimativa da variância dentro dos grupos

$$\begin{aligned} s_E^2 &= \frac{(n_1 - 1)s_1^2 + \ldots + (n_5 - 1)s_5^2}{n_1 + n_2 + n_3 + n_4 + n_5 - 5} \\ &= \frac{(5-1)(3,3)^2 + \ldots + (5-1)(2,9)^2}{25-5} \\ &= 8,06. \end{aligned}$$

Fontes de variação

$$
\begin{aligned}
s_T^2 &= \frac{n_1(\bar{y}_{1.} - \bar{y}_{..})^2 + \ldots + n_5(\bar{y}_{5.} - \bar{y}_{..})^2}{5-1} \\
&= \frac{5(9,8-15,0)^2 + \ldots + 5(10,8-15,0)^2}{5-1} \\
&= 118,95.
\end{aligned}
$$

Fontes de variação

- Desta forma, a estatística de teste F é

$$\begin{aligned} F &= \frac{s_T^2}{s_E^2} \\ &= \frac{118,95}{8,06} \\ &= 14,76. \end{aligned}$$

- Comparar com a distribuição F com $k - 1 = 5 - 1 = 4$ e $N - k = 25 - 5 = 20$ graus de liberdade, o $valor - p < 0.001$.

- Rejeitamos a hipótese nula ao nível de 5% de significância.

- Existe alguma diferença entre as médias dos valores do resistência entre as 5 porcentagens de algodão.

Fontes de variação

- De forma análoga, temos a seguinte tabela de análise de variância.

Tabela: Tabela da Análise de Variância - ANOVA.

Fontes de variação	SQ	GL	QM	*F*	*p-valor*
Entre os grupos	475,8	4	$118,94$	14,765	< 0.001
Dentro dos grupos	161.2	20	$8,06$	-	-
Total	637,0	24	-	-	-

Modelo Estatístico para Efeitos Fixos

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

em que Y_{ij} é o valor da resposta da j-ésima observação ($j = 1, \ldots, n$) no i-ésimo nível do fator ($i = 1, \ldots, k$); μ: efeito geral da média: τ_i: efeito do i-ésimo tratamento.