

## Mais detalhes

### SA15: Comparando Subpopulações

**Tópicos:** Gráficos e tabelas, Inferência Estatística.

**Recursos:** tabela fornecida com dados sobre altura e gênero.

**Nível de ensino:** Médio e Superior.

**Duração:** 2 horas-aula.

Em geral, coleta-se informações de uma população através de amostras. Uma das formas, chamada *amostra aleatória simples com reposição*, atribui mesma possibilidade de presença na amostra para todos os elementos da população. Ela é a mais simples, teoricamente falando, mas nem sempre a mais prática. Se for possível enumerar os elementos da população, pode-se usar algum mecanismo que escolha os números sem preferência, repondo o número sorteado após cada escolha. Se for difícil, ou impossível, busca-se alguma alternativa que imite esse tipo de escolha.

#### 1. Amostragem

A tabela *Estudantes* (ver ao final do texto) contém altura e gênero de estudantes da Licenciatura do IME-USP. Imprima cópias da tabela de dados para distribuir na classe. Solicite que cada grupo imite (ver observação ao final do texto) a coleta de uma *amostra aleatória simples com reposição*, obtendo 30 valores da variável altura em cada subpopulação de gênero. O professor pode escolher um tamanho de amostra menor (10 ou 20) se houver restrições ao tempo disponível em sala de aula. Cada grupo preenche uma tabela como a apresentada abaixo:

Observações	01	02	03	...	...	...	30
Masculino							
Feminino							

Cada grupo vai usar sua amostragem para comparar as alturas nas duas subpopulações. A comparação será feita usando a "caixa" de um gráfico denominado *box-plot*. O gráfico usa as medidas dos quartis, sendo o 1º quartil ( $Q_1$ ) o limite para as 25% menores observações. De modo análogo, a mediana ou 2º quartil ( $md_{obs}$ ) é o limite para 50% e o 3º quartil ( $Q_3$ ) para as 75% menores. A "caixa" do *box-plot* é apresentada a seguir:



Cada grupo deve escrever os valores coletados em ordem crescente e anotar em uma tabela do tipo:

Observações Ordenadas	1º	2º	3º	...	...	...	30º
Masculino							
Feminino							

Em termos gerais, para se obter os quartis, primeiro obtém-se a mediana (existem regras diferentes para número par ou ímpar de observações), separando o conjunto original de dados em duas partes, inferior e superior. Calcule as medianas de cada parte para obter os quartis (para detalhes, consulte a página 19 do livro Magalhães, M. N. e Lima, A. C. P. (2015): *Noções de Probabilidade e Estatística*, 7ª ed., revista, 3ª reimpressão. São Paulo: EDUSP).

No caso de 30 observações, os quartis seriam:  $Q_1 = 8^{\text{a}}$  observação ordenada;  $md_{\text{obs}} =$  média entre a 15ª e 16ª observações ordenadas e  $Q_3 = 23^{\text{a}}$  observação ordenada.

Cada grupo desenha as “caixas” correspondentes às subpopulações e compara os resultados.

## 2. Discussão

Para facilitar a discussão geral, os grupos desenharam na lousa as suas “caixas”. Note que os grupos imitaram uma *amostra aleatória simples com reposição*. Cabe discutir as diversas imitações utilizadas, comentando suas eventuais deficiências. Algumas perguntas para orientar a discussão:

- a) Os gráficos são todos iguais?
- b) Por que há diferenças entre eles?
- c) Será que houve uma falha no procedimento que usamos?
- d) A escolha deu preferência a alguns valores?
- e) Ela foi mesmo aleatória?

É importante salientar aos estudantes que, em geral, a amostragem envolve aleatoriedade. Em fenômenos aleatórios, a repetição pode produzir respostas diferentes, dependendo da probabilidade de ocorrência associada a cada valor da população.

Para concluir indique que sempre em processos aleatórios, mesmo se tivesse sido feita a coleta de uma *amostra aleatória simples com reposição*, existiriam os chamados *erros amostrais*. Esses erros dependem da variabilidade dos dados e diminuem à medida que se aumenta o tamanho da amostra (para aprofundar essa discussão consulte a referência acima).

**Observação:** *Uma forma de gerar amostras aleatórias simples com reposição seria usar um aplicativo no Smartphone. O Random Number Generator, encontrado gratuitamente no Google Play, faz sorteios de números, supostamente, com igual probabilidade.*

## 3. Tabela Estudantes

Na próxima página apresentamos uma tabela com dados, das variáveis Gênero e Altura, de 212 estudantes da Licenciatura em Matemática do IME-USP.

Tabela *Estudantes* (estudantes da Licenciatura- IME-USP- 2012/2013)

Aluno	Gênero	Alt. (m)
1	M	1,89
2	F	1,61
3	M	1,80
4	M	1,70
5	M	1,74
6	M	1,75
7	M	1,68
8	F	1,62
9	F	1,68
10	M	1,65
11	M	1,67
12	M	1,96
13	F	1,68
14	M	1,80
15	M	1,76
16	F	1,64
17	M	1,79
18	F	1,70
19	M	1,65
20	M	1,60
21	M	1,68
22	M	1,94
23	M	1,82
24	M	1,90
25	M	1,76
26	M	1,88
27	M	1,83
28	M	1,71
29	F	1,65
30	F	1,63
31	M	1,72
32	M	1,81
33	M	1,78
34	M	1,74
35	M	1,80
36	M	1,75
37	F	1,65
38	F	1,58
39	M	1,70
40	M	1,65
41	M	1,85
42	M	1,77
43	M	1,90
44	M	1,81
45	F	1,65
46	M	1,75
47	F	1,67
48	M	1,78
49	M	1,78
50	M	1,75
51	M	1,67
52	M	1,86
53	M	1,82
54	M	1,83
55	M	1,77
56	F	1,80
57	F	1,65
58	M	1,79
59	F	1,59
60	M	1,72
61	M	1,71
62	M	1,85
63	M	1,90
64	M	1,77
65	M	1,69
66	M	1,81
67	F	1,58
68	F	1,59
69	F	1,57
70	M	1,70
71	M	1,81

Aluno	Gênero	Alt. (m)
72	M	1,78
73	F	1,55
74	M	1,55
75	F	1,57
76	F	1,63
77	F	1,70
78	M	1,65
79	F	1,64
80	M	1,88
81	M	1,85
82	M	1,70
83	M	1,65
84	M	1,73
85	F	1,65
86	F	1,69
87	M	1,79
88	F	1,70
89	F	1,78
90	M	1,90
91	M	1,71
92	M	1,75
93	F	1,65
94	F	1,62
95	F	1,68
96	F	1,77
97	M	1,84
98	M	1,87
99	M	1,85
100	M	1,90
101	M	1,78
102	M	1,81
103	M	1,82
104	F	1,80
105	F	1,65
106	M	1,88
107	F	1,57
108	F	1,58
109	M	1,72
110	M	1,75
111	M	1,70
112	M	1,75
113	F	1,58
114	M	1,80
115	F	1,65
116	M	1,75
117	M	1,70
118	M	1,71
119	M	1,69
120	M	1,76
121	M	1,75
122	M	1,75
123	F	1,65
124	F	1,69
125	F	1,63
126	M	1,78
127	F	1,63
128	M	1,76
129	F	1,62
130	M	1,80
131	F	1,58
132	M	1,71
133	M	1,92
134	F	1,57
135	F	1,56
136	F	1,73
137	M	1,70
138	M	1,82
139	M	1,72
140	F	1,72
141	F	1,65
142	F	1,68

Aluno	Gênero	Alt. (m)
143	M	1,87
144	M	1,68
145	M	1,69
146	F	1,53
147	M	1,78
148	M	1,85
149	M	1,76
150	M	1,98
151	M	1,74
152	F	1,65
153	F	1,66
154	M	1,67
155	M	1,68
156	M	1,85
157	F	1,62
158	F	1,63
159	M	1,77
160	M	1,84
161	M	1,90
162	F	1,53
163	F	1,80
164	M	1,67
165	F	1,70
166	M	1,86
167	M	1,72
168	F	1,69
169	M	1,83
170	M	1,58
171	F	1,54
172	M	1,82
173	M	1,65
174	M	1,69
175	M	1,78
176	F	1,57
177	F	1,73
178	M	1,77
179	M	1,71
180	M	1,76
181	F	1,60
182	F	1,74
183	M	1,76
184	M	1,88
185	M	1,75
186	F	1,66
187	F	1,52
188	M	1,80
189	M	1,70
190	F	1,68
191	M	1,73
192	F	1,68
193	M	1,85
194	M	1,75
195	M	1,71
196	M	1,84
197	F	1,63
198	M	1,83
199	M	1,75
200	F	1,84
201	M	1,73
202	M	1,72
203	M	1,78
204	M	1,82
205	M	1,73
206	M	1,75
207	M	1,75
208	F	1,56
209	F	1,62
210	F	1,67
211	M	1,94
212	M	1,75