

**Escola Superior de Agricultura "Luiz de Queiroz" - Universidade de São Paulo**  
**LCE 0211 - Estatística Geral**

Prof. Idemauro Antonio Rodrigues de Lara

Estagiárias: Cristiane Rodrigues, Janaina Marques e Laura Torres

Trabalho Semestral de Estatística - individual  
Análise de Dados relativos a um Experimento com milho, ano base 2017.

Atividades em três módulos

**MÓDULO I - Análise descritivas**

1. Seleção da amostra aleatória com reposição, com  $n=50$ , usando como semente aleatória o número USP.
2. Com os dados de sua amostra aleatória, construir tabelas de frequências e gráficos apropriados para todas as variáveis (unidimensionais).
3. Calcular e interpretar as medidas descritivas de tendência central, dispersão, bem como assimetria e curtose para as variáveis quantitativas.
4. Construir o resumo dos cinco números e o "*Boxplot*" para as variáveis quantitativas contínuas. Identificar pontos atípicos caso existam e interpretar os resultados.
5. Análise conjunta: investigar se existe associação entre as variáveis: número de plantas, altura da planta, altura da espiga e produção. Apresente métodos descritivos adequados. Construir uma tabela de contingência e um gráfico apropriado para: Categoria planta x Categoria espiga.

**MÓDULO II-Trabalhando distribuições amostrais**

Considere o arquivo, medidas\_fase2\_csv, contendo as médias, variâncias, proporções e coeficientes de correlação encontra por cada um, com base no total de alunos que participam da atividade (turmas 1 e 4)

1. Selecione uma amostra aleatória de tamanho 50. Use o programa já executado na Fase 1.
- 2) Com base no seu arquivo, selecione uma variável quantitativa referente à média (altura planta, altura espiga, produção), uma variável quantitativa referente à variância (altura planta, altura espiga, produção) e uma variável quantitativa referente à proporção.

3) Com base no seu arquivo de médias, qual é a probabilidade de se encontrar uma média maior do que aquela que você observou na sua amostra?

4) Calcule a "média populacional", "variância populacional" e a "proporção populacional" das variáveis escolhidas mas com base no arquivo original (completo- Fase1). Agora calcule a média das estatísticas selecionadas (item 1 -Fase 2). Compare os valores.

4) Construa um histograma para apresentação de cada uma de suas variáveis. Para a variável "Variância", considere a seguinte transformação:

$$\frac{(n - 1)S^2}{\sigma^2}$$

em que  $S^2$  corresponderá a cada uma das variâncias de seu arquivo e  $\sigma^2$  é a variância populacional calculada no item 4 (Crie este vetor de dados).

Comente sobre a distribuição das estatísticas. Há sugestão de algum modelo probabilístico para estas distribuições? Os valores populacionais estão contemplados na distribuição apresentada pelos respectivos histogramas?

### MÓDULO III-INFERÊNCIAS

1. Construir intervalos de 95% de confiança para a média referente às variáveis: altura da planta, altura da espiga e produção. Interpretar os intervalos.

2. Construir um intervalo de 99% de confiança para a proporção de plantas baixas.

3. Teste a hipótese de que a média da altura da planta é 230 cm. Use nível de significância igual a 0,05.

4. Teste a hipótese de que a média da altura da espiga é superior a 120 cm. Use nível de significância igual a 0,05.

5. Construa um teste t para comparar as médias de alturas e produção das turmas 1 e 2.