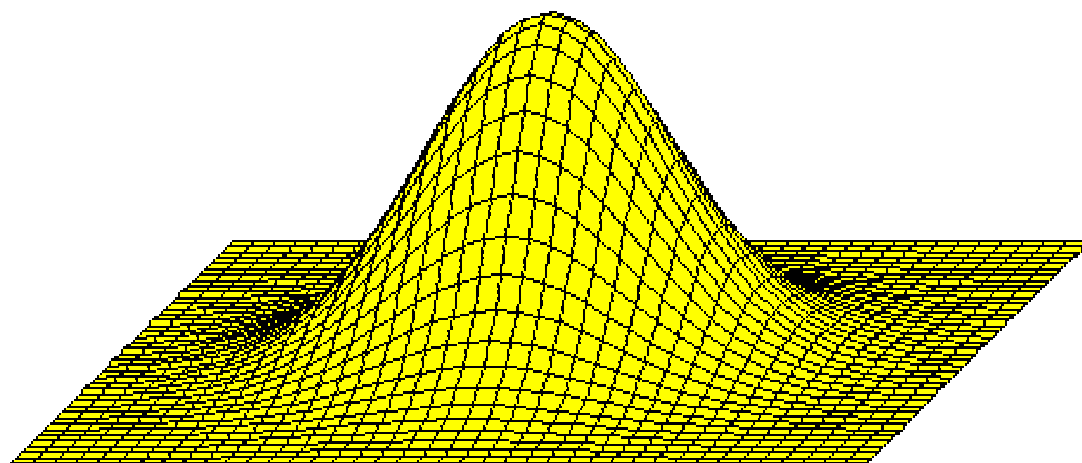


Probabilidades e Estatística no Ensino Médio:

conceitos básicos
e uso da planilha eletrônica



1. Introdução aos métodos estatísticos para a tomada de decisão

1. Para que estudar Estatística? Livro, p. 2
2. Campos da Estatística: Livro, p. 11
3. Prática inicial do EXCEL: Livro, p. 156
 - iniciar o aplicativo / células / identificação / célula ativa / inclusão : números, texto e células
 - identificação do “Colar Função” e estudo do seu potencial: Livro, p. 13
 - identificação da Ferramenta “Análise de Dados”: Livro, p. 14
 - Atenção: após digitar os dados, escolher uma célula diferente para os resultados

2. Início de um estudo: retirada de uma amostra

1. Para que retirar um amostra. **Cuidado: a amostra deve ser representativa da população**
2. como retirar (um dos vários tipos de amostragem): Aleatória (pseudoaleatória):
Exemplo: retirar uma amostra aleatória de tamanho 10 de uma população com 6750 elementos.

3. Transformando características em números: as escalas de medidas

- Escalas nominal, ordinal, intervalar e proporcional

4. Estatística Descritiva: medidas de representatividade e de dispersão

1. **Dados brutos e rol:** Livro, p. 22
2. **Medidas de representatividade (tendência central)**
 - MÉDIA ARITMÉTICA amostral, \bar{X}
 - MEDIANA AMOSTRAL
 - MODA AMOSTRAL

CUIDADO: NÃO USAR O EXCEL PARA CALCULAR A MODA!**3. Medidas de dispersão absoluta**

- AMPLITUDE TOTAL: Livro, p. 33
- VARIÂNCIA AMOSTRAL s^2
- DESVIO-PADRÃO AMOSTRAL s

4. Medidas de dispersão relativa

- COEFICIENTE DE VARIAÇÃO: Livro, p. 42

5. A variável reduzida: uma questão de justiça

1. A variável $Z = \frac{X - \mu}{s}$, que mede o desvio de um determinado valor em relação à média em unidades de desvio-padrão, é denominada **variável reduzida**, sendo uma quantidade abstrata (isto é, independe das unidades usadas).

2. No Excel, Livro, p. 92, "Colar função"/**PADRONIZAR**

- Exercício: Determine o valor da variável reduzida z correspondente ao valor 648 de uma distribuição com média 600 e desvio-padrão 12.
- Exercício: No Provão, os resultados são expressos com base em uma distribuição com média 500 e desvio-padrão 100.. Antes, os valores eram colocados em uma escala com média 80 e desvio-padrão 5. Qual resultado é o melhor: 670 no sistema atual ou 89 no sistema antigo?

6. Estatística Descritiva: apresentando os valores em uma tabela

1. **Tabela de frequências SEM perda de informação:** Livro, p. 43

valores	freqüência absoluta	freqüência relativa	freqüência absoluta acumulada	freqüência relativa acumulada
Total				

2. **Tabela de frequências COM perda de informação**

Tabela 3

classes	freqüência absoluta	freqüência relativa	freqüência acumulada
—			
—			
—			
—			
—			
—			
—			
Total			

7. Mais gráficos no Excel

8. Introdução às probabilidades, a segunda ferramenta para a Inferência

- conceito experimental: regularidade estatística
- conceito clássico
- conceito axiomático
- condições a respeitar
- OBSERVAÇÃO: para melhor compreensão, as probabilidades devem ser expressas em porcentagens.

9. Usando probabilidades para otimizar esforços

Considere uma prova de 50 questões de múltipla escolha com 5 opções cada uma. Qual a probabilidade de um aluno, marcando ao acaso, obter nota maior ou igual a 6 (ou seja, acertar 30 ou mais questões)? E se cada questão tivesse apenas 4 opções, haveria muita diferença? Este tipo de problema pode ser modelado pela chamada **distribuição binomial**.

No EXCEL, a **DISTRBINOM** fornece a resposta.

Na situação apresentada, a probabilidade do aluno tirar uma nota igual ou maior que 6 (equivalente a acertar 30 ou mais questões, e para o cálculo dessa probabilidade, deve-se ter o somatório de $x = 30$ até $x = 50$). Entretanto, o EXCEL retorna apenas o somatório de $x=0$ até o valor que se deseja. Tendo em vista que a soma de todas as probabilidades é

igual a 1, tem-se que o cálculo da probabilidade do aluno tirar nota igual ou maior que 6 é feito da seguinte forma:

a) calcula-se a probabilidade dele acertar até 29 questões, e subtrai-se de 1. Desta forma, tem-se: $\text{DISTRBINOM}(29;50;0,2;\text{VERDADEIRO}) = 0,999999999$. Por esse resultado, a chance de obter nota mínima igual a 6, marcando as respostas aleatoriamente, é igual a 0,000000001.

se houvesse 4 opções, ter-se-ia: $\text{DISTRBINOM}(29;50;0,25;\text{VERDADEIRO}) = 0,999999836$. Por esse resultado, a chance de obter nota mínima igual a 6, marcando as respostas aleatoriamente, é igual a 0,000000164, a qual, embora 164 vezes maior que a de 5 opções, é praticamente zero.

Exercício: Um professor deseja que, no máximo, haja 10% de chance de que um aluno tire nota maior ou igual a 6 em uma prova de 10 questões de múltipla escolha, marcando as respostas ao acaso. Determine o menor número de opções que garante tal probabilidade.

10. Inferência estatística: questão de confiança e risco de erro

1. Inferência Estatística: Livro, p. 105
 - a) estimacão: nada se sabe a respeito da populacão
 - b) testes de hipóteses: afirma-se algo a respeito da populacão
2. O que é necessário: conhecimento mínimo da populacão, da amostra a ser retirada (tamanho e característica) e dos riscos a correr se uma decisão errada for tomada.
3. TODA AFIRMACÃO DEVE VIR ACOMPANHADA DE UM GRAU DE CERTEZA (ou confiança)!!!
4. Toda decisão tem um risco, que é a probabilidade associada a uma decisão errada.

11. Estimando parâmetros da populacão

1. O que influencia a estimacão de maneira diretamente proporcional? E de maneira inversamente proporcional?
2. O conceito de intervalo de confiança.

12. Estimando proporções

1. Intervalo de 95% confiança para a proporção da população:

$$\bar{p} - 1,96 \sqrt{\frac{p(1-p)}{n}} \text{ --- } \bar{p} + 1,96 \sqrt{\frac{p(1-p)}{n}}$$

Exemplo: Considere uma amostra de 25 pessoas, das quais 56% delas tiveram um ataque cardíaco. Determine os limites de 95% de confiança para a verdadeira proporção de ataques cardíacos da população.

13. Testes de hipóteses

1. O que são: Livro, p. 121
2. Estrutura clássica de um teste
3. Tipos de testes: unilateral e bilateral.
4. Região de rejeição e região de não-rejeição.
5. **Termos equivalentes:**
 - a) Estatisticamente significativo = Rejeitar a hipótese nula = O valor amostral não é compatível com o valor da hipótese nula = A variação amostral não é uma explicação razoável da discrepância entre os valores da hipótese nula e os valores amostrais
 - b) Não estatisticamente significativo = Não rejeitar a hipótese nula = O valor amostral é compatível com o valor da hipótese nula = A variação amostral é uma explicação razoável da discrepância entre os valores da hipótese nula e os valores amostrais
6. Conceito moderno: valor-p.

É comum que os pesquisadores indiquem as suas conclusões com o respectivo nível de significância da seguinte forma: "**estatisticamente significativo (P<0,05)**" ou "**não estatisticamente significativo (P<0,05)**", onde P representa probabilidade no sentido de significância estatística.

A afirmação P<0,05 indica que a discrepância entre o valor amostral e o valor da hipótese nula é significativa (rejeita-se a hipótese nula) mesmo se uma significância conservadora (risco baixo) é adotada. A afirmação "P=0,05" indica que o resultado é significativo em qualquer nível até 5%.

14. Regressão linear e regressão não-linear

BIBLIOGRAFIA

LOPES, Paulo Afonso, **Probabilidades e Estatística: conceitos, modelos, aplicações em Excel**, Reichmann&Affonso, 1999.