

Métodos Quantitativos para a Gestão Ambiental

Probabilidades e Distribuições Estatísticas – Parte 2

(4/13)

Luiz Carlos Estraviz Rodriguez



Distribuição de probabilidades



Distribuição de probabilidades - **variáveis contínuas**

Para intervalos de x limitados entre A e B

Distribuição Uniforme

Distribuição Beta

Para intervalos infinitos

Cauchy

Weibull

Pareto

Logística

Log-normal

Normal

Para intervalos semi-infinitos

Distribuição Gama

Chi-quadrado (caso especial da Gama)

Exponencial (caso especial da Weibull)

A distribuição Normal

A distribuição normal descreve muitos fenômenos aleatórios e muito comuns.
A pdf de uma distribuição normal é definida como:

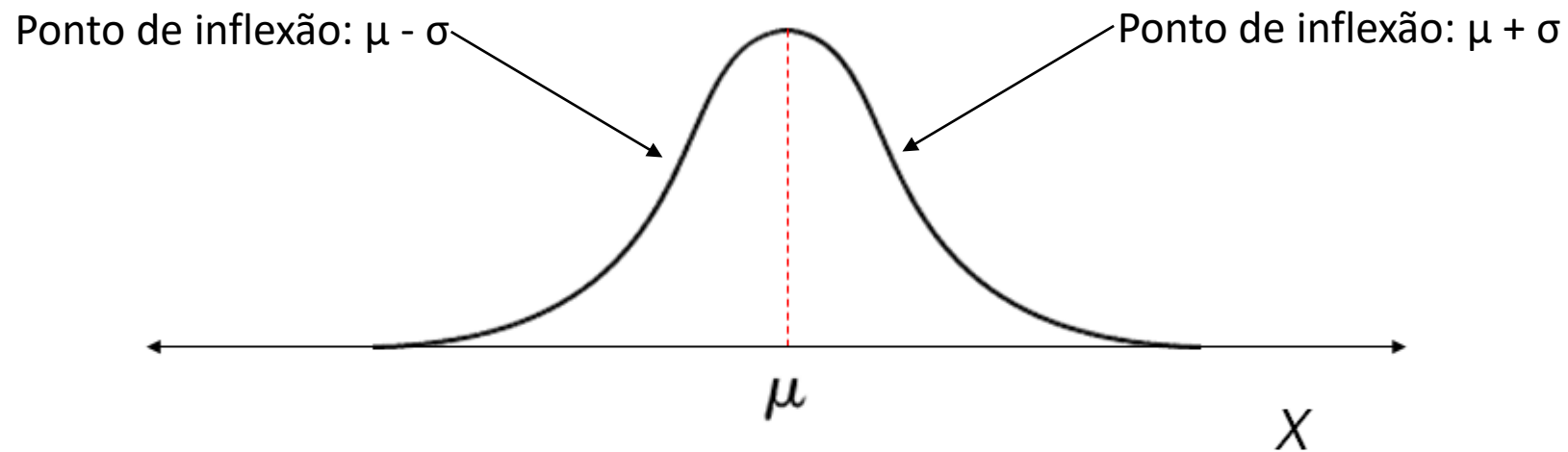
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

onde,

$$\begin{aligned} E\{x\} &= \mu \\ \text{var } \{x\} &= \sigma^2 \end{aligned}$$

A notação $N(\mu, \sigma)$ é geralmente utilizada para representar uma distribuição normal com média μ e desvio padrão σ . A importância da distribuição normal se dá também pelo fato da média das amostras tiradas de quaisquer distribuições seguirem sempre distribuição normal.

A distribuição Normal



- **Suas média, mediana e moda são iguais.**
- **Tem forma de sino e é simétrica em torno da média.**
- **A área total sob a curva é de 100%.**

Medidas de posição (ou de tendência central)

Média Aritmética:
$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$$

Mediana: valor tal que metade do conjunto de dados é igual ou inferior a esse valor e a outra metade é igual ou superior a esse valor. É o valor que divide o conjunto de dados ao meio.

Se n é IMPAR, a mediana é o valor central. Posição da mediana em uma amostra de tamanho n impar: $(n+1)/2$.

Se n é PAR, a mediana é a média aritmética dos dois valores centrais. Posições dos dois valores centrais para o cálculo da mediana em uma amostra de tamanho n par: $n/2$ e $(n+2)/2$.

Moda: valor que ocorre com maior frequência no conjunto de dados. É o valor mais comum, se existir. A moda pode não existir; o conjunto de dados pode ser bimodal, ou multimodal.

Quartis: separatrizes que dividem o conjunto de dados em quatro partes iguais

Percentis: separatrizes que dividem o conjunto de dados em cem partes iguais

Medidas de dispersão (ou de variabilidade)

Usadas para determinar a variação dos dados em torno da média, da mediana e da moda, ou qualquer outra medida de tendência central.

Amplitude total: $x_{\text{máx}} - x_{\text{mín}}$

Desvio Padrão da amostra : é a média quadrática dos desvios observados. Ou seja, se $x_1, x_2, x_3 \dots x_n$ são valores de uma amostra aleatória de tamanho n , então:

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{X})^2}{n - 1}}$$

O valor $n-1$ na fórmula explica-se pelo fato de ser conhecida a média e, portanto, o total ser conhecido a priori. Assim, conhecido o total, temos a liberdade de perder uma observação x_i da amostra e, mesmo assim, podermos recuperar toda a informação da mostra.

Variância: é o quadrado do desvio padrão

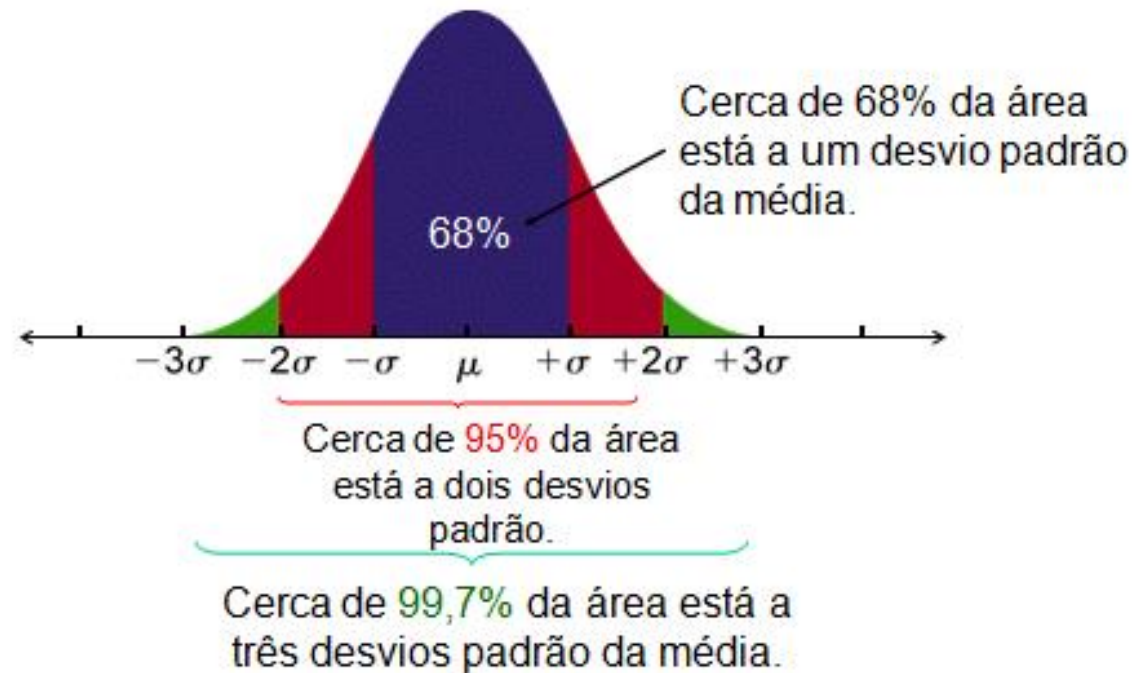
Dispersão Absoluta: A variação ou dispersão real, determinada a partir do desvio padrão, ou qualquer outra medida de dispersão.

Dispersão Relativa: dispersão absoluta / média

Coeficiente de Variação: (dispersão absoluta da amostra / média da amostra) x 100

Regra Empírica da Distribuição Normal

Regra empírica



teste de assimilação – segunda parte



Luiz Estraviz