# Distribuições Teóricas mais Importantes:

## Distribuição Normal

- É uma das distribuições mais utilizadas (talvez mesmo a mais utilizada);

- Caracteriza-se pela sua função densidade apresentar uma forma em sino, simétrica em torno da respectiva média;

- Esta curva é conhecida pela curva de Gauss, sendo a distribuição Normal conhecida também como a Gaussiana;

- Muitas características físicas se distribuem desta forma (pesos, alturas, etc.) – daqui resultam até as tabelas de percentis utilizadas nos boletins médicos infantis...;

- Também descreve bem a distribuição de erros de medição...

- Esta distribuição depende de dois parâmetros caracterizadores: a média, $\mu$, ($\mu \in$ |R) e a variância, $\sigma^2$.

- A sua função densidade é

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- Tem a particularidade de não ser possível determinar explicitamente a sua primitiva. Assim, não dispomos de uma expressão algébrica para a função de distribuição.

    - O cálculo de valores de F(x) faz-se por métodos numéricos;

    - É necessário utilizar tabelas.

- O **valor esperado** de uma normal, $\mu$, pode ser qualquer valor real, e define o eixo de simetria da função densidade respectiva.

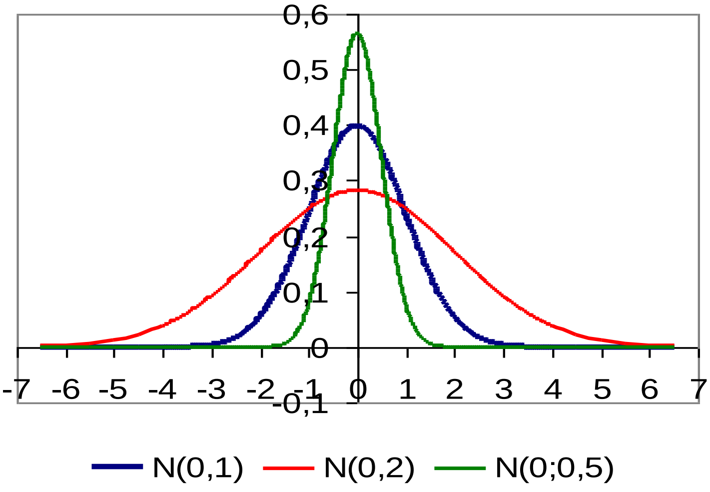
  É um parâmetro de localização.

- A **variância**, $\sigma^2$, define a dispersão em torno da média.

  Valores maiores corresponderão a distribuições mais dispersas (curva em sino mais baixa e com pontas que se estendem mais).

- Qualquer variável X que tenha distribuição normal é definida como se segue:

$$X \sim N(\mu,\sigma)$$




- Sendo X uma v.a com distribuição Normal, qualquer transformação linear de X,(ex.: aX+b), tem ainda distribuição Normal.

  Se $X \sim N(\mu, \sigma)$ e $Y = aX + b$, onde a e b são duas constantes reais, sendo a não nula, então,

  $E[Y] = E[aX+b] = a\,E[X] + b = a\,\mu + b$

  $Var[Y] = Var[aX+b] = a^2\,Var[X] = a^2\,\sigma^2$

  sendo, assim,

$$Y \sim N(a\mu+b;\ a\sigma)$$

- Em particular,

Se X ~ N(μ, σ) então, $\dfrac{X-\mu}{\sigma}$ tem distribuição N(0,1)

- Apenas uma **distribuição** normal está tabelada: a da N(0,1), que se designa usualmente por **normal-padrão**, ou **normal standard**.

Isto é:

$$z = \frac{X-\mu}{\sigma} \qquad \text{onde} \quad Z \sim N(0, 1)$$

- **Teorema da aditividade na distribuição Normal**

Sejam

$X_1, X_2, ..., X_n$, n v.a. independentes com $X_i \sim N(\mu_i, \sigma_i)$

e $a_1, a_2, ..., a_n$ constantes reais.

Então:

$$T = \sum_{i=1}^{n} a_i X_i \quad \sim \quad N\left(\sum_{i=1}^{n} a_i \mu_i ; \sqrt{\sum_{i=1}^{n} a_i^2 \sigma_i^2}\right)$$

- **Corolários**

1. Se as n v.a. forem iid N(μ, σ) então:

$$S_N = \sum_{i=1}^{n} X_i \ \ \textit{segue}\ N\left(n\mu ; \sigma\sqrt{n}\right)$$

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \ \ \textit{segue}\ N\left(\mu ; \sigma/\sqrt{n}\right)$$

2. Se X seguir N($\mu_X$, $\sigma_X$) e Y seguir N($\mu_Y$, $\sigma_Y$), X e Y independentes então:

$$X + Y \ \ segue\ N\left(\mu_X + \mu_Y \ ; \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

$$X - Y \ \ segue\ N\left(\mu_X - \mu_Y \ ; \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

- **Distribuição Normal – como aproximação da binomial**

A distribuição binomial converge para a distribuição Normal.

Se X ~ bi(n; p),

quando n → ∞ (na prática, n>20) e p "central" (entre 0,1 e 0,9) tem-se que:

$$X \ \overset{o}{\sim} \ N\left(\mu = np; \sigma = \sqrt{np(1-p)}\right)$$

ou seja,

$$\frac{X - np}{\sqrt{np(1-p)}} \ \overset{o}{\sim} \ N(0,1)$$

- **Distribuição Normal – como aproximação da Poisson**

A distribuição de Poisson converge para a distribuição Normal.

Se X ~ P(λ), quando λ → ∞ (em termos práticos, λ >20) tem-se que:

$$X \ \overset{o}{\sim} \ N\left(\mu = \lambda; \sigma = \sqrt{\lambda}\right)$$

ou seja,

$$\frac{X - \lambda}{\sqrt{\lambda}} \ \overset{o}{\sim} \ N(0,1)$$

- **Distribuição Normal: Teorema do Limite Central**

  Porque é que a distribuição normal é tão importante?

**Teorema do limite central**

i.e.,

Se tivermos uma qualquer população, X, podemos não conhecer a sua distribuição, mas, se recolhermos uma amostra suficientemente grande (*i.e.*, n → ∞), a soma de todos os $X_i$ tem um comportamento aproximadamente normal.

Ou seja,

Dada a sucessão de v. a. independentes e igualmente distribuídas, $X_1$, $X_2$, ..., $X_n$, com

$E[X_i] = \mu$ e $Var[X_i] = \sigma^2$

Então, quando n → ∞ (na prática, quando n > 30):

$$Y_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \;\dot{\cap}\; N(0;1)$$

Como consequência, nas mesmas condições

$$\frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \dot{\cap} N(0;1) \qquad \text{(se n for suficientemente grande)}$$

Nota:

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$