UNIDADE III - Elementos de probabilidades

3.1. Introdução à teoria das probabilidades
- 3.1.1. Introdução
- 3.1.2. Conceitos fundamentais
- 3.1.3. Conceitos de probabilidade
- 3.1.4. Teoremas para o cálculo de probabilidades
- 3.1.5. Probabilidade condicional e independência

3.2. Variáveis aleatórias
- 3.2.1. Introdução e conceito
- 3.2.2. Variáveis aleatórias discretas
- 3.2.3. Variáveis aleatórias contínuas

3.3. Distribuições de probabilidade
- 3.3.1. Distribuições de probabilidade de variáveis discretas
- 3.3.2. Distribuições de probabilidade de variáveis contínuas

Por que a variável é denominada aleatória?

Porque seus valores ocorrem de forma aleatória ou probabilística.

Uma das mais importantes maneiras pelas quais a aleatoriedade nos afeta é por meio de sua influência nas medições.

Por exemplo, quando um professor avalia um trabalho de um aluno a nota não é uma descrição do grau de qualidade do trabalho, mas uma medição dessa qualidade. Nesse caso, o instrumento de medição é o professor, e a avaliação de tal profissional, como qualquer medição, está sujeita a variações e erros aleatórios.

A incerteza da medição é ainda mais problemática quando a quantidade medida é subjetiva.

Para entender as medições é fundamental conhecer a forma como os dados estão distribuídos.

As medidas descritivas são uma ferramenta essencial para obter essas informações.

Outro exemplo: suponha que 15 enólogos vão avaliar um vinho.

Situação 1. Os 15 enólogos concordam que a nota do vinho é 90.

Situação 2. Os enólogos expressam as notas 80, 81, 82, 87, 89, 89, 90, 90, 90, 91, 91, 94, 97, 99 e 100.

Uma das maneiras de gerar um número único a partir de uma série de medições discordantes é calcular a média. Assim, podemos resumir as opiniões utilizando a média das notas atribuídas ao vinho. Mas essa informação é suficiente?

Os dois conjuntos de dados tem a mesma média, mas diferem no quanto variam a partir dessa média.

O desvio padrão caracteriza o quanto um conjunto de dados se aproxima da média, o que também pode significar a incerteza dos dados.

Na situação 2, o desvio padrão é 6 e o que podemos realmente dizer sobre o vinho é que sua nota provavelmente se situe entre 84 e 96.

Notamos assim que a aleatoriedade (os erros aleatórios) gera variabilidade.

A Estatística busca a regularidade presente na variabilidade

Esta regularidade permite que as variáveis aleatórias sejam representadas por modelos matemáticos, que são denominados distribuições de probabilidade.

As distribuições de probabilidade constituem a espinha dorsal da metodologia estatística, pois é com base nesses modelos que a Estatística cria técnicas que possibilitam tomar decisões válidas apesar da variabilidade e na presença de incerteza.

Na pesquisa científica o pesquisador precisa tomar decisões e obter conclusões na presença de variabilidade.

Distribuições de probabilidade

O que é uma distribuição de probabilidade?

Uma distribuição de probabilidade é essencialmente um modelo de descrição probabilística de uma população.

⇨ População estatística é o conjunto de todos os valores de uma variável aleatória.

População estatística

X=x	0	1	2	Σ
P(X=x)	0,1	0,6	0,3	1

Distribuição de probabilidade

As ideias de **população** e **distribuição de probabilidade** são indissociáveis e serão tratadas como sinônimos. O termo "distribuição" se refere ao modo como as probabilidades se **distribuem** aos diferentes valores do espaço amostral.

- **Distribuições de probabilidade:** são modelos matemáticos para descrição probabilística da ocorrência de valores de uma variável aleatória

  ⇨ Os modelos têm aplicações gerais e são individualizados por meio dos parâmetros

- **Parâmetros:** caracterizações numéricas que permitem a individualização de um modelo em determinado contexto

⇨ No estudo de uma variável aleatória é importante saber:

1. O **tipo de distribuição**, que é determinado pela **função de probabilidade** da variável
2. Os **parâmetros** da distribuição
3. As **medidas descritivas** da distribuição (média, variância, assimetria)

Distribuições discretas

1. **Distribuição de Bernoulli**
2. **Distribuição Binomial**
3. **Distribuição Hipergeométrica**
4. **Distribuição de Poisson**
5. Distribuição Multinomial
6. Distribuição Geométrica
7. Distribuição Binomial Negativa
8. Distribuição Hipergeométrica Negativa
9. Distribuição Uniforme

Distribuições contínuas

1. Distribuição Uniforme
2. **Distribuição Exponencial**
3. **Distribuição Normal**
   * **Distribuição Normal Padrão**
4. Distribuição Gama
5. Distribuição Beta
6. Distribuição Lognormal
7. Distribuição Seminormal
8. Distribuição Weibull
9. Distribuição Gumbel

Distribuições de probabilidade de variáveis discretas

1. Distribuição de Bernoulli

Jakob Bernoulli
(1654 - 1705)

Deduzida no final do século XVII pelo matemático suíço Jakob Bernoulli.

Definição: Modelo que descreve probabilisticamente os resultados de um experimento de Bernoulli.

O experimento (ou ensaio) de Bernoulli é definido como o experimento aleatório que **possui apenas dois resultados possíveis**.

Profa. Clause Piana 11

Exemplo: Uma semente é colocada para germinar

S = {germinar, não germinar}

Consideramos um dos resultados como sucesso:

sucesso = germinar

fracasso = não germinar

Definimos a variável X como número de sucessos em uma repetição do experimento.

X = número de sucessos

$$X = \begin{cases} 0 \text{ , se não germinar} \\ 1 \text{ , se germinar} \end{cases} \longrightarrow S_X = \{0, 1\}$$

π = **probabilidade de sucesso**

$1 - \pi$ = **probabilidade de fracasso**

Função de probabilidade

De modo geral, se X é uma variável que tem distribuição de Bernoulli, sua função de probabilidade será:

X = x	0	1	Σ
P(X = x)	$1-\pi$	π	1

Representação tabular

Representação analítica

$$P(X = x) = \pi^{x}.(1-\pi)^{1-x} \quad , \quad \text{para } S_X = \{0, 1\}$$

parâmetro

Parâmetro

A distribuição de Bernoulli tem apenas um parâmetro:

π = **probabilidade de sucesso**

$X \sim Ber(\pi)$

X tem distribuição de Bernoulli com parâmetro π

Medidas descritivas

- **Média ou valor esperado**

$$E(X) = \mu = \sum_{x \in S_X} x\,p(x)$$

Para $S_X = \{0, 1\}$

$X = x$	0	1	Σ
$P(X = x)$	$1-\pi$	π	1

$$E(X) = \mu = \sum_{x \in S_X} x\,p(x) = 0 \times (1-\pi) + 1 \times \pi = \pi$$

Teorema: $E(X)=\mu=\pi$

- **Variância**

$$V(X) = \sigma^2 = E(X^2) - \mu^2$$

Para $S_X = \{0, 1\}$

$X = x$	0	1	Σ
$P(X = x)$	$1-\pi$	π	1

$$E(X^2) = \sum_{x \in S_X} x^2\,p(x) = 0^2 \times (1-\pi) + 1^2 \times \pi = \pi$$

$$V(X) = \sigma^2 = E(X^2) - \mu^2 = \pi - \pi^2 = \pi\,(1-\pi)$$

Teorema: $V(X)=\sigma^2=\pi\,(1-\pi)$

- **Coeficiente de assimetria**

$$a_3 = \frac{\mu_3}{\mu_2\sqrt{\mu_2}}$$

$$\mu_2 = E(X-\mu)^2 = E(X^2) - \mu^2$$

$$\mu_3 = E(X-\mu)^3 = E(X^3) - 3\mu\, E(X^2) + 2\mu^3$$

Teorema: $a_3 = \dfrac{(1-\pi)-\pi}{\sqrt{\pi(1-\pi)}} \leftarrow \boxed{\text{desvio padrão}}$

$$a_3 = \frac{\mu_3}{\mu_2\sqrt{\mu_2}}$$

$$\mu_2 = E(X^2) - \mu^2 = \pi - \pi^2 = \pi(1-\pi)$$

$$\mu_3 = E(X^3) - 3\mu\, E(X^2) + 2\mu^3 = \pi - 3\pi\pi + 2\pi^3 = \pi - 3\pi^2 + 2\pi^3$$

$$a_3 = \frac{\pi - 3\pi^2 + 2\pi^3}{\pi(1-\pi)\sqrt{\pi(1-\pi)}}$$

$$= \frac{\pi(1 - 3\pi + 2\pi^2)}{\pi(1-\pi)\sqrt{\pi(1-\pi)}}$$

$$\begin{array}{l|l} 1 - 3\pi + 2\pi^2 & 1-\pi \\ \hline 1+\pi & 1-2\pi \\ -2\pi + 2\pi^2 & \\ 2\pi - 2\pi^2 & \end{array}$$

$$= \frac{(1 - 3\pi + 2\pi^2)}{(1-\pi)\sqrt{\pi(1-\pi)}} = \frac{(1-\pi)(1-2\pi)}{(1-\pi)\sqrt{\pi(1-\pi)}} = \frac{(1-2\pi)}{\sqrt{\pi(1-\pi)}} = \frac{(1-2\pi)}{\sqrt{\pi(1-\pi)}}$$

Teorema: $a_3 = \dfrac{(1-\pi)-\pi}{\sqrt{\pi(1-\pi)}} \leftarrow \boxed{\text{desvio padrão}}$

2. Distribuição binomial

Definição: Modelo que descreve probabilisticamente os resultados de uma seqüência de experimentos de Bernoulli **independentes** entre si, ou seja, onde a probabilidade de sucesso é **constante** em todas as repetições do experimento.

Se $X = Y_1 + Y_2 + ... + Y_n$

onde:

$Y_i \sim$ Ber (π);

$Y_{i's}$ são independentes;

então, a variável X tem distribuição binomial.

Distribuição binomial ⇨ processo finito de Bernoulli

⇨ **n** experimentos de Bernoulli independentes, com probabilidade de sucesso π constante para todos eles

⇨ É importante no contexto de **amostragem com reposição**

Exemplo: Em uma estância 60% dos bovinos foram vacinados contra uma determinada doença. Se um bovino dessa estância é escolhido ao acaso e sua situação em relação a vacinação é registrada, temos um experimento de Bernoulli.

S = {vacinado, não vacinado}

onde:

p(vacinado) = 0,6

p(não vacinado) = 1 - 0,6 = 0,4

Se três bovinos são escolhidos, um a um, e o resultado é registrado, temos uma seqüência de três experimentos de Bernoulli **independentes**, pois, a cada escolha, a probabilidade de sucesso permanecerá inalterada.

$\#S = 2^3 = 8$

V = vacinado
N = não vacinado

S = {VVV, VVN, VNV, NVV, NNV, NVN , VNN, NNN}

Sucesso = vacinado

A variável X é definida como o número de sucessos em **n** experimentos de Bernoulli independentes, com probabilidade de sucesso igual a π.

$n=3$ e $\pi=0,6$

$S_X = \{0, 1, 2, 3\}$

Qual é a função de probabilidade P(X=x) associada a variável X?

S = {VVV, VVN, VNV, NVV, NNV, NVN, VNN, NNN}

$S_X = \{0,1,2,3\}$

P(X=x) = ?

$P(X=0) = 0,4^3 = 1 \times \pi^0 \times (1 - \pi)^3 = 0,064$

$P(X=1) = 3 \times 0,6^1 \times 0,4^2 = 3 \times \pi^1 \times (1 - \pi)^2 = 0,288$

$P(X=2) = 3 \times 0,6^2 \times 0,4^1 = 3 \times \pi^2 \times (1 - \pi)^1 = 0,432$

$P(X=3) = 0,6^3 = 1 \times \pi^3 \times (1 - \pi)^0 = 0,216$

Como podemos determinar de quantas maneiras diferentes teremos **x** sucessos e **3-x** fracassos?

$P_3^{x,3-x} \leftarrow$ **Permutação com repetição**

Qual é a função de probabilidade P(X=x) associada a variável X?

$P(X=0) = P_3^{0,3}\,\pi^0\,(1-\pi)^3$

$P(X=1) = P_3^{1,2}\,\pi^1\,(1-\pi)^2$

$P(X=2) = P_3^{2,1}\,\pi^2\,(1-\pi)^1$

$P(X=3) = P_3^{3,0}\,\pi^3\,(1-\pi)^0$

$$P(X=x) = P_3^{x,3-x}\pi^x(1-\pi)^{3-x}$$

Generalizando para n repetições do ensaio de Bernoulli:

$$P(X=x) = P_n^{x,n-x}\pi^x(1-\pi)^{n-x}$$

$$P_n^{x,n-x} = \frac{n!}{x!(n-x)!}$$

Representação tabular

X = x	0	1	2	3	Σ
P(X = x)	0,064	0,288	0,432	0,216	1

Representação analítica

$$P(X=x) = \underbrace{P_3^{x,3-x}}_{\text{Número de casos}}\underbrace{0{,}6^x(1-0{,}6)^{3-x}}_{\text{Probabilidade de um caso}}\text{, para } S_X = \{0, 1, 2, 3\}$$

Por que esta distribuição é denominada binomial?

Fazendo:

$\pi = p$
$1-\pi = q$
n = repetições

Temos:

$(p+q)^n$ ← Binômio de Newton

$(p+q)^0 = 1p^0q^0 = 1$

$(p+q)^1 = 1p^1q^0 + 1p^0q^1 = p + q$

$(p+q)^2 = 1p^2q^0 + 2p^1q^1 + 1p^0q^2 = p^2 + 2pq + q^2$

$(p+q)^3 = 1p^3q^0 + 3p^2q^1 + 3p^1q^2 + 1p^0q^3$

$P(X=0) = 1\ \pi^0\ (1 - \pi)^3$

$P(X=1) = 3\ \pi^1\ (1 - \pi)^2$

$P(X=2) = 3\ \pi^2\ (1 - \pi)^1$

$P(X=3) = 1\ \pi^3\ (1 - \pi)^0$

$(p+q)^n$ ← Binômio de Newton

$(p+q)^0 = 1p^0q^0 = 1$

$(p+q)^1 = 1p^1q^0 + 1p^0q^1 = p + q$

$(p+q)^2 = 1p^2q^0 + 2p^1q^1 + 1p^0q^2 = p^2 + 2pq + q^2$

$(p+q)^3 = 1p^3q^0 + 3p^2q^1 + 3p^1q^2 + 1p^0q^3 = p^3 + 3p^2q + 3pq^2 + q^3$

$(p+q)^4 = 1p^4q^0 + 4p^3q^1 + 6p^2q^2 + 4p^1q^3 + 1p^0q^4$

$(p+q)^5 = 1p^5q^0 + 5p^4q^1 + 10p^3q^2 + 10p^2q^3 + 5p^1q^4 + 1p^0q^5$

Coeficientes
↓
Número binomial

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

$(p+q)^n$ ← Binômio de Newton

$$(p+q)^0 = 1p^0q^0 = 1$$

$$(p+q)^1 = 1p^1q^0 + 1p^0q^1 = p+q$$

$$(p+q)^2 = 1p^2q^0 + 2p^1q^1 + 1p^0q^2 = p^2 + 2pq + q^2$$

$$(p+q)^3 = 1p^3q^0 + 3p^2q^1 + 3p^1q^2 + 1p^0q^3 = p^3 + 3p^2q + 3pq^2 + q^3$$

$$(p+q)^4 = 1p^4q^0 + 4p^3q^1 + 6p^2q^2 + 4p^1q^3 + 1p^0q^4$$

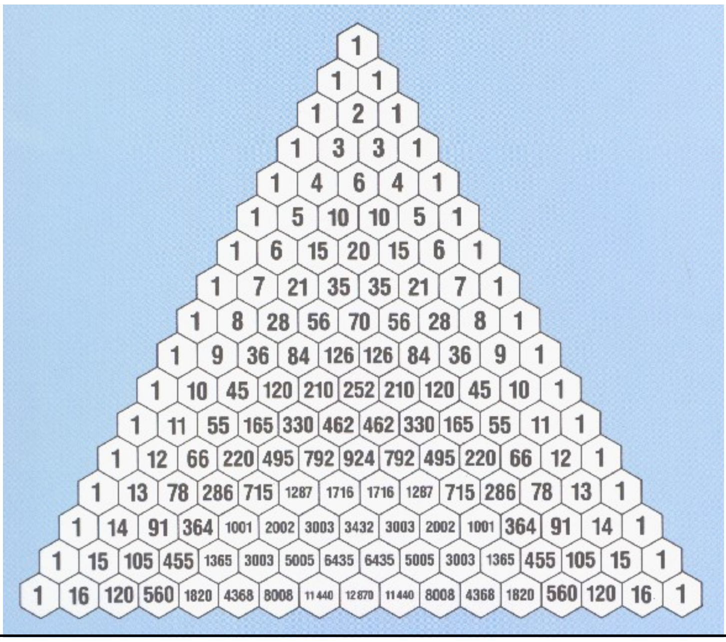
$$(p+q)^5 = 1p^5q^0 + 5p^4q^1 + 10p^3q^2 + 10p^2q^3 + 5p^1q^4 + 1p^0q^5$$

Outro método para obter os coeficientes:

Triângulo de Pascal

1
1 1
1 2 1
1 3 + 3 1
1 + 4 6 4 1
1 5 10 10 5 1

Triângulo de Pascal



Representação tabular

X = x	0	1	2	3	Σ
P(X = x)	0,064	0,288	0,432	0,216	1

Representação analítica

$$P(X = x) = P_3^{x,3-x} 0,6^x (1 - 0,6)^{3-x}, \text{ para } S_X = \{0, 1, 2, 3\}$$

Função de probabilidade

De modo geral, se X é uma variável que tem distribuição binomial, sua função de probabilidade será:

$$P(X = x) = P_n^{x,n-x} \pi^x (1 - \pi)^{n-x}, \text{ para } S_X = \{0, 1, ..., n\}$$

Parâmetros

$$P(X = x) = P_n^{x,n-x} \pi^x (1 - \pi)^{n-x}$$

parâmetros

A distribuição binomial tem dois parâmetros:

n = número de repetições do experimento de Bernoulli

π = probabilidade de sucesso

X ~ Bin (n,π)

X tem distribuição binomial com parâmetros n e π

Profa. Clause Piana 30

Medidas descritivas

- **Média ou valor esperado**

Bernoulli
$E(Y) = \pi$

$$E(X) = \mu = \sum_{x \in S_X} x\,p(x)$$

$$X = Y_1 + Y_2 + \ldots + Y_n$$

$$E(X) = E(Y_1 + Y_2 + \ldots + Y_n)$$

$$E(X) = E(Y_1) + E(Y_2) + \ldots + E(Y_n)$$

$$E(X) = \pi + \pi + \ldots + \pi \quad = n\pi$$

Teorema: $E(X)=\mu=n\pi$

Profa. Clause Piana 31

- **Variância**

Bernoulli
$E(Y) = \pi$
$V(Y) = \pi(1-\pi)$

$$V(X) = \sigma^2 \;= E(X^2) - \mu^2$$

$$X = Y_1 + Y_2 + \ldots + Y_n$$

$$V(X) = V(Y_1 + Y_2 + \ldots + Y_n)$$

$$V(X) = V(Y_1) + V(Y_2) + \ldots + V(Y_n)$$

$$V(X) = \pi(1-\pi) + \pi(1-\pi) + \ldots + \pi(1-\pi) \quad = n\pi(1-\pi)$$

Teorema: $V(X) = \sigma^2 = n\pi(1-\pi)$

- **Coeficiente de assimetria**

$$a_3 = \frac{\mu_3}{\mu_2\sqrt{\mu_2}}$$

Bernoulli

$$a_3 = \frac{(1-\pi)-\pi}{\sqrt{\pi(1-\pi)}}$$

Teorema: $$a_3 = \frac{(1-\pi)-\pi}{\underbrace{\sqrt{n\,\pi(1-\pi)}}_{\text{desvio padrão}}}$$

Interpretação:

Se $\pi < (1-\pi)$, a distribuição binomial é **assimétrica positiva**
Se $\pi = (1-\pi) = 0{,}5$, a distribuição binomial é **simétrica**
Se $\pi > (1-\pi)$, a distribuição binomial é **assimétrica negativa**

- **Coeficiente de curtose**

$$a_4 = \frac{\mu_4}{\mu_2^2} \quad \begin{cases} \text{Se } a_4<3\text{, a distribuição é } \textbf{platicúrtica} \\ \text{Se } a_4=3\text{, a distribuição é } \textbf{mesocúrtica} \\ \text{Se } a_4>3\text{, a distribuição é } \textbf{leptocúrtica} \end{cases}$$

$$a_4' = a_4 - 3$$

Teorema: $$a_4' = \frac{1-6\pi(1-\pi)}{n\pi(1-\pi)}$$

Interpretação:

Se $a_4'<0$, a distribuição binomial é **platicúrtica**
Se $a_4'=0$, a distribuição binomial é **mesocúrtica**
Se $a_4'>0$, a distribuição binomial é **leptocúrtica**

X = x	0	1	2	3	Σ
P(X = x)	0,064	0,288	0,432	0,216	1

$$E(X) = \mu = \sum_{x \in S_X} x\,p(x)$$

$$E(X) = \mu = 0 \times 0{,}064 + 1 \times 0{,}288 + 2 \times 0{,}432 + 3 \times 0{,}216 = 1{,}8$$

$$V(X) = \sigma^2 = E(X^2) - \mu^2 = 3{,}24 - 1{,}8^2 = 0{,}72$$

$$E(X^2) = 0^2 \times 0{,}064 + 1^2 \times 0{,}288 + 2^2 \times 0{,}432 + 3^2 \times 0{,}216 = 3{,}24$$

Profa. Clause Piana 35

Utilizando os teoremas: $X \sim Bin\ (n=3, \pi=0{,}6)$

$E(X)=n\pi=3\times0{,}6=1{,}8$ bovinos vacinados

Significado: Se o experimento (escolher três bovinos) for repetido um grande número de vezes, o número médio de sucessos (bovinos vacinados) obtidos nesses experimentos será 1,8.

$V(X) = n\pi(1-\pi) = 3\times0{,}6\times0{,}4 = 0{,}72$ bovinos vacinados^2

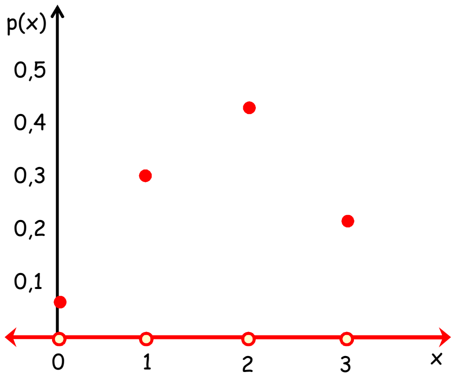
$\sigma = \sqrt{V(X)} = \sqrt{0{,}72} = 0{,}8485$ bovinos vacinados

Significado: Se o experimento for repetido um grande número de vezes, a variação média do número de sucessos em torno do valor esperado será de 0,85.

$$a_3 = \frac{(1-\pi)-\pi}{\sqrt{n\pi(1-\pi)}} = \frac{0,4-0,6}{\sqrt{3\times 0,6\times 0,4}} = -0,24$$

A distribuição é **assimétrica negativa**.

Significado: As chances para os valores maiores que a média (1,8) são maiores que para os menores.



$$a'_4 = \frac{1-6\pi(1-\pi)}{n\pi(1-\pi)}$$
$$= \frac{1-6\times 0,6\times 0,4}{3\times 0,6\times 0,4}$$
$$= -0,61$$

A distribuição é **platicúrtica**.

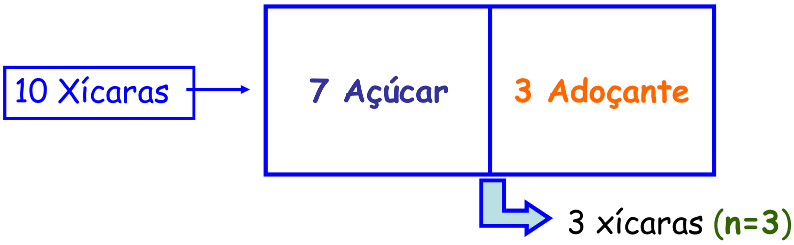
3. Distribuição hipergeométrica

Definição: Modelo que descreve probabilisticamente os resultados de uma seqüência de experimentos de Bernoulli **dependentes**. Refere-se a experimentos que se caracterizam por retiradas **sem reposição**, onde a probabilidade de sucesso **se altera** a cada retirada.

⇨ A **distribuição hipergeométrica não configura um processo de Bernoulli** porque a probabilidade de sucesso muda de um experimento para o outro

⇨ Essa distribuição é extremamente importante no contexto de **amostragem sem reposição**

Exemplo: Uma bandeja contém 10 xícaras de cafezinho, sete com açúcar e três com adoçante. Três pessoas se aproximam da bandeja e se servem. Se a variável aleatória X é definida como o número de xícaras de café com açúcar, construa a distribuição de probabilidade de X.



X = número de xícaras de café **com açúcar**

$S = \{A_1A_2A_3,\ A_1A_2A_4, \ldots,\ A_1A_2a_1,\ \ldots,\ a_1a_2a_3\}$ A= açúcar

$\#S = C_{10}^{3} = 120$ a= adoçante

$S = \{A_1A_2A_3,\ A_1A_2A_4, \ldots,\ A_1A_2a_1,\ \ldots,\ a_1a_2a_3\}$ A= açúcar

a= adoçante

$\#\ S = C_{10}^{3} = 120$

X = número de xícaras de café **com açúcar**

$S_X = \{0, 1, 2, 3\}$

$$P(X=0) = \frac{C_7^0\, C_3^3}{C_{10}^3} = \frac{1 \times 1}{120} = \frac{1}{120} = 0{,}008333$$

$$P(X=1) = \frac{C_7^1\, C_3^2}{C_{10}^3} = \frac{7 \times 3}{120} = \frac{21}{120} = 0{,}175$$

$$P(X=2) = \frac{C_7^2\, C_3^1}{C_{10}^3} = \frac{21 \times 3}{120} = \frac{63}{120} = 0{,}525$$

$$P(X=3) = \frac{C_7^3\, C_3^0}{C_{10}^3} = \frac{35 \times 1}{120} = \frac{35}{120} = 0{,}2917$$

Representação tabular

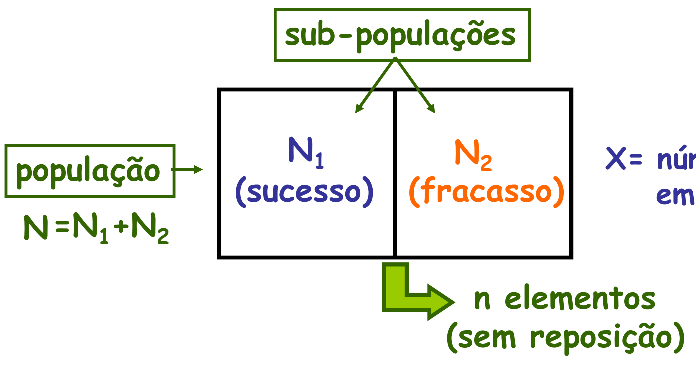
X = x	0	1	2	3	Σ
P(X = x)	0,00833	0,175	0,525	0,2917	1

Representação analítica

$$P(X = x) = \frac{C_7^x \, C_3^{3-x}}{C_{10}^3} \text{ , para } S_X = \{0, 1, 2, 3\}$$

Como generalizar esta função para qualquer variável X que tenha distribuição hipergeométrica?

Convencionamos usar a seguinte notação:



X= número de sucessos em n retiradas

- Cada grupo de tamanho n formado é denominado **amostra**.
- Todas as amostras têm a mesma chance de ocorrência.
- Do ponto de vista probabilístico não faz diferença considerar retiradas individuais sem reposição ou retirada conjunta de grupos

Representação tabular

X = x	0	1	2	3	Σ
P(X = x)	0,00833	0,175	0,525	0,2917	1

Representação analítica

$$P(X=x)=\frac{C_7^x\,C_3^{3-x}}{C_{10}^3}\text{ , para } S_X = \{0, 1, 2, 3\}$$

Função de probabilidade

De modo geral, se X é uma variável que tem distribuição hipergeométrica, sua função de probabilidade será:

$$P(X=x)=\frac{C_{N_1}^x\,C_{N_2}^{n-x}}{C_N^n}\text{ , para } S_X=\{\max(0, n-N_2), \ldots, \min(n,N_1)\}$$

Parâmetros

$$P(X=x)=\frac{C_{N_1}^x\,C_{N_2}^{n-x}}{C_N^n}$$

$\boxed{\text{parâmetros}}$

A distribuição hipergeométrica tem três parâmetros:

n = número de repetições do experimento de Bernoulli

N = tamanho da população

N_1 = tamanho da sub-população de interesse

$$X \sim Hip\,(n, N, N_1)$$

X tem distribuição hipergeométrica com parâmetros n, N e N_1

Medidas descritivas

Binomial
$E(X) = n\pi$
$V(X) = n\pi(1-\pi)$

- **Média ou valor esperado**

$$E(X) = \mu = \sum_{x \in S_X} x\,p(x)$$

Teorema: $E(X) = \mu = n\dfrac{N_1}{N}$ ← probabilidade de sucesso

- **Variância**

$$V(X) = \sigma^2 = E(X^2) - \mu^2$$

probabilidade de fracasso

Teorema: $V(X) = \sigma^2 = n\dfrac{N_1}{N}\left(\dfrac{N_2}{N}\right)\left(\dfrac{N-n}{N-1}\right)$ ← Fator de correção para populações finitas

População finita é aquela que pode ser esgotada por processo de amostragem.

⇨ Uma população será considerada finita quando tiver um **número finito** de elementos **e** a amostragem for efetuada **sem reposição**.

População infinita é aquela que não se esgota por processo de amostragem.

⇨ Uma população será considerada infinita quando tiver um **número infinito** de elementos **ou** quando amostragem for efetuada **com reposição**.

Medidas descritivas

- **Média ou valor esperado**

Teorema: $E(X) = n\dfrac{N_1}{N}$

A dependência não altera a média, mas altera a variância

- **Variância**

Teorema: $V(X) = n\dfrac{N_1}{N}\dfrac{N_2}{N}\left(\dfrac{N-n}{N-1}\right)$

Fator de correção é irrelevante para N grande

No exemplo:

X = x	0	1	2	3	Σ
P(X = x)	0,00833	0,175	0,525	0,2917	1

$$E(X) = \mu = \sum_{x \in S_X} x\,p(x)$$

$$E(X) = \mu = 0 \times 0{,}00833 + 1 \times 0{,}175 + 2 \times 0{,}525 + 3 \times 0{,}2917 = 2{,}1$$

$$V(X) = \sigma^2 = E(X^2) - \mu^2 = 4{,}9 - 2{,}1^2 = 0{,}49$$

$$E(X^2) = 0^2 \times 0{,}00833 + 1^2 \times 0{,}175 + 2^2 \times 0{,}525 + 3^2 \times 0{,}525 = 4{,}9$$

Utilizando os teoremas: X ~ Hip (n=3, N=10, N_1=7)

$$E(X) = n\frac{N_1}{N} = 3\frac{7}{10} = 2,1$$ xícaras com açúcar

Significado: Se o experimento (escolher três xícaras) for repetido um grande número de vezes, o número médio de sucessos (xícaras de café com açúcar) obtidos será 2,1.

$$V(X) = n\frac{N_1}{N}\frac{N_2}{N}\left(\frac{N-n}{N-1}\right) = 3\frac{7}{10}\frac{3}{10}\left(\frac{10-3}{10-1}\right) = 0,49$$ xícaras com açúcar²

$$\sigma = \sqrt{V(X)} = \sqrt{0,49} = 0,7$$ xícaras com açúcar

Significado: Se o experimento for repetido um grande número de vezes, a variação média do número de sucessos em relação ao valor esperado será 0,7.

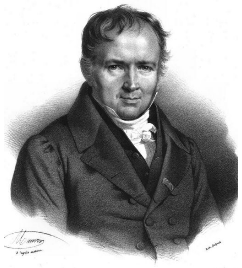
4. Distribuição de Poisson

Siméon D. Poisson
(1781 – 1840)

A distribuição de Poisson foi assim designada em homenagem ao matemático e físico francês Siméon Denis Poisson.

Modelo de descrição de acontecimentos de ocorrência rara

Exemplo clássico da distribuição de Poisson: número de soldados do exército da Prússia mortos anualmente por coice de cavalo, entre 1875 e 1894.



Siméon Denis Poisson
(1781 – 1840)

Matemático e físico francês, nascido em Pithiviers, considerado o sucessor de **Laplace** no estudo da mecânica celeste e da atração de esferóides. Filho de um administrador público, entrou para a École Polytechnique (1798), em Palaiseau, onde se formou, estudando com professores como **Joseph Louis Lagrange**, **Pierre Simon Laplace** e **Jean Baptiste Fourier**, dos quais se tornou amigo pessoal. Ocupou cargos acadêmicos na Ecole Polytechnique e na Sorbonne e contribuiu para as teorias da eletricidade e do magnetismo e estudou também o movimento da lua. Desenvolveu pesquisas sobre mecânica, eletricidade (a *constante de Poisson*), elasticidade (*razão de Poisson*), calor, som e estudos matemáticos (*integral de Poisson* na teoria do potencial e o *colchete de Poisson* nas equações diferenciais), com aplicação na medicina e na astronomia.

Produziu escritos sobre movimentos de ondas em geral e coeficientes de contração e a relação entre estes e a extensão. Publicou trabalhos (1812) que ajudaram a eletricidade e o magnetismo tornarem-se um ramo da física matemática. Ganhou o título de barão (1825). Na hidrodinâmica seu mais notável trabalho foi *Mémoire sur les équations générales de l'équilibre et du mouvement des corps solides élastiques et des fluides* (1829), relacionando equilíbrio de sólidos elásticos e correntes de fluidos compressíveis. Publicou o importante tratado *Traité de mécanique* (1833), em dois volumes, na termodinâmica a *Teoria matemática do calor* (1835) e em *Recherches sur la probabilité des jugements* (1837) apareceu a famosa *distribuição de Poisson*, de intensa aplicação em estatística. Na teoria de probabilidades descobriu a forma limitada da distribuição binomial que posteriormente recebeu o seu nome e hoje considerada uma das mais importantes distribuições na probabilidade, sendo o *método de Poisson* é um processo randômico de importância fundamental. Publicou cerca de quatrocentos trabalhos e morreu em Sceaux, próximo a Paris, França.

(Só Biografias. Disponível em http://www.dec.ufcg.edu.br/biografias/SimeonDe.html)

4. Distribuição de Poisson

Definição: descreve probabilisticamente a seqüência de um grande número de fenômenos independentes entre si, cada um com probabilidade de sucesso muito pequena.

⇨ Ocorre naturalmente quando se deseja contar o número de um tipo particular de eventos que ocorrem por unidade de tempo, de superfície ou de volume.

⇨ Pode ser considerada como uma binomial onde o número de experimentos (**n**) é grande, π é pequeno (sucesso raro) e $n\pi$ (média de sucessos) é constante.

Distribuição de Poisson ⇨ processo infinito de Bernoulli

Exemplos:

- número de peças defeituosas observadas em uma linha de produção num determinado período de tempo;
- número de acidentes de trabalho ocorridos numa grande empresa num determinado período de tempo;
- número de ciclones ocorridos em certa região num determinado período de tempo;
- número de formigueiros por unidade de área em uma região;
- número de bactérias por unidade de área em uma lâmina com extratos de uma planta;
- número de espermatozóides inviáveis de um touro por unidade de volume de sêmen.

⇨ A distribuição de Poisson tem inúmeras aplicações na **simulação de sistemas** modelando o número de eventos ocorridos num intervalo de tempo, quando os eventos ocorrem a uma taxa constante.

Função de probabilidade

De modo geral, se X é uma variável que tem distribuição de Poisson, sua função de probabilidade será:

$$P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!} \text{ , para } S_X = \{0, 1, 2, ...\}$$

espaço amostral infinito

onde:

X: número de sucessos

e = 2,718 (base dos logaritmos neperianos)

λ: **número médio de sucessos** (sempre maior que zero)

Demonstra-se que $p(x) = e^{-\lambda}\dfrac{\lambda^x}{x!}$ é uma função de probabilidade, provando que $\sum_{x=0}^{\infty} p(x) = 1$. Como S_X = {0, 1, 2, ...}, temos:

$$\sum_{x=0}^{\infty} p(x) = \sum_{x=0}^{\infty} e^{-\lambda}\frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\left(\frac{\lambda^0}{0!}+\frac{\lambda^1}{1!}+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\ldots\right)$$

$$= e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2!}+\frac{\lambda^3}{3!}+\ldots\right)$$

$$= e^{-\lambda}e^{\lambda}$$

$$= e^{-\lambda+\lambda} = e^0 = 1$$

Série de Taylor

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

$$e^\lambda = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \ldots$$

Parâmetros

$$P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}$$

parâmetro

A distribuição de Poisson tem apenas um parâmetro:

λ = número médio de sucessos

$X \sim$ Poi (λ)

X tem distribuição de Poisson com parâmetro λ

Medidas descritivas

- **Média ou valor esperado:** $E(X) = \mu = \sum_{x \in S_X} x\, p(x)$

Teorema: $E(X) = \mu = \lambda$

- **Variância:** $V(X) = \sigma^2 = E(X^2) - \mu^2$

Teorema: $V(X) = \sigma^2 = \lambda$

Na Poisson média e variância são iguais!!

Demonstração

- **Média ou valor esperado:** $E(X) = \mu = \sum_{x \in S_X} x\, p(x)$

$$E(X) = \sum_{x \in S_X} x\, p(x) = \sum_{x=0}^{\infty} x\, e^{-\lambda} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} x \frac{\lambda}{x} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= e^{-\lambda} \sum_{x=1}^{\infty} \lambda \frac{\lambda^{x-1}}{(x-1)!} = e^{-\lambda} \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$, fazendo $y = x - 1$, temos

$$= e^{-\lambda} \lambda \boxed{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!}} = e^{-\lambda} \lambda\, e^{\lambda} = e^{-\lambda+\lambda} \lambda = e^0 \lambda = \lambda$$

$$\boxed{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = e^{\lambda}}$$

Teorema: $E(X) = \mu = \lambda$

Demonstração

- **Variância:** $V(X) = \sigma^2 = E(X^2) - \mu^2$

$$\boxed{\sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} = \lambda e^{\lambda}} \qquad \boxed{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!} = e^{\lambda}}$$

$$E(X^2) = \sum_{x \in S_X} x^2 p(x) = \sum_{x=0}^{\infty} x^2 e^{-\lambda} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} x^2 \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=1}^{\infty} x^2 \frac{\lambda}{x} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= e^{-\lambda} \lambda \sum_{x=1}^{\infty} x \frac{\lambda^{x-1}}{(x-1)!}, \text{ fazendo } y = x - 1, \text{ temos}$$

$$= e^{-\lambda} \lambda \sum_{y=0}^{\infty} (y+1) \frac{\lambda^y}{y!} = e^{-\lambda} \lambda \sum_{y=0}^{\infty} \left(y \frac{\lambda^y}{y!} + \frac{\lambda^y}{y!} \right) = e^{-\lambda} \lambda \left(\boxed{\sum_{y=0}^{\infty} y \frac{\lambda^y}{y!}} + \boxed{\sum_{y=0}^{\infty} \frac{\lambda^y}{y!}} \right)$$

$$= e^{-\lambda} \lambda (\lambda e^{\lambda} + e^{\lambda}) = \lambda e^{\lambda} e^{-\lambda} \lambda + e^{\lambda} e^{-\lambda} \lambda = e^0 \lambda^2 + e^0 \lambda = \lambda^2 + \lambda$$

$$V(X) = E(X^2) - \mu^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Teorema: $V(X) = \sigma^2 = \lambda$

- **Coeficiente de assimetria**

Teorema: $a_3 = \dfrac{1}{\sqrt{\lambda}}$

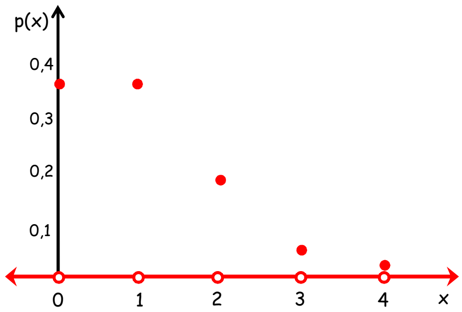
A Poisson é assimétrica positiva, tendendo para a simetria quando μ cresce.

- **Coeficiente de curtose**

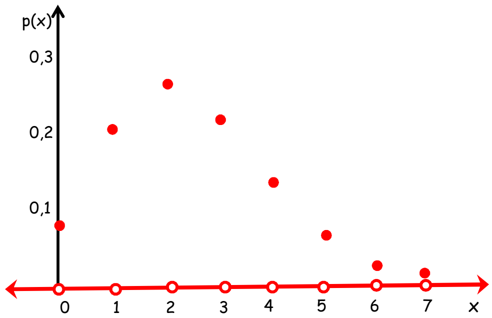
Teorema: $a_4 = \sqrt{\lambda}$

A Poisson é platicúrtica, tendendo para mesocúrtica quando μ cresce.

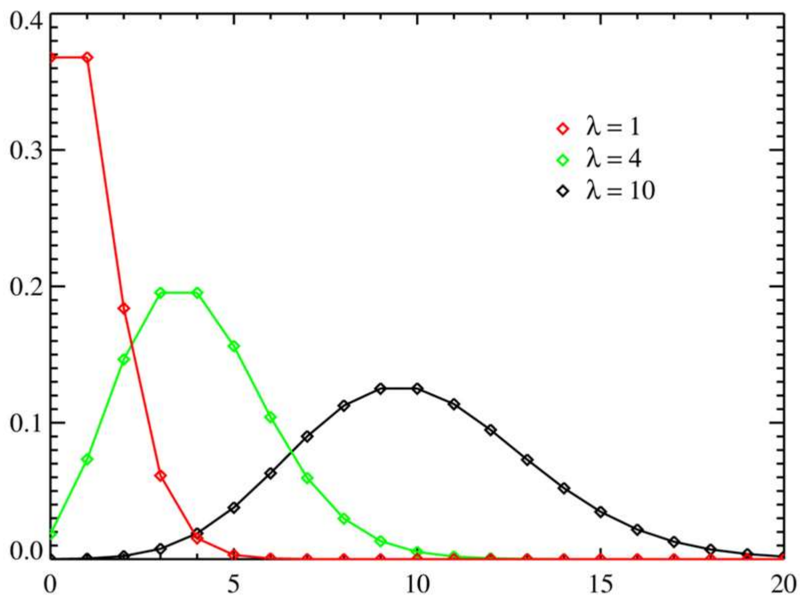
$X \sim Poi\ (\lambda=1)$



$X \sim Poi\ (\lambda=2,5)$



A Poisson tende para a simetria quando $\mu=\lambda$ cresce.



Exercício proposto:

A taxa média de chegada de clientes em um posto de serviços é de 0,5 por minuto. Calcular a probabilidade de, em um dado minuto, chegarem dois clientes.

Resumo das distribuições discretas

Distribuição de Bernoulli

Descreve probabilisticamente resultados de experimentos que possuem apenas dois resultados possíveis.

Função de probabilidade

$$P(X=x)=\pi^{x}(1-\pi)^{1-x}\ , \quad \text{para } S_X = \{0, 1\}$$

Parâmetro: π = probabilidade de sucesso

Medidas descritivas

$E(X) = \mu = \pi$

$V(X) = \sigma^2 = \pi\,(1-\pi)$

$$a_3=\frac{(1-\pi)-\pi}{\sqrt{\pi(1-\pi)}}$$

Distribuição binomial

Descrição probabilística de uma seqüência de experimentos de Bernoulli independentes. Importante no contexto de amostragem com reposição.

Função de probabilidade

$$P(X=x)=P_n^{x,n-x}\,\pi^{x}(1-\pi)^{n-x}\ , \text{ para } S_X = \{0, 1, ..., n\}$$

Parâmetros: n = número de repetições no experimento

π = probabilidade de sucesso

Medidas descritivas

$E(X) = \mu = n\pi$

$$a_3=\frac{(1-\pi)-\pi}{\sqrt{n\pi(1-\pi)}}$$

$V(X) = \sigma^2 = n\pi(1-\pi)$

$$a_4'=\frac{1-6\pi(1-\pi)}{n\pi(1-\pi)}$$

Distribuição hipergeométrica

Descrição probabilística de uma seqüência de experimentos de Bernoulli dependentes. Importante no contexto de amostragem sem reposição.

Função de probabilidade

$$P(X=x)=\frac{C_{N_1}^{x}\,C_{N_2}^{n-x}}{C_N^n}\ ,\quad \text{para } S_X = \{0, 1, ..., n\}$$

Parâmetros: n = número de repetições do experimento
N = tamanho da população
N_1 = tamanho da sub-população de interesse

Medidas descritivas

$$E(X)=\mu=n\frac{N_1}{N} \qquad V(X)=\sigma^2=n\frac{N_1}{N}\frac{N_2}{N}\left(\frac{N-n}{N-1}\right)$$

Distribuição de Poisson

Descrição probabilística da seqüência de um grande número de fenômenos independentes, todos com probabilidade de sucesso muito pequena, e média de sucessos constante.

Função de probabilidade

$$P(X=x)=e^{-\lambda}\frac{\lambda^x}{x!}\ ,\quad \text{para } S_X = \{0, 1, 2, ...\}$$

Parâmetro: λ = número médio de sucessos

Medidas descritivas

$E(X) = \mu = \lambda$ $\qquad a_3=\frac{1}{\sqrt{\lambda}}$

$V(X) = \sigma^2 = \lambda$ $\qquad a_4=\sqrt{\lambda}$

Formas limites da distribuição binomial

- Em determinadas circunstâncias, uma distribuição de probabilidade pode tender para outra.
- Os casos mais importantes de aproximações entre distribuições são:

1. **Hipergeométrica → Binomial**
2. **Binomial → Poisson**
3. **Binomial → Normal**

1. Hipergeométrica → Binomial

Quando N (tamanho da população) é muito grande (ou tende para +∞), a distribuição **hipergeométrica** se aproxima da distribuição **binomial**.

fator de correção

Binomial

$$V(X) = n\pi(1-\pi)$$

Hipergeométrica

$$V(X) = n\,\frac{N_1}{N}\,\frac{N_2}{N}\left(\frac{N-n}{N-1}\right)$$

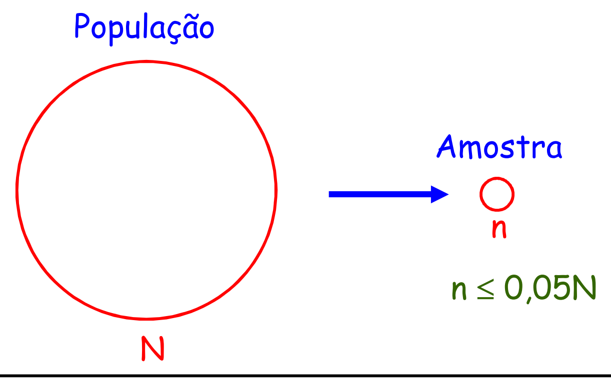
Quando N tende para infinito, o fator de correção se aproxima de 1.

tende a zero

$$\frac{N-n}{N-1} \cong \frac{N-n}{N} = \frac{N}{N} - \frac{n}{N} = 1 - \frac{n}{N} \cong 1$$

tende a +∞

Esta aproximação é considerada satisfatória quando o número de elementos retirados (n) não excede 5% da população (N), isto é, $n \leq 0{,}05N$.



Quando o tamanho da amostra representa menos de 5% do tamanho da população, a população é tão grande em relação a amostra que pode ser considerada infinita.

2. Binomial → Poisson

Quando n (número de repetições do experimento) é muito grande (ou tende para $+\infty$) e π (probabilidade de sucesso) é muito pequena (ou tende para 0) a distribuição **binomial** se aproxima da distribuição de **Poisson**.

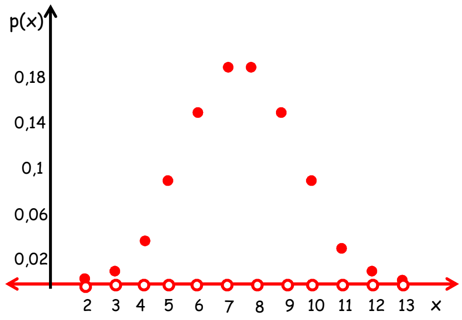
$$\underbrace{n\pi}_{E(X)} < 10 \quad e \quad n \geq 100$$

3. Binomial → Normal

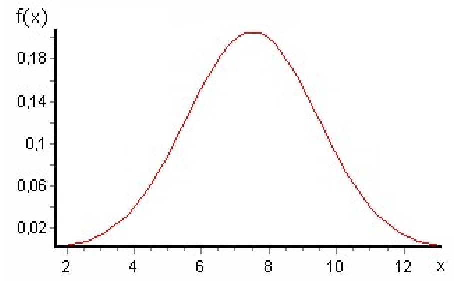
Quando n (número de repetições do experimento) é grande (ou tende para $+\infty$) e π (probabilidade de sucesso) se aproxima de 0,5, a distribuição **binomial** se aproxima da distribuição **normal**.

Se $\pi=(1-\pi)=0{,}5$, a distribuição binomial será simétrica.

X ~ Bin (n=15,π=0,5)



X ~ Nor (μ=7,5,σ²=3,75)



Exercícios propostos:

Um auditor foi contratado para examinar uma coleção de 6.000 faturas, das quais 128 contêm erros. Se foi selecionada uma amostra de 120 faturas, qual é a probabilidade de esta amostra conter exatamente duas faturas com erros?

Sendo de 1% o percentual de canhotos numa população, qual é a probabilidade de haver apenas um canhoto numa classe de 30 alunos?

Distribuições de probabilidade de variáveis contínuas

Em geral é difícil identificar o tipo de distribuição de probabilidade de uma variável contínua. Inicialmente, uma revisão bibliográfica pode útil para verificar se a variável de interesse já foi estudada e seu modelo de descrição probabilística já é conhecido.
A observação do campo de variação (espaço amostrar) da variável também pode ajudar nesta identificação.

Existem vários tipos de distribuições contínuas. Por exemplo:

- ⇨ **Distribuição gama:** descreve variáveis que só assumem valores positivos.
- ⇨ **Distribuição beta:** descreve variáveis que assumem valores no intervalo [0, 1].
- ⇨ **Distribuição normal** → mais importante

Distribuições contínuas

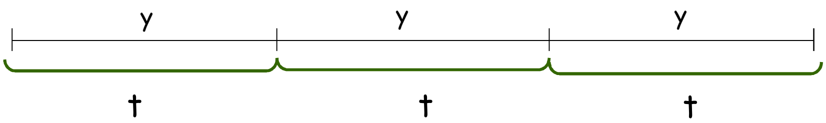
1. Distribuição Exponencial
2. Distribuição Normal
   * Distribuição Normal Padrão
3. Distribuição Uniforme
4. Distribuição Gama
5. Distribuição Beta
6. Distribuição Lognormal
7. Distribuição Seminormal
8. Distribuição Weibull
9. Distribuição Gumbel

(8 e 9) → eventos climáticos extremos

1. Distribuição exponencial

A distribuição exponencial está relacionada com a distribuição de Poisson.
Na distribuição de Poisson, a variável aleatória é definida como o número de sucessos em determinado período de tempo t, sendo a média de sucessos no período definida como λ.
Na distribuição exponencial, a variável aleatória é definida como o tempo entre dois sucessos.

Processo de Poisson $E(Y) = \lambda$



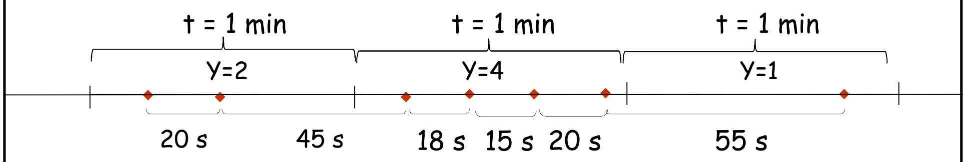
Y = número de sucessos → variável discreta

X = tempo decorrido entre sucessos sucessivos → variável contínua

Exemplo: Se a média de atendimentos no caixa de uma loja é de 3 clientes por minuto, então, o tempo decorrido entre atendimentos é uma variável com distribuição exponencial.

Y = número de atendimentos → distribuição de Poisson

$S_Y = \{0, 1, 2, \ldots\}$ $E(Y) = \lambda = 3$



X = tempo entre atendimentos (em segundos) → **distribuição exponencial**

$S_X = (0, \infty)$

A distribuição exponencial é extensivamente utilizada para modelar o tempo entre ocorrências de eventos num sistema, onde os eventos ocorrem a uma taxa constante. Desse modo, tem grande aplicabilidade em estudos de sobrevivência, confiabilidade, teoria das filas e simulação.

Função densidade de probabilidade

De modo geral, se X é uma variável aleatória contínua que tem distribuição exponencial, então, sua função densidade de probabilidade será:

$$f(x) = \lambda e^{-\lambda x}, \quad \text{para } S_X = [0, \infty)$$

X só assume valores reais não negativos

onde:

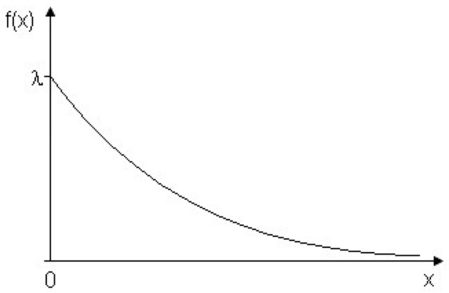
- X: tempo decorrido entre dois sucessos
- e = 2,718 (base dos logaritmos neperianos)
- λ: número médio de sucessos (sempre maior que zero)

Podemos demonstrar que f(x) é uma função densidade de probabilidade provando que $\int_{S_x} f(x)\,dx = 1$

$$\int_{S_x} f(x)\,dx = \int_0^{\infty} \lambda\, e^{-\lambda x} dx = \left[e^{-\lambda x} \right]_0^{\infty} = e^{-\lambda \infty} - e^{-\lambda 0} = 0 - (-1) = 1$$

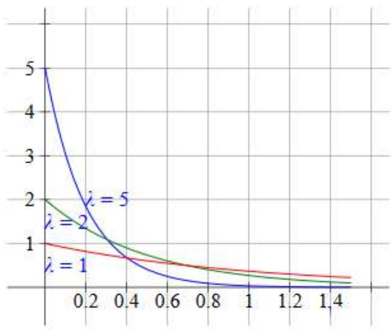
Função densidade de probabilidade

Representação gráfica: $f(x) = \lambda e^{-\lambda x}$



$$f(0) = \lambda e^{-\lambda 0} = \lambda e^{0} = \lambda$$

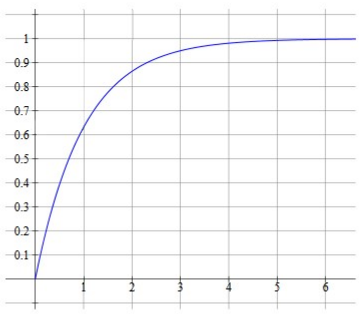
Exemplos:



Função de distribuição ou de probabilidade acumulada

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$$

Representação gráfica:



$$P(X > x) = 1 - F(x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

Parâmetros

$$f(x) = \lambda e^{-\lambda x}$$

$$\boxed{\text{parâmetro}}$$

A distribuição exponencial tem apenas um parâmetro:

λ = número médio de sucessos

$X \sim \text{Exp}(\lambda)$

X tem distribuição exponencial com parâmetro λ

Medidas descritivas

- **Média ou valor esperado:** $E(X) = \mu = \int_{S_X} x\, f(x) dx$

$$E(X) = \mu = \int_{S_x} x\, f(x)\, dx$$

$$= \int_0^\infty x\lambda\, e^{-\lambda x} dx = \frac{1}{\lambda}$$

Teorema: $E(X) = \mu = \dfrac{1}{\lambda}$

Medidas descritivas

- **Variância:** $V(X) = \sigma^2 = E(X^2) - \mu^2$

$$V(X) = \sigma^2 = \left[\int_{S_x} x^2 f(x) dx \right] - \mu^2$$

$$= \left[\int_0^{\infty} x^2 \left(\lambda\, e^{-\lambda x} \right) dx \right] - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Teorema: $V(X) = \sigma^2 = \dfrac{1}{\lambda^2}$

Exercícios propostos:

1. Os tempos até a falha de um dispositivo eletrônico seguem o modelo exponencial, com uma taxa de falha λ = 0,012 falhas/hora. Indique qual a probabilidade de um dispositivo escolhido ao acaso sobreviver a 50 horas? E a 100 horas? Qual o tempo esperado até que ocorra falha?

2. Suponha que um componente eletrônico tenha um tempo de vida X (em unidades de 1000 horas) que segue uma distribuição exponencial de parâmetro λ = 1. Suponha que o custo de fabricação do item seja 2 reais e que o preço de venda seja 5 reais. O fabricante garante devolução total se $X < 0{,}9$. Qual o lucro esperado por item?

Resolução:

Suponha que um componente eletrônico tenha um tempo de vida X (em unidades de 1000 horas) que segue uma distribuição exponencial de parâmetro $\lambda = 1$. Suponha também que o custo de fabricação do item seja 2 reais e que o preço de venda seja 5 reais. O fabricante garante devolução total se X<0,9. Qual o lucro esperado por item?

X = tempo de vida (1000 h)

Y = lucro por item (reais)

Se X < 0,9, então Y = -2 reais

Se X > 0,9, então Y = 3 reais

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}$$

$$\text{Sendo } \lambda = 1, \quad F(0{,}9) = P(X \leq 0{,}9) = 1 - e^{-0{,}9} = 0{,}5934$$

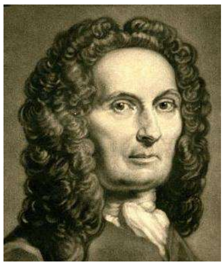
X	x<0,9	x>0,9	
Y = y	-2	3	Σ
P(Y = y)	0,5934	0,4066	1

$E(Y) = -2 \times 0{,}5934 + 3 \times 0{,}4066 = 0{,}03$ reais

2. Distribuição normal

Síntese histórica

A distribuição normal foi primeiramente obtida por Abraham De Moivre, num artigo de 1734, no contexto de aproximação da distribuição binomial quando o número n de experimentos de Bernoulli cresce.
De Moivre se deparou com a curva normal ao buscar uma aproximação para os números que habitam regiões do triângulo de Pascal.
Os gráficos abaixo representam as magnitudes de algumas linhas do triângulo de Pascal.



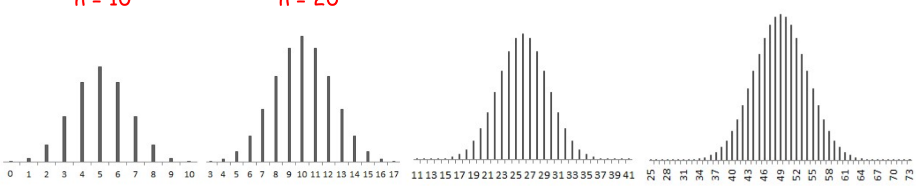
Abraham De Moivre
(1667 - 1754)

Linha 11
n = 10

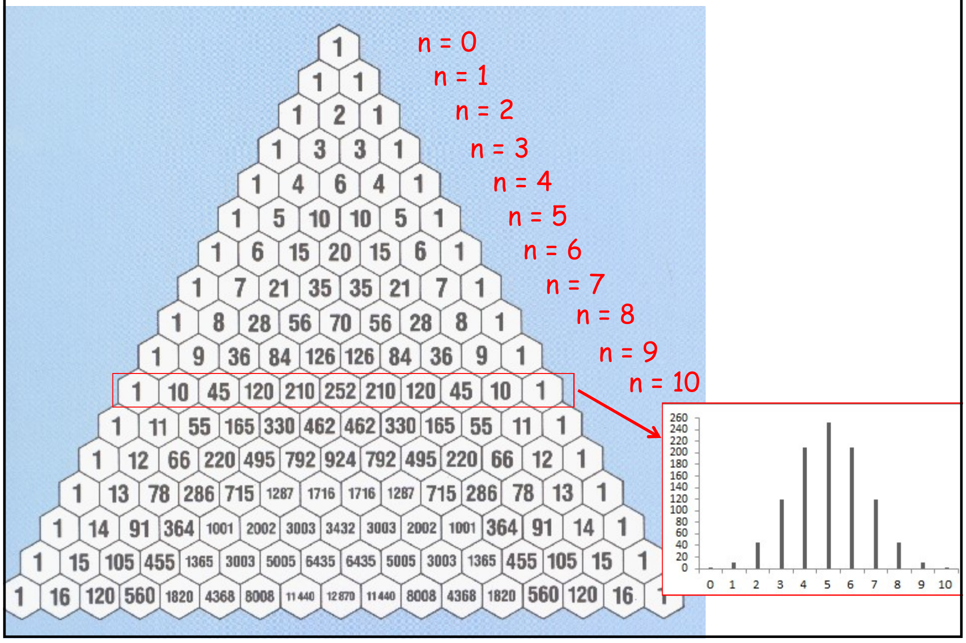
Linha 21
n = 20

Linha 51
n = 50

Linha 101
n = 100



Triângulo de Pascal



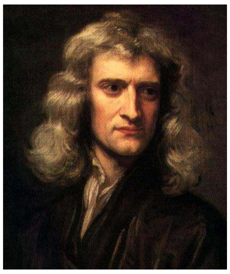
Carl Friedrich Gauss
(1777 –1855)

Em 1809, Gauss assumiu que os erros de medida poderiam ser modelados pela distribuição normal. Ele teve essa percepção ao realizar medições astronômicas, enquanto trabalhava no problema dos movimentos planetários.

Em 1812, em seu livro “Theórie Analytique des Probabilités”, Laplace melhorou o resultado encontrado por De Moivre para o que hoje é denominado Teorema de De Moivre-Laplace. Também foi Laplace quem tirou a distribuição normal da obscuridade e criou um argumento melhor do que o de Gauss para sustentar a noção de que a distribuição normal é, de fato, a lei dos erros.

Pierre-Simon Laplace
(1749 – 1827)

Sobre os erros de medida…



Isaac Newton
(1642 - 1727)

Newton foi o primeiro a empregar a média para obter um único valor a partir de uma série de medições discordantes.
Mas a maior parte dos cientistas daquela época e no século seguinte não calculava a média, em vez disso, escolhia dentre suas medições um único "número áureo" considerado, essencialmente por palpite, o mais confiável dos resultados obtidos.
Isso porque não consideravam a variação como um subproduto inevitável do processo de mensuração, e sim como uma evidência de fracasso.

medições discordantes

$x_i = x_1, x_2, x_3, \ldots, x_n$

$$\bar{x} = \frac{\sum x}{n}$$

valor mais razoável para a medida

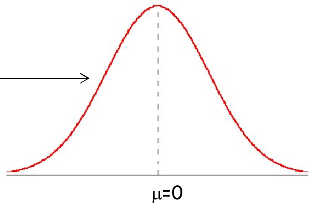
desvio = erro de medida

$(x_i - \bar{x})$

média dos erros

$$\frac{\sum (x_i - \bar{x})}{n} = 0$$

Distribuição dos erros de medida



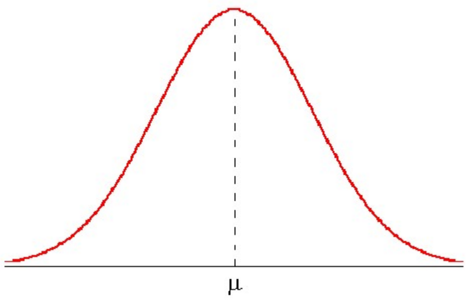
Atualmente, a distribuição normal é importante tanto do ponto de vista teórico como nas aplicações. Essa importância se deve a um conjunto de aspectos:

- ⇨ Suas propriedades matemáticas
- ⇨ É útil para descrever uma grande quantidade de fenômenos naturais físicos, ambientais, psicométricos etc, além dos erros de medida
- ⇨ Distribuições de um grande número de variáveis aleatórias convergem para a distribuição normal
- ⇨ Muitas variáveis não normais podem ser tratadas como normais após transformações simples
- ⇨ Uma grande quantidade de métodos e procedimentos de inferência estatística são derivados tendo-a como pressuposição básica

O conjunto de métodos desenvolvidos para tratar variáveis que têm distribuição normal forma a chamada **Estatística Clássica** ou **Estatística Paramétrica**.

2. Distribuição normal

Definição: É uma distribuição teórica de frequências, onde a maioria das observações se situa em torno da **média** (centro) e diminui gradual e simetricamente no sentido dos extremos.
A distribuição normal é representada graficamente pela curva normal (curva de Gauss) que tem a forma de sino e é **simétrica** em relação ao centro, onde se localiza a média μ.



Profa. Clause Piana 95

Função densidade de probabilidade

De modo geral, se X é uma variável contínua que tem distribuição normal, sua função densidade de probabilidade será:

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

parâmetros

- A distribuição normal é um membro da família exponencial.

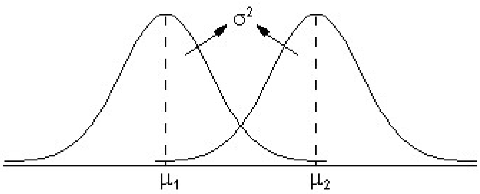
Parâmetros

A distribuição normal tem dois parâmetros:

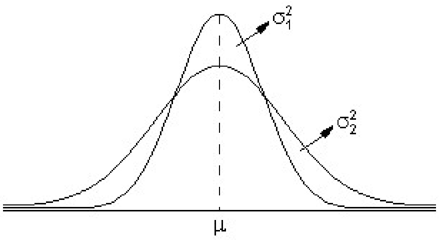
μ = média (determina o centro da distribuição)
σ^2 = variância (determina a dispersão da distribuição)

$$X \sim N(\mu, \sigma^2)$$

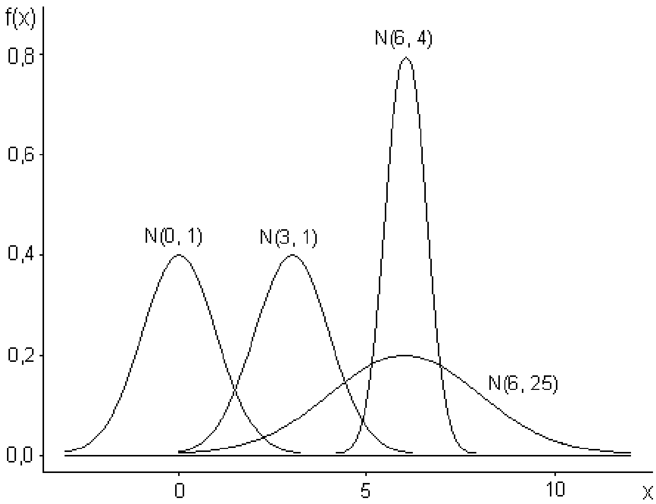
X tem distribuição normal com parâmetros μ e σ^2



Populações normais com médias diferentes e mesma variância



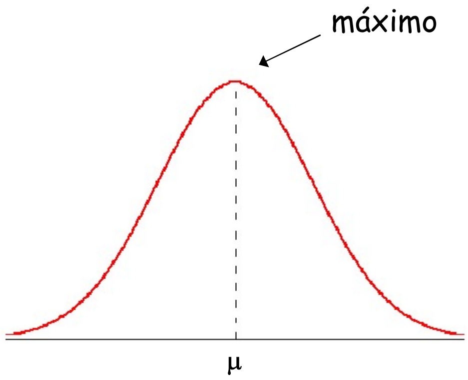
Populações normais com variâncias diferentes e mesma média



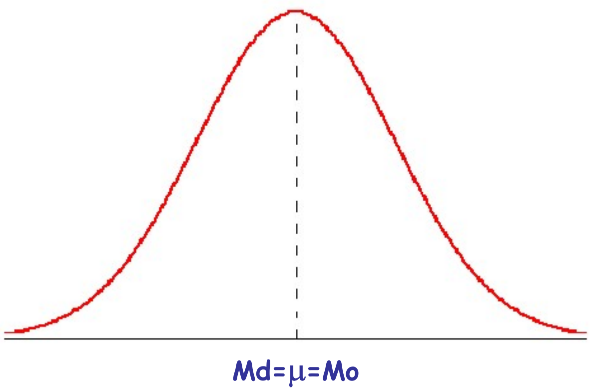
Existe um número infinito de curvas normais

Propriedades da distribuição normal

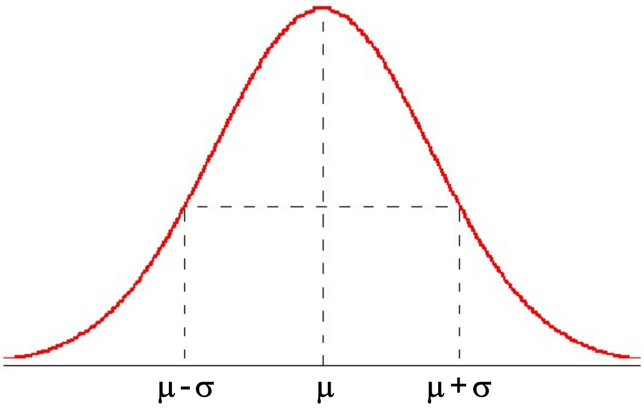
1. O máximo da função densidade de probabilidade se dá no ponto **x=μ** .



2. A distribuição é simétrica em relação ao centro onde coincidem a média, a moda e a mediana.



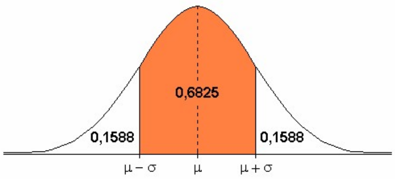
3. Os pontos de inflexão são exatamente $\mu-\sigma$ e $\mu+\sigma$.



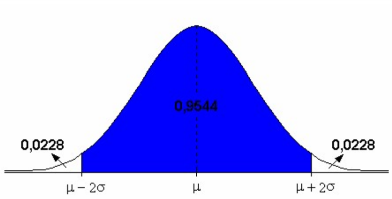
Ponto de inflexão: ponto onde a concavidade à direita tem sinal diferente ao da concavidade à esquerda

4. Verifica-se na distribuição normal que:

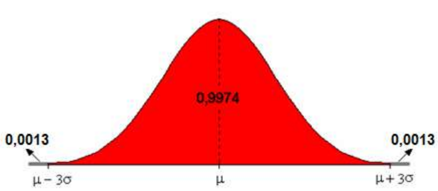
$P(\mu-\sigma < X < \mu+\sigma) = 0,6825$



$P(\mu-2\sigma < X < \mu+2\sigma) = 0,9544$



$P(\mu-3\sigma < X < \mu+3\sigma) = 0,9974$



Se X é uma variável que tem distribuição normal com $\mu = 5$ e $\sigma^2 = 4$, calcule:

a) $P(3 < X < 7)$

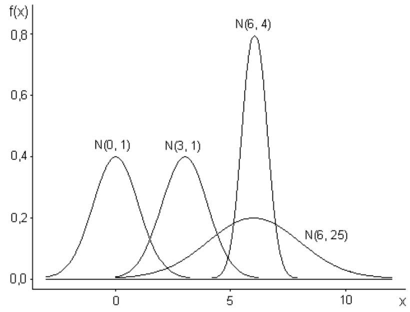
b) $P(5 < X < 9)$

c) $P(4 < X < 10)$

Cálculo de áreas

⇨ Para cada valor de μ e de σ, existe uma distribuição normal diferente

⇨ Existem infinitas distribuições (e curvas) normais, pois basta que mude um dos parâmetros para termos uma distribuição diferente



⇨ O cálculo de áreas sob a curva normal, deverá ser feito sempre em função dos valores particulares de μ e σ

⇨ Para evitar a trabalhosa tarefa de calcular as áreas foi determinada uma distribuição normal padrão ou reduzida

⇨ As áreas sob a curva normal padrão foram calculadas e apresentadas numa tabela

Distribuição normal padrão

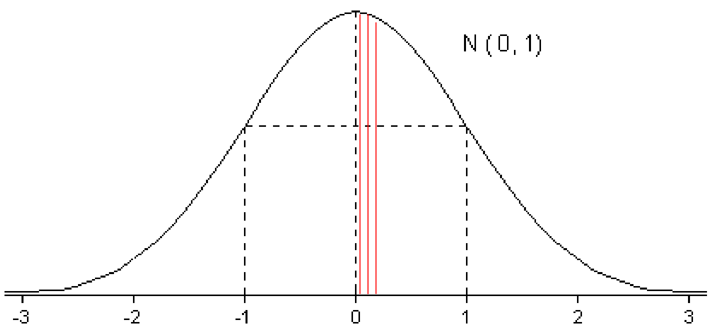
Definição: é a distribuição normal de uma variável Z que tem média igual a zero (μ=0) e desvio padrão igual a um (σ=1).

Função densidade de probabilidade de uma variável X que tem distribuição normal

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

Função densidade de probabilidade da variável Z que tem normal padrão

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{z^2}{2}} \quad , \text{ para } S_Z = (-\infty,+\infty)$$



A curva normal padrão foi dividida em pequenas tiras, cujas áreas foram calculadas e apresentadas numa tabela.

Na tabela da distribuição normal padrão, podemos encontrar as áreas correspondentes aos intervalos de 0 a z.

Tabela - Área sob a curva normal padrão de 0 a z, $P(0 \leq Z \leq z)$.

z	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
0,3	0,1179	0,1217	0,1255	0,1293	0,1331
0,4	0,1554	0,1591	0,1628	0,1664	0,1700
0,5	0,1915	0,1950	0,1985	0,2019	0,2054
0,6	0,2258	0,2291	0,2324	0,2357	0,2389
0,7	0,2580	0,2612	0,2642	0,2673	0,2704
0,8	0,2881	0,2910	0,2939	0,2967	0,2996

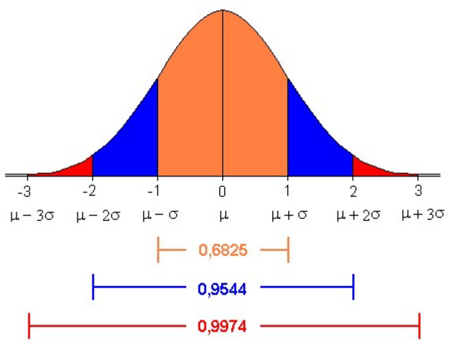
probabilidade

$P(0 < Z < 0,62) = 0,2324$

Profa. Clause Piana 107

⇨ Os valores negativos não são apresentados na tabela porque a curva é simétrica; assim, as áreas correspondentes a esses valores são exatamente iguais às dos seus simétricos positivos, por exemplo **$P(-1<Z<0)=P(0<Z<1)$**.

⇨ Na tabela da distribuição normal padrão, os valores de Z vão de 0 a 3,9. Este limite é estabelecido com base na quarta propriedade da distribuição normal.



Profa. Clause Piana 108

Exercício proposto:

Seja Z uma variável aleatória com distribuição normal padrão. Determine as seguintes probabilidades:

a) $P(0 < Z < 1{,}73)$

b) $P(0{,}81 < Z < +\infty)$

c) $P(-1{,}25 \leq Z \leq -0{,}63)$

⇨ Através da distribuição normal padrão é possível estudar qualquer variável X que tenha distribuição normal, com quaisquer valores para μ e σ.

⇨ Para utilizarmos os valores da tabela, devemos **padronizar** a variável X, ou seja, transformar X em Z.

$$X \sim N(\mu, \sigma^2)$$

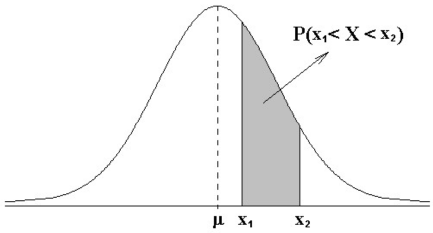
$$\downarrow \text{ transformar } \rightarrow \quad Z = \frac{X - \mu}{\sigma}$$

$$Z \sim N(0, 1)$$

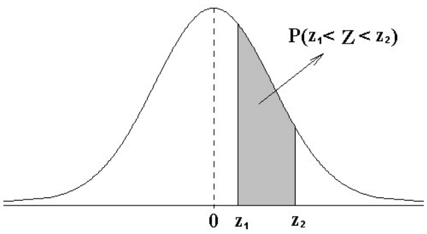
⇨ Após a transformação, procuramos na tabela a área compreendida entre **0** e **z**, que corresponderá a área entre μ e **x**.

A transformação muda as variáveis, mas não altera a área sob a curva.

$X \sim N(\mu, \sigma^2)$



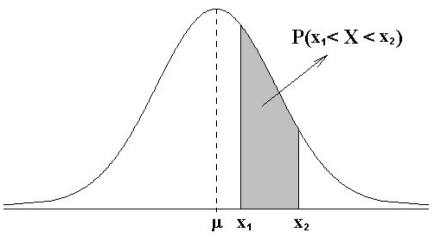
$Z \sim N(0, 1)$



$$P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

Como determinar um quantil x_p?

$$X_p = \mu + Z_p \, \sigma$$



Exercício proposto:

A resistência à compressão de amostras de cimento pode ser modelada por uma distribuição normal, com μ = 4000kg e σ = 80 kg/cm^2. Determine:

a) a probabilidade de que a resistência da amostra seja menor que 4150 kg/cm^2;

b) a probabilidade de que a resistência da amostra esteja entre 3950 e 4100 kg/cm^2;

c) o número de amostras com resistência inferior à $\mu+\sigma$ kg/cm^2, das 1200 amostras avaliadas no período de um mês.

Bibliografia

BUSSAB, W.O.; MORETTIN, P.A. **Estatística Básica**. São Paulo: Saraiva. 2006. 526p.

FERREIRA, D.F. **Estatística Básica**. Lavras: Editora UFLA, 2005, 664p.

FREUND, J.E., SIMON, G.A. **Estatística Aplicada. Economia, Administração e Contabilidade**. 9.ed., Porto Alegre: Bookman, 2000. 404p.

MEYER, P. L. **Probabilidade: aplicações à estatística**. Rio de Janeiro: LTC, 1976.

MLODINOW, L. **O andar do bêbado. Como o acaso determina nossas vidas**. Rio de Janeiro: Editora Zahar, 2009, 264p.

SILVEIRA JÚNIOR, P., MACHADO, A.A., ZONTA, E.P., SILVA, J.B. da **Curso de Estatística**. v.2, Pelotas: Universidade Federal de Pelotas, 1992. 234p.

Sistema Galileu de Educação Estatística. Disponível em: http://www.galileu.esalq.usp.br/topico.html