



ESCOLA SUPERIOR DE
TECNOLOGIA DA SAÚDE
DE LISBOA
INSTITUTO POLITÉCNICO DE LISBOA

QFE

T3

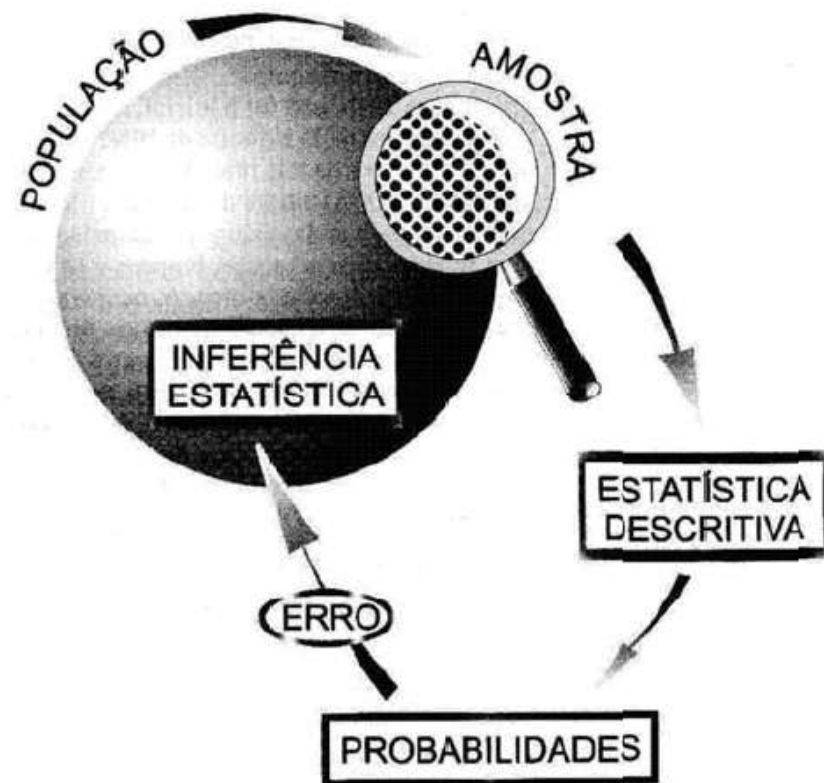
Carina Silva

carina.silva@estesl.ipl.pt

SUMÁRIO

Distribuição Normal

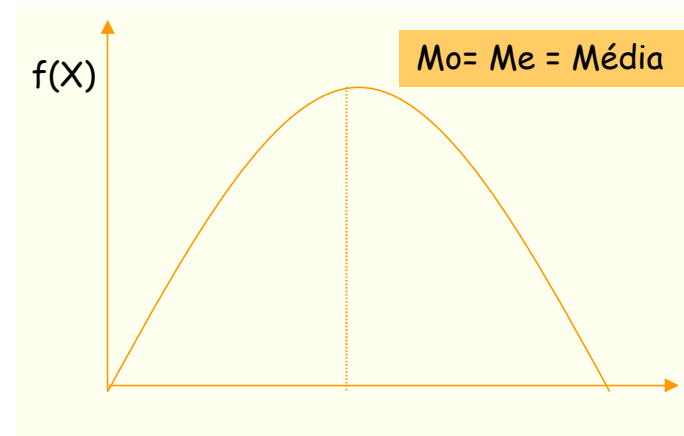
Estimação



Distribuição Normal ou Distribuição Gaussiana

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} ; x \in \mathbb{R}$$

$$F_X(x) = P(X \leq x) = P(X < x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$



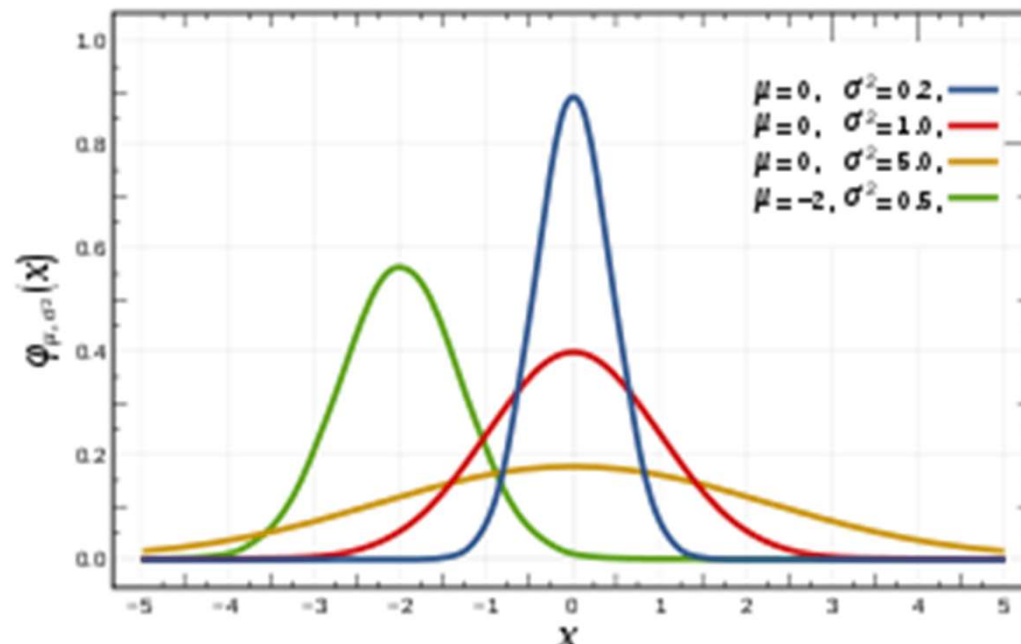
A distribuição normal é caracterizada por **2**
PARÂMETROS: média e desvio padrão

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

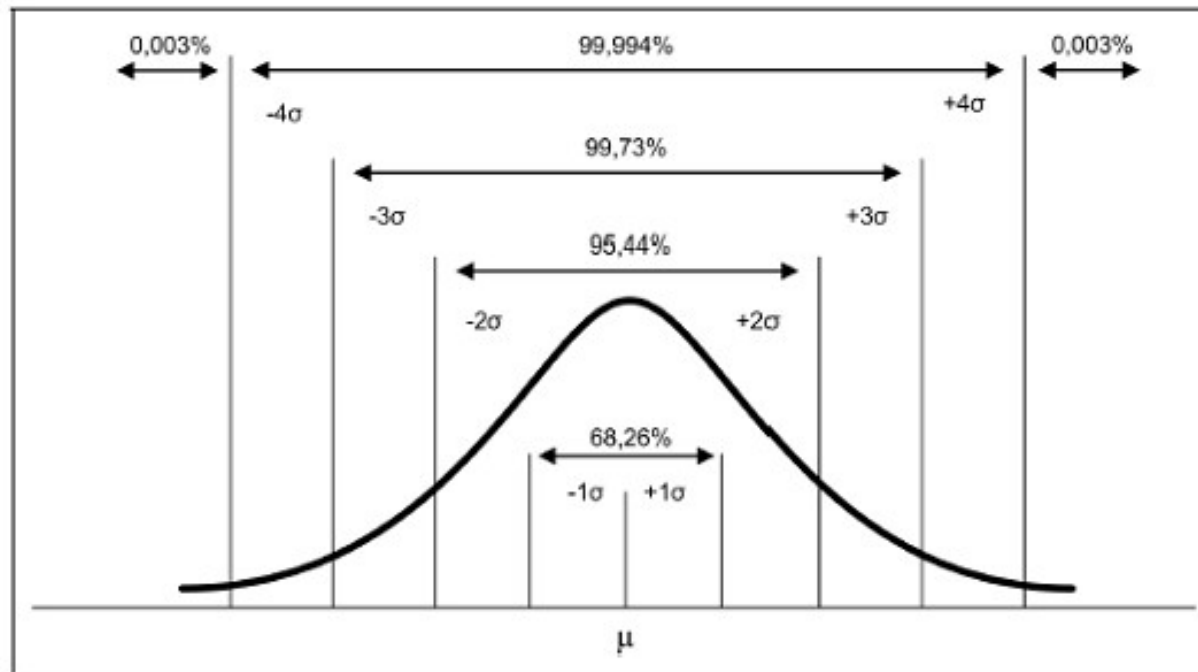
O ponto máximo da curva normal está na MÉDIA, que também é mediana e moda da distribuição

A média poderá ser um valor positiva, zero ou negativo.

A distribuição normal é **SIMÉTRICA**. Extremos tendem ao infinito.



As probabilidades da variável aleatória normal são dadas pelas áreas sob a curva. Área total = 1.
Área à esquerda da média é igual a área à direita (ambas = 0,5)



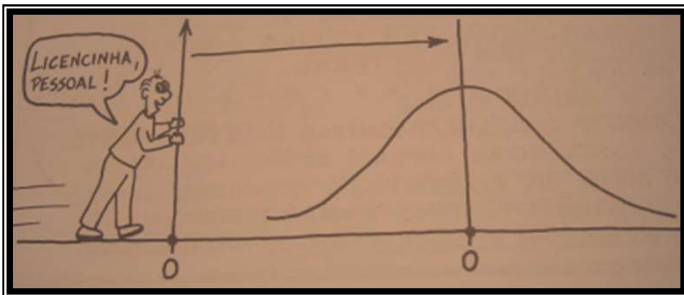
DISTRIBUIÇÃO NORMAL REDUZIDA (PADRÃO)

Distribuição normal padrão ou normal reduzida

$$\text{Se } X \sim \mathcal{N}(\mu, \sigma) \text{ então } Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

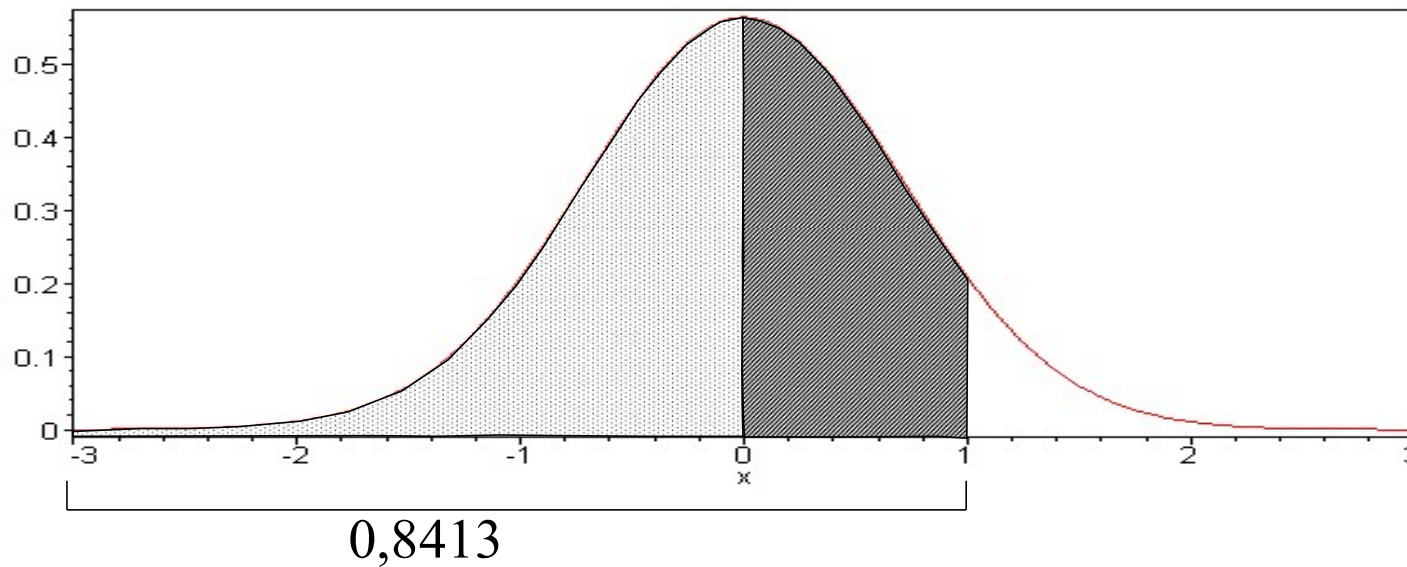
A f.d. da normal reduzida tem uma notação especial. Assim se Z for uma v.a. Normal standard apresentamos:

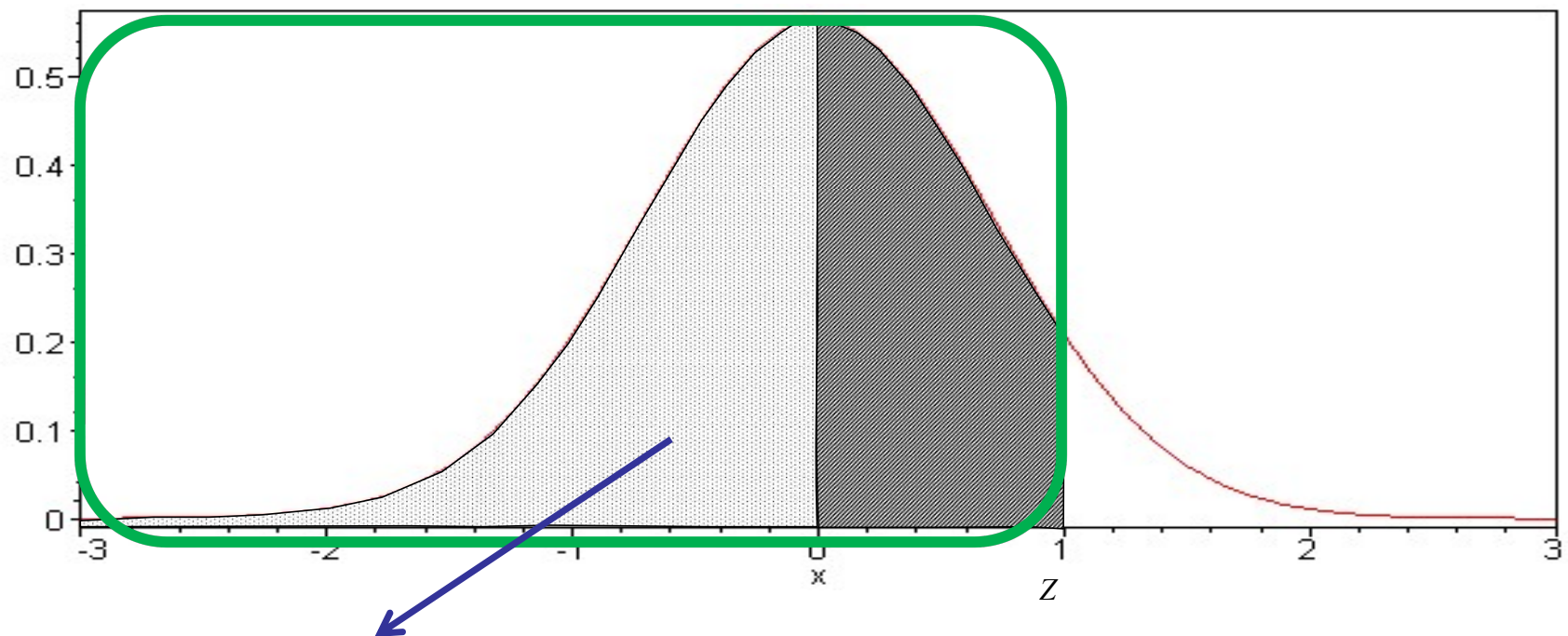
$$P(Z \leq z) = \Phi(z)$$



1) Seja X uma variável com distribuição normal de média 20 e desvio-padrão 4, determinar $P(X < 24)$:

$$P(X < 24) = P(Z < (24 - 20)/4) = \Phi(1) = 0.8413$$





E sabemos que essa área pontilhada tem probabilidade 0,5

$$P(X \leq x_1) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x_1 - \mu}{\sigma}\right) = P(Z \leq z_0)$$

$$P(X > x_0) = P\left(\frac{X - \mu}{\sigma} > \frac{x_0 - \mu}{\sigma}\right) = P(Z \geq z_0) = 1 - P(Z < z_0) =$$

$$= 1 - \phi(z_0)$$

$$P(x_0 < X < x_1) = P\left(\frac{x_0 - \mu_0}{\sigma_0} < \frac{X - \mu}{\sigma} < \frac{x_1 - \mu_0}{\sigma_0}\right) =$$

$$= P(z_0 < Z < z_1) = \phi(z_1) - \phi(z_0)$$

Propriedades:

$$\phi(-z_0) = 1 - \phi(z_0)$$

$$\phi(z_0) = \text{área à esquerda de } z_0$$

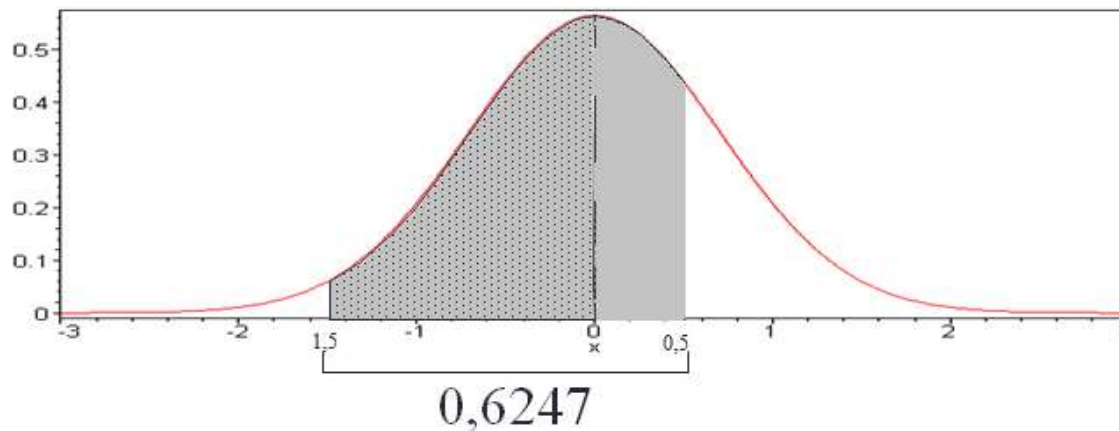
$$\phi(z_0) = 1 - \phi(-z_0)$$

x	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767

Seja X uma variável com distribuição normal de média 20 e desvio-padrão 4, determinar $P(14 < X < 22)$:

Seja X uma variável com distribuição normal de média 20 e desvio-padrão 4, determinar $P(14 < X < 22)$:

$$P((14-20)/4 < Z < (22-20)/4) = P(-1.5 < Z < 0.5) = \Phi(0.5) - \Phi(-1.5) = 0.6915 - (1 - \Phi(1.5)) = 0.6915 - 1 + 0.9332 = 0.6247$$



Quantis da normal

DENOTAMOS POR $z_{1-\alpha}$ O QUANTIL DE PROBABILIDADE $1-\alpha$ DA GAUSSIANA PADRÃO, OU SEJA $P(Z \leq z_{1-\alpha}) = 1-\alpha$

POR OUTRAS PALAVRAS $z_{1-\alpha} = \Phi^{-1}(1 - \alpha)$

A CONSULTA DA TABELA DA FUNÇÃO DISTRIBUIÇÃO NORMAL REDUZIDA FORNECE OS PERCENTIS DA NORMAL PADRÃO

EXEMPLO: $z_{0.95} = \Phi^{-1}(0.95) = 1.645$

Exercício

Um laboratório de análises clínicas utiliza reagentes que estão acondicionados em embalagens de vidro. Estas embalagens devem conter 300 ml de reagente. Considere que a quantidade de reagente presente em cada embalagem é uma variável aleatória que segue uma distribuição Normal com valor médio de 298 ml e desvio padrão 3 ml.

1. Qual é a probabilidade de uma embalagem de reagente, retirada ao acaso, conter no máximo 295 ml?
2. Calcule o quantil de ordem 0,90 desta distribuição e diga qual o seu significado.

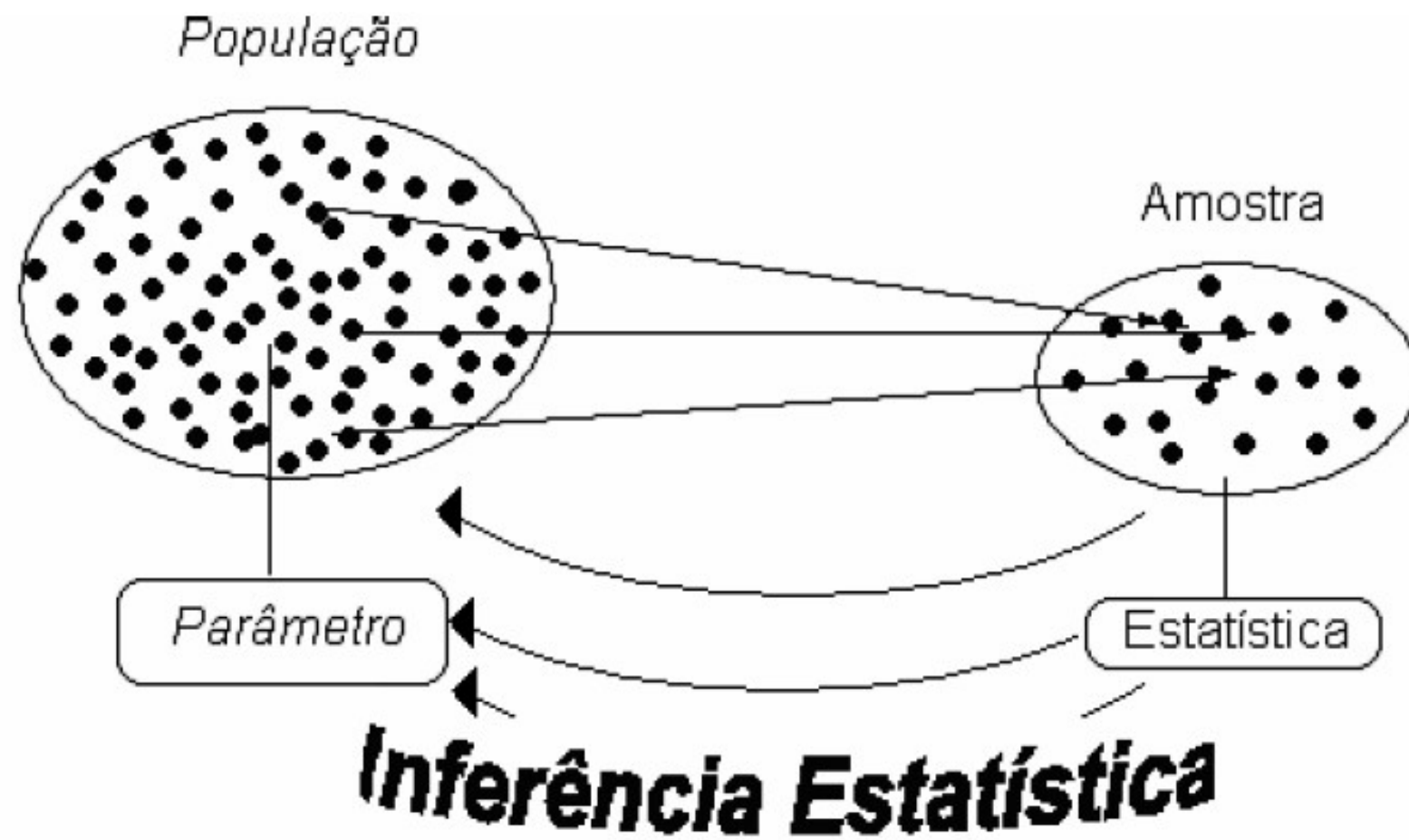
Inferência Estatística

O que é? Quando se utiliza? Para que serve?

É um processo de raciocínio indutivo, em que se procuram tirar conclusões indo do particular, para o geral. É um tipo de raciocínio contrário ao tipo de raciocínio matemático, essencialmente dedutivo.

Utiliza-se quando se pretende estudar uma **população**, **estudando só alguns** elementos dessa população, ou seja, uma **amostra**.

Serve para, a partir das propriedades verificadas na **amostra**, **inferir propriedades** para a **população**.



Estimativa Pontual



Sabe-se que o nível de mercúrio segue uma Distribuição Normal.

Diga qual é o valor do mercúrio metílico médio ingerido por todos os sujeitos desta população:

- Recorrendo a uma estimativa pontual
- Recorrendo a um intervalo de confiança de 95% e explicando como foram obtidos os resultados apresentados.

Descriptives

			Statistic	Std. Error
VAR00001	Mean		329,02	19,44
	95% Confidence Interval for Mean	Lower Bound	290,03	
		Upper Bound	368,00	
	5% Trimmed Mean		320,68	
	Median		287,50	
	Variance		20400,132	
	Std. Deviation		142,83	
	Minimum		123	
	Maximum		693	
	Range		570	
	Interquartile Range		115,25	
	Skewness		1,187	,325
	Kurtosis		1,130	,639

Estimação Intervalar

Na estimação por intervalos, em vez de se propor apenas um valor concreto para certo parâmetro da população, constrói-se um intervalo de valores $(w_1; w_2)$ que, com um certo grau de certeza, previamente estipulado, contenha o verdadeiro valor do parâmetro.

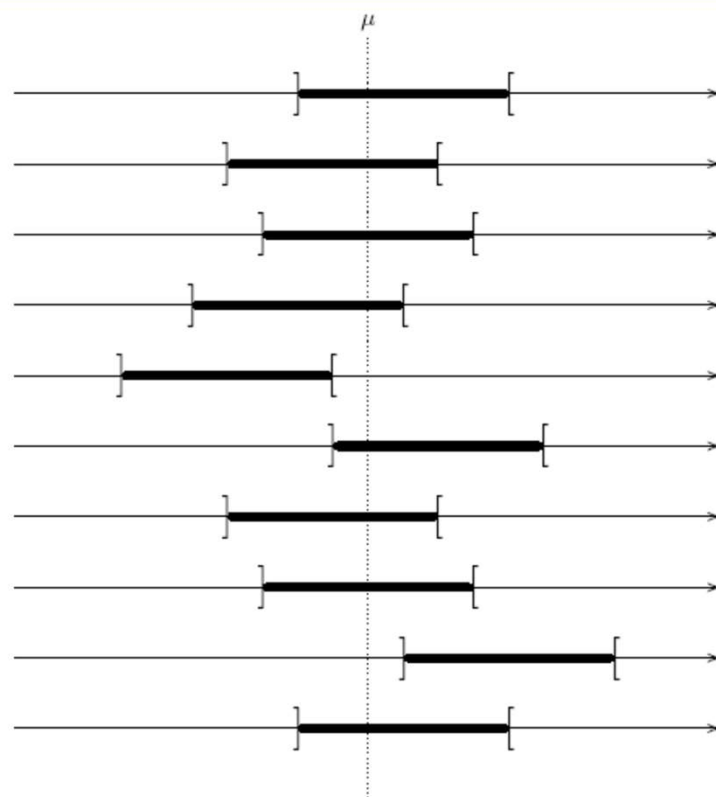
Em muitos casos, o intervalo é da forma $(\hat{\theta} - \varepsilon; \hat{\theta} + \varepsilon)$, sendo $\hat{\theta}$ uma estimativa para o parâmetro de interesse θ , e ε é considerado uma medida de precisão ou medida do erro inerente à estimativa $\hat{\theta}$. Usualmente, ε é designado por erro de estimativa ou margem de erro (absoluta). Desta forma, este método de estimação incorpora a confiança que se pode atribuir às estimativas.

O **erro de estimativa** corresponde ao erro máximo que, com a confiança especificada, se pode cometer na estimativa de θ . Nos I. C. centrados em torno da estimativa $\hat{\theta}$, o erro de estimativa corresponde à semiampitude do I. C.

A amplitude do I. C. pode ser reduzida se:

- Aumentar a dimensão da amostra (n);
- Mantendo a dimensão da amostra, se diminuir o grau de confiança $(1 - \alpha)$.

Para cada amostra que se observa, obtém-se em geral um intervalo de confiança diferente para o mesmo parâmetro. Quando dizemos que um intervalo tem confiança $1 - \alpha$, estamos a dizer que se observarmos muitas amostras distintas, digamos m , os intervalos que se obtêm contêm o verdadeiro valor do parâmetro $(1 - \alpha) \times 100\%$ das vezes, quando $m \rightarrow \infty$.



Sabe-se que o nível de mercúrio segue uma Distribuição Normal.

Diga qual é o valor do mercúrio metílico médio ingerido por todos os sujeitos desta população:

- Recorrendo a uma estimativa pontual
- Recorrendo a um intervalo de confiança de 95% e explicando como foram obtidos os resultados apresentados.

Descriptives

			Statistic	Std. Error
VAR00001	Mean		329,02	19,44
	95% Confidence Interval for Mean	Lower Bound	290,03	
		Upper Bound	368,00	
	5% Trimmed Mean		320,68	
	Median		287,50	
	Variance		20400,132	
	Std. Deviation		142,83	
	Minimum		123	
	Maximum		693	
	Range		570	
	Interquartile Range		115,25	
	Skewness		1,187	,325
	Kurtosis		1,130	,639

$\mu \in]290,03; 368[$ com uma confiança de 95% e uma margem de erro de 38,99

Intervalo de confiança para o valor médio (μ)

→ se $n > 30$ (grandes amostras)

⇒ se σ conhecido, X segue uma **distribuição Normal** ou **arbitrária**

$$\bar{x} - z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}$$

⇒ se σ desconhecido, qualquer X segue uma distribuição Normal ou arbitrária

$$\bar{x} - z_{\left(1-\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\left(1-\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}}$$

→ se $n < 30$ (pequenas amostras) e X é uma variável aleatória com **distribuição normal**

⇒ se σ conhecido

$$\bar{x} - z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}$$

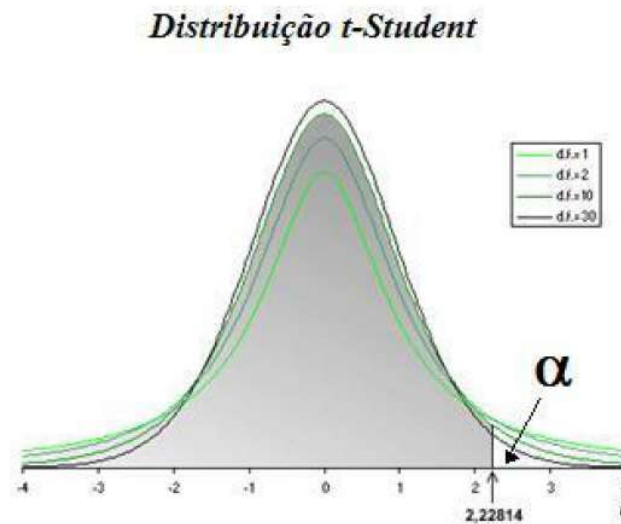
⇒ se σ desconhecido

$$\bar{x} - t_{n-1; 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1; 1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

DISTRIBUIÇÃO T-STUDENT

PROPRIEDADES:

- $E(T)=0$
- $VAR(T)=n/(n-2)$ PARA $n>2$, APROXIMANDO-SE DE UM À MEDIDA QUE n AUMENTA
- É SIMÉTRICA EM RELAÇÃO AO SEU VALOR MÉDIO E É UNIMODAL
- PARA DIFERENTES VALORES DO N.º DE G.L. TEMOS DIFERENTES DISTRIBUIÇÕES DE t
- ASSUME VALORES ENTRE $-\infty$ E $+\infty$
- APROXIMA-SE DA DISTRIBUIÇÃO NORMAL À MEDIDA QUE n CRESCE
- É MAIS ACHATADA E TEM AS CAUDAS MAIS PESADAS EM RELAÇÃO À DISTRIBUIÇÃO NORMAL



$$P(T_3 < 3,5)$$

Distribuição t de Student - t_n

Os valores tabelados correspondem aos pontos x tais que: $P(t_n \leq x)$

n	P($t_n \leq x$)							
	0,600	0,750	0,900	0,950	0,975	0,990	0,995	0,9995
1	0,325	1,000	3,078	6,314	12,706	31,821	63,657	636,619
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	31,598
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	12,924
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	8,610
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	6,869
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,959
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	5,408
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	5,041
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,781
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,587
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,437

Exercício

Considere a seguinte informação relativa à pressão arterial sistólica (BPs1) de indivíduos de ambos os géneros. A informação disponível permite-lhe afirmar que:

Descriptives				
Género			Statistic	Std. Error
BPs1	Masculino	Mean	114,563	3,7438
		95% ConfidenceInterval	106,583	
		for Mean	122,543	
	Feminino	Mean	109,612	1,8346
		95% ConfidenceInterval	105,461	
		for Mean	113,762	

- (A) A estimativa pontual da pressão arterial sistólica é mais precisa no grupo masculino.
- (B) A estimativa da pressão arterial sistólica por intervalo de confiança apresenta uma margem de erro menor no grupo masculino.
- (C) Para um nível de significância de 1% podemos afirmar que a pressão arterial sistólica no grupo das mulheres difere significativamente de 114,5.
- (D) Nenhuma das opções anteriores.

Para refletir!!!!

Que é que significa aumentar a confiança?

Efeito da dimensão da amostra na amplitude do intervalo de confiança

Margem de erro de uma sondagem, como diminuir.

Qual a influência da dimensão da População na determinação de um intervalo de confiança?

Tem sentido utilizar um intervalo de confiança para estimar um parâmetro quando a amostra coincide com a População?

Tem sentido construir um intervalo de confiança com base numa amostra enviesada?

TPC

Diga se são verdadeiras ou falsas as afirmações seguintes:

- a) Se uma amostra é aleatória qualquer elemento da população tem a mesma probabilidade de pertencer à amostra.
- b) Uma estatística é um número que se calcula a partir da amostra.
- c) Os parâmetros utilizam-se para estimar estatísticas.
- d) A média populacional é um parâmetro.
- e) Reduz-se o enviesamento, aumentando a dimensão da amostra.