Inferência Estatística

Modelos Probabilísticos

Prof. Antonio Estanislau Sanches
2017

Modelos Probabilísticos Padrões

Veremos adiante alguns **Modelos Probabilísticos Padrões**, que nos auxiliarão em diversas situações práticas.

MODELOS PROBABILÍSTICOS PADRÕES

1. DISTRIBUIÇÃO BINOMIAL

Trata-se de uma distribuição de probabilidade adequada aos ***experimentos*** que apresentam ***apenas dois resultados*** possíveis: ***sucesso*** ou ***fracasso***. Fornece a base para ***inferências sobre proporções***.

HIPÓTESES DO MODELO BINOMIAL

1. O experimento é repetido **n vezes** nas mesmas condições.

2. Os **resultados** das repetições são **independentes**, ou seja, uma repetição não interfere nas subseqüentes.

3. Cada repetição admite apenas dois resultados: **sucesso** ou **fracasso**.

4. As probabilidades de sucesso "**p**" e de insucesso "**q**" **(q=1-p)** se mantêm **constantes** durante as repetições.

Por exemplo:

a) Lançar uma moeda 5 vezes e observar o número de caras.

b) Numa linha de produção, observar 10 itens tomados ao acaso e verificar o número de defeituosos.

c) Verificar o número de bits que não estão afetados por ruído num pacote com n bits.

MODELOS PROBABILÍSTICOS PADRÕES

1. DISTRIBUIÇÃO BINOMIAL

Define-se a **VARIÁVEL BINOMIAL X** como o número de sucessos em **n repetições** do experimento. A expressão geral da Distribuição Binomial é:

$$P\{X = x\} = \binom{n}{x} p^x q^{n-x}$$

NOTAS:

1. O nome **BINOMIAL** se deve ao fato da expressão acima corresponder ao termo geral do desenvolvimento do **BINÔMIO DE NEWTON**.

2. Para **p=0,5** a distribuição é **simétrica**. Para **P<0,5**, a distribuição tem inclinação para a **direita**.

3. No caso de **n grande** (**n ≥ 30**) e **p não muito pequena nem muito grande** (valores centrais, com alguns autores recomendando **np>5** e **nq>5**), a **DISTRIBUIÇÃO BINOMIAL** pode ser aproximada pela **DISTRIBUIÇÃO DE POISON**, que será vista adiante.

MODELOS PROBABILÍSTICOS PADRÕES

1. DISTRIBUIÇÃO BINOMIAL

PRINCIPAIS CARACTERÍSTICAS DA DISTRIBUIÇÃO BINOMIAL

Média $\mu_x = E\{X\} = np$ e **Variância $\sigma^2_x = V\{X\} = npq$**.

No EXCEL

Função DISTRBINOM(núm_s ; tentativas ; probabilidade_s ; cumulativo), onde

núm_s: número de sucessos

tentativas: número de tentativas independentes

probabilidade_s: probabilidade de sucesso em uma tentativa

cumulativo: um valor lógico que define o tipo de distribuição:

VERDADEIRO (**1**): retorna o valor da função de probabilidade acumulada **$P(X \leq num\_s)$**

FALSO (**0**): retorna o valor da função de probabilidade no ponto num_s: **$P(X = num\_s)$**

EXEMPLO: Uma moeda não viciada é lançada 5 vezes. Encontre a probabilidade de:

a) dar exatamente 3 caras

b) pelo menos uma cara

c) no máximo 2 caras

d) calcular o valor esperado e o desvio padrão

MODELOS PROBABILÍSTICOS PADRÕES

1. DISTRIBUIÇÃO BINOMIAL

EXEMPLO: Uma moeda não viciada é lançada 5 vezes. Encontre a probabilidade de:

a) dar exatamente 3 caras

b) pelo menos uma cara

c) no máximo 2 caras

d) calcular o valor esperado e o desvio padrão

SOLUÇÃO: Seja X a variável **BINOMIAL** com os parâmetros: **n=5**, **p=1/2** (e portanto **q=1/2**).

a) Desejamos ***P{X = 3}*** = C(5,3)x$(1/2)^3$x$(1/2)^2$ ⇨ $P\{X=3\} = \frac{5!}{3!\,2!} \times \left(\frac{1}{2}\right)^3 \times \left(\frac{1}{2}\right)^2$ e portanto

P{X = 3} = 10 x $(1/2)^5$ = 10/32 = **31,25%**

No **EXCEL**, a chamada à função ***DISTRBINOM(3;5;50%;0)*** fornece o valor 0,3125.

b) Desejamos ***P{X ≥ 1}*** que é o mesmo que 1 – P{X < 1}, equivalente a 1 – P{X = 0} = 1 – 0,03125 = **96,88%**

No **EXCEL**, a função ***1-DISTRBINOM(0;5;50%;0)*** fornece o valor 0,96875.

c) Desejamos ***P(X ≤ 2)*** que equivale a P{X = 0} + P{X = 1} + P{X = 2} = **50%**

No **EXCEL**, a chamada à função ***DISTRBINOM(2;5;50%;1)*** fornece o valor 0,5000.

d) **E{X} =**np e portanto **E{X} = 2,5** caras, e **V{X} =** npq = 5/4 = **1,25** . Logo **DesvPad{X} = 1,12** caras.

MODELOS PROBABILÍSTICOS PADRÕES

2. DISTRIBUIÇÃO DE POISSON

A distribuição de Poisson é uma aproximação da **DISTRIBUIÇÃO BINOMIAL**, quando o número de repetições *n* é muito grande (*tende ao infinito*) e a probabilidade de sucesso *p* do evento, em uma tentativa, é muito pequena (*tende a zero*), mantendo-se constante, finita e não nulo o produto entre *n* e *p* (média dos sucessos).

Considere as situações em que se avalia o número de ocorrências de um determinado evento por unidade de tempo, de comprimento, de área ou de volume (*genericamente denominados de área de oportunidade*). Em muitos casos, conhece-se o número de sucessos, mas às vezes é muito difícil ou até mesmo impossível determinar o número de fracassos. Imagine por exemplo, o número de automóveis que passam por uma esquina: pode-se anotar o número de veículos que passaram num determinado intervalo de tempo, mas não se pode determinar quantos deixaram de passar.

MODELOS PROBABILÍSTICOS PADRÕES

2. DISTRIBUIÇÃO DE POISSON

A distribuição de Poisson é aplicada nos tipos de situações em que nos interessa o ***número de vezes em que um evento pode ocorrer durante um intervalo de tempo*** ou em determinado ambiente físico (*área de oportunidade*). Tomando como referência o número de ocorrências em determinado intervalo de tempo, em um processo de Poisson podem ser observados eventos discretos num intervalo de tempo, de tal forma que, reduzindo suficientemente este intervalo, tenhamos:

HIPÓTESES DO MODELO DE POISSON

1. A probabilidade de observar apenas um sucesso no intervalo é estável.
2. A probabilidade de observar mais que um sucesso no intervalo é zero.
3. A ocorrência de um sucesso em qualquer intervalo é independente da ocorrência de sucesso em qualquer outro intervalo.

MODELOS PROBABILÍSTICOS PADRÕES

2. DISTRIBUIÇÃO DE POISSON

A distribuição de Poisson é caracterizada apenas pelo parâmetro λ, que representa o **valor esperado** ou **média**, do **número de sucessos** por **intervalo t**. Em outras palavras, λ é a taxa de ocorrência dos eventos no intervalo de tempo.

A **função de probabilidade** da distribuição de Poisson é:

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

e é uma constante (*base do logarítmo neperiano*) valendo 2,718...

$\boldsymbol{\lambda}$ é o número esperado de sucessos no intervalo considerado

x é o número de sucessos (x = 0, 1, 2, ...,∞.)

Modelos Probabilísticos Padrões

2. Distribuição de Poisson

Principais Características da Distribuição de Poisson

Média $\mu_x = E\{X\} = \lambda$ e **Variância $\sigma^2_x = V\{X\} = \lambda$.**

No Excel

DIST.POISSON(x ; média ; cumulativo), onde

x: número de sucessos

média: valor esperado no intervalo

cumulativo: um valor lógico que define o tipo de distribuição:

VERDADEIRO (1): retorna o valor da função de probabilidade acumulada **P**($X \leq x$)

FALSO (0): retorna o valor da função de probabilidade no ponto x: **P**($X = x$)

Exemplo: As consultas a um banco de dados ocorrem de forma independente e aleatória, à base de 3 consultas por minuto. Calcule as probabilidades:

a) no próximo minuto ocorrerem exatamente 3 consultas

b) no próximo minuto ocorrerem menos de 3 consultas

c) nos próximos dois minutos, ocorrerem mais do que 5 consultas

MODELOS PROBABILÍSTICOS PADRÕES

2. DISTRIBUIÇÃO DE POISSON

EXEMPLO: As consultas a um banco de dados ocorrem de forma independente e aleatória, à base de 3 consultas por minuto. Calcule as probabilidades:

a) no próximo minuto ocorrerem exatamente 3 consultas

b) no próximo minuto ocorrerem menos de 3 consultas

c) nos próximos dois minutos, ocorrerem mais do que 5 consultas

SOLUÇÃO: Seja X a variável **Poisson** com ocorrência média de **3 consultas por minuto** (λ**=3**)

a) Desejamos **P(X = 3)** = $[e^{-3} . 3^3]/3!$ = **22,4%**

No **EXCEL**, a chamada à função ***POISSON(3;3;0)*** fornece o valor 0,22404.

b) Desejamos **P(X < 3) = P(X ≤ 2)** = P(X=0) + P(X=1) + P(X=2) = **42,32%**

No **EXCEL**, a chamada à função ***POISSON(2;3;1)*** fornece o valor 0,42319.

c) Observe que a unidade de tempo alterou de 1 para 2 minutos. Como a taxa média é de 3 por minuto, então em dois minutos teremos **λ=6**. Desejamos assim **P(X > 5) = 1 – P(X ≤ 5)** = 1 – 0,44568 = **55,43%**

No **EXCEL**, a chamada à função ***1 - POISSON(5;6;1)*** fornece o valor 0,55432.

Modelos Probabilísticos Padrões

3. Distribuição Exponencial

Todo fenômeno aleatório descrito por esta distribuição se caracteriza pela total imprevisibilidade e assimetria, mesmo que se conheça seu passado, por isso ela tem grande aplicabilidade em sistemas de filas. A distribuição exponencial é muito utilizada na modelagem de tempos decorridos entre dois eventos, particularmente se estes forem causados por grande quantidade de fatores independentes.

A distribuição Exponencial também é importante na teoria da confiabilidade, pois serve para descrever as características da vida útil de certo componente, principalmente, os eletrônicos. Destaca-se ainda, para esta aplicação, que esta distribuição não tem "memória", isto é, pode ser usada para modelos de duração de vida que não desgastam com o tempo. Sendo assim, um componente novo não é mais confiável do que outro que já esteja em funcionamento.

MODELOS PROBABILÍSTICOS PADRÕES

3. DISTRIBUIÇÃO EXPONENCIAL

Se o número de ocorrências de um evento por unidade de tempo seguir a Distribuição de Poisson, então automaticamente a ***distribuição do intervalo de tempo entre ocorrências*** do evento segue a **DISTRIBUIÇÃO EXPONENCIAL**. Assim, se λ é a taxa à qual ocorrem eventos de Poisson, a distribuição do tempo entre chegadas sucessivas é expressa por:

$$\boxed{f(x) = \lambda e^{-\lambda x}, x > 0}$$

A distribuição exponencial também é caracterizada apenas pelo parâmetro λ, que representa a taxa de ocorrência dos eventos no intervalo de tempo.

A **função de probabilidade acumulada** da distribuição exponencial é:

$$\boxed{P\{X \leq a\} = \int_0^a \lambda e^{-\lambda x} dx = 1 - e^{-\lambda a}}$$

onde:

e é constante de Euler (*base do logarítmo neperiano*) valendo 2,71828...

λ é a taxa de ocorrência do evento por unidade de tempo

x é o tempo entre chegadas sucessivas

MODELOS PROBABILÍSTICOS PADRÕES

3. DISTRIBUIÇÃO EXPONENCIAL

PRINCIPAIS CARACTERÍSTICAS DA DISTRIBUIÇÃO EXPONENCIAL

Média $\mu_x = E\{X\} = 1/\lambda$ e Variância $\sigma^2_x = V\{X\} = 1/\lambda$.

NO EXCEL

Função DISTEXPON(x ; lambda ; cumulativo), onde

x: intervalo de tempo entre chegadas sucessivas

lambda: inverso da taxa de chegada por unidade de tempo

cumulativo: um valor lógico que define o tipo de distribuição:

VERDADEIRO (1): retorna o valor da função de probabilidade acumulada **P**(X ≤ x)

FALSO (0): retorna o valor da função densidade de probabilidade no ponto x: **P**(X **=** x)

EXEMPLO: Carros chegam a um posto de gasolina aleatoriamente a cada 2 minutos em média. Determine a probabilidade de o tempo entre chegadas não exceder 1 minuto.

SOLUÇÃO: Seja X a variável Exponencial que representa o tempo entre chegadas. Queremos determinar:

P{X≤1} com λ=1/2, o que implica em calcular $\int_0^1 (1/2)e^{-(1/2)}$. Note que primeiramente é necessário ***converter a taxa de chegadas por unidade de tempo*** para a mesma unidade de tempo empregada no valor da variável do problema.

No **EXCEL**, a chamada à função ***DISTEXPON(1;1/2;1)*** fornece o valor 0,3935 ou **39,35%**.
Note que se x=1/2 (de 2 minutos) e λ=1 (em 2 minutos), ***DISTEXPON(1/2;1;1)*** também retorna **39,35%**.

MODELOS PROBABILÍSTICOS PADRÕES

3. DISTRIBUIÇÃO EXPONENCIAL

Existe uma forte relação entre a distribuição Exponencial e a distribuição de Poisson. Se uma variável aleatória *x* de Poisson tem média de λ ocorrências em um intervalo de tempo, então o intervalo de tempo T entre ocorrências segue uma distribuição exponencial e tem média de $1/\lambda$.

EXEMPLO: Uma máquina opera, em média, durante 2 horas sem necessitar de paradas para reajustes de configuração. Considerando-se que esse fenômeno possa ser representado por uma distribuição Exponencial, qual é a probabilidade dela funcionar durante 1h sem paradas?

$P(x \geq 1h) = ?$

$E(x) = 1/\lambda = 2 \rightarrow \lambda = 0{,}5$

$P(x \geq 1h) = e^{-0,5*1} = 0{,}6065 = 60{,}65\%$

Usando o Excel:

P(X≥1) = 1 - DISTR.EXPON(1;0,5;1) = 60,65%

Modelos Probabilísticos Padrões

4. Distribuição Normal

É considerada a distribuição de probabilidades mais importante, pois ***permite modelar uma infinidade de fenômenos naturais***, e além disso, possibilita realizar aproximações para calcular probabilidades de muitas variáveis aleatórias que têm outras distribuições, tais como a **Binomial** (**$n \geq 30$, $np>5$** e **$nq>5$**). Também conhecida como distribuição de **Gauss**, é muito importante também na inferência estatística.

A distribuição Normal é caracterizada por uma **Função de Densidade de Probabilidade** cujo gráfico descreve uma ***curva em forma de sino***, que evidencia maior probabilidade.

Função Densidade de Probabilidade da ***Distribuição NORMAL***

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

onde:

μ = **média** da distribuição

σ = **desvio padrão** da distribuição

π e e são constantes (3,1416... e 2,718...)

MODELOS PROBABILÍSTICOS PADRÕES

4. DISTRIBUIÇÃO NORMAL

PARÂMETROS DA DISTRIBUIÇÃO NORMAL

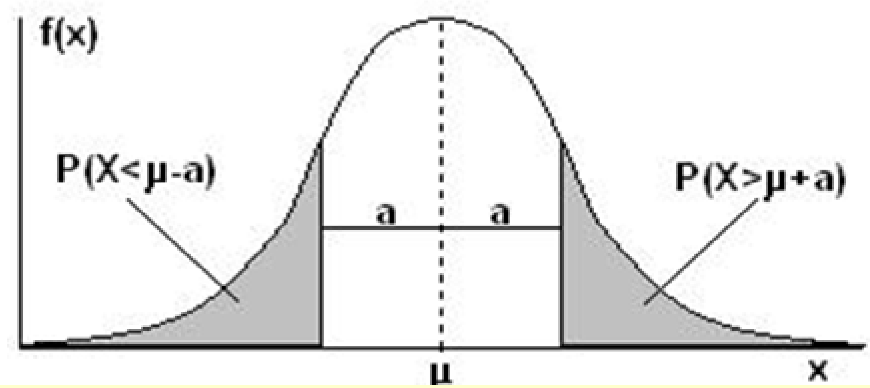
Média ou **Valor Esperado** $\mu_x = E\{X\} = \mu$ e **Variância** $\sigma^2_x = V\{X\} = \sigma^2$

PRINCIPAIS CARACTERÍSTICAS:

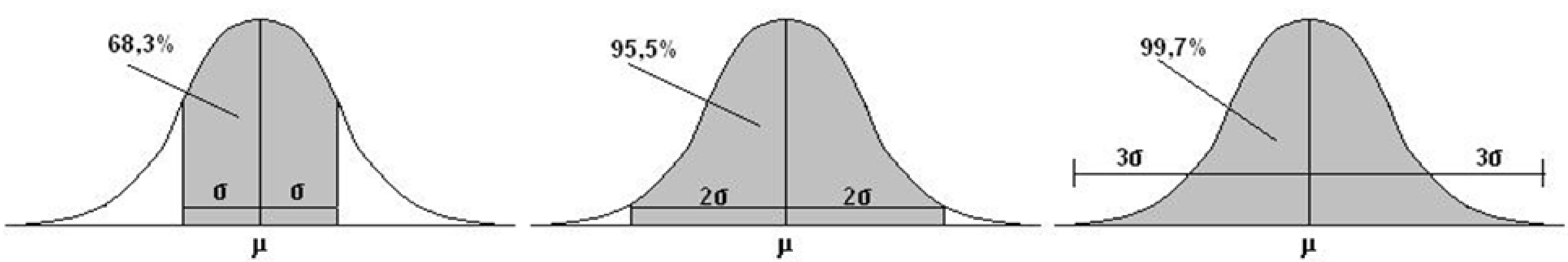
1. Teoricamente, a curva prolonga-se de $-\infty$ a $+\infty$, sendo que **lim f(x) =0** para x tendendo a $\pm\infty$.

2. A **área total** sob a curva é igual a 1, ou seja: $\boxed{\int_{-\infty}^{+\infty} f(x)dx = 1}$

3. A curva é simétrica em torno de μ, o que faz com que **média = mediana = moda**. Adicionalmente, temos também que **P(X < μ - a) = P(X > μ + a)**.



4. A curva tem dois pontos de inflexão, respectivamente em **μ-σ** e **μ+σ** . Cerca de **68%** dos valores recaem no **intervalo de um desvio padrão** de cada lado da média, **95%** recaem no **intervalo média ± 2 desvios** e **99,7%** recaem no **intervalo média ± 3 desvios**.



MODELOS PROBABILÍSTICOS PADRÕES

4. DISTRIBUIÇÃO NORMAL

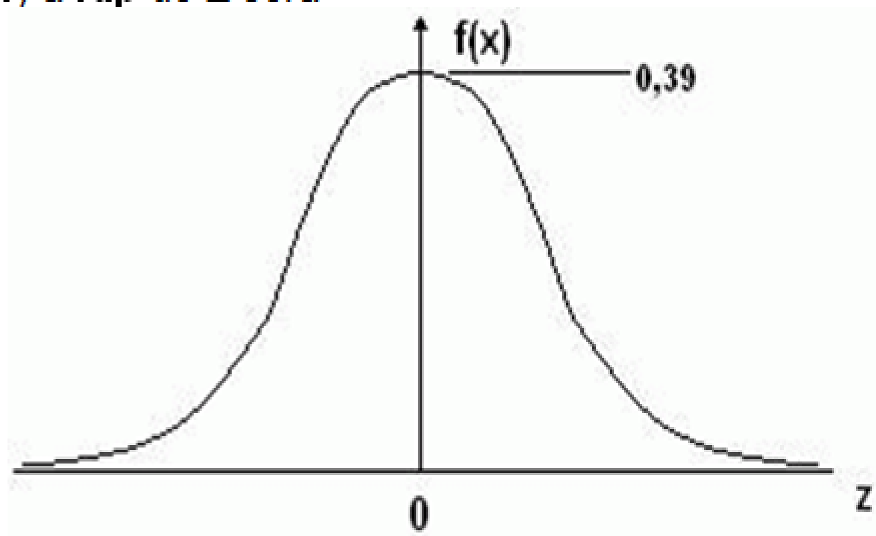
Considerando a enorme dificuldade de calcularmos probabilidades pela integração da Função de Densidade de Probabilidade (**fdp**) para as infinitas combinações de valores de **μ e σ**·, utiliza-se a **DISTRIBUIÇÃO NORMAL PADRÃO** ou **REDUZIDA**, definida conforme a seguir.

DISTRIBUIÇÃO NORMAL PADRÃO

Seja **Z** a variável com distribuição **normal** com **média = 0** e **variância = 1**, geralmente denotada por **N(0;1)**. Neste caso (lembrando que **desvio-padrão = variância = 1**) a **fdp** de Z será

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

com a forma:



Observe-se a conveniência de termos a **média igual a zero** e o **desvio padrão igual a 1**, fazendo com que esta distribuição passe a representar os valores de z_i como **número de desvios** em relação à média (origem). Assim, esta distribuição nos permite trabalhar com **valores *relativos* de desvios em relação à média**.

MODELOS PROBABILÍSTICOS PADRÕES

4. DISTRIBUIÇÃO NORMAL

TRANSFORMAÇÃO DE UMA DISTRIBUIÇÃO NORMAL $N(\mu;\sigma^2)$ PARA A NORMAL PADRÃO (OU REDUZIDA) N(0;1)

Qualquer distribuição normal com média **μ** e desvio padrão **σ pode ser transformada**, para efeito de cálculo de probabilidades, na distribuição **normal padrão**, através de uma mudança de variável conforme a seguir.

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

TABELAS DA FUNÇÃO NORMAL PADRÃO

Há vários tipos de **tabelas** que fornecem as áreas (probabilidades) sob a curva **NORMAL PADRÃO**. O tipo mais comum é a tabela de **FAIXA CENTRAL.** Este tipo de tabela fornece a área sob a curva normal padrão entre z=0 e qualquer valor positivo de z. A simetria em torno de z=0 permite-nos obter a área entre quaisquer valores de z, sejam positivos ou negativos, não sem razoável esforço na identificação correta de intervalos.

FUNÇÕES DO EXCEL PARA A DISTRIBUIÇÃO NORMAL

O **Excel** disponibiliza as unções para cálculos com a **Distribuição Normal**:

Função DIST.NORMP.N(z), onde, **z**: valor da VA Normal Padrão ou Reduzida.

Esta função retorna a ***probabilidade*** **P(-∞<Z< z) = P(Z< z)**, p/ **qualquer valor de z**.

Para um intervalo genérico **P(a<z<b)**, pode-se aplicar F(b) – F(a) diretamente:

P(a<z<b) = DIST.NORMP.N(b) – DIST.NORMP.N(a)

Para **P(z>a)**, usa-se **1 – P(z<a)** e portanto **P(z>a) = 1 – DIST.NORMP.N(a)**.

4. Distribuição Normal

Funções do Excel para a Distribuição Normal

Exemplos do emprego da **Função DIST.NORMP.N(z)**:

a) Calcule P(z < 0,85) ► DIST.NORMP.N(0,85;1) ► **= 0,8023** (ou **80,23%**).

b) Calcule P(0 < z < 1,25) ► DIST.NORMP.N(1,25;1)-DIST.NORMP.N(0;1) ► **= 0,3944** (ou **39,44%**).

c) Calcule P(z>2,39) ► 1-DIST.NORMP.N(2,39;1) ► **= 0,0084** ou **0,84%.**

d) Calcule P(-2,55 < z < 1,2) ► DIST.NORMP.N(1,2;1)-DIST.NORMP.N(-2,55;1) ► **= 0,8795** ou **87,95%.**

Função DIST.NORM.N(x ; média ; desv_padrão ; cumulativo),
Onde **x**: valor da VA Normal; **média**: média da VA X ; **desv_padrão**: desvio padrão da VA X ; **cumulativo**: um valor lógico que define o tipo de distribuição: **VERDADEIRO (1)**: retorna o valor da função de distribuição acumulada (FDA) **F(x) = P(X ≤ x**) e **FALSO (0)**: retorna o valor da função densidade de probabilidade (fdp) no ponto x: **f(x)**
Essa é a função mais completa para tratamento de distribuição normal. No caso de média=0, desvio=1 e cumulativo=1(*verdadeiro*), esta função retorna o mesmo valor da DIST.NORMP.N

Modelos Probabilísticos Padrões

Funções do Excel para a Distribuição Normal

Função INV.NORMP.N(probabilidade),

Retorna o valor **z** da VA Normal Padrão, abaixo do qual se tem a probabilidade informada. É o **inverso** da função ***DIST.NORMP.N*(z)**

Exemplo do emprego da **Função INV.NORMP.N (p)**:

Sendo, P(z<a) = 0,3015 ache "**a**" ► INV.NORMP.N(0,3015)► **= -0,524401**

Função INV.NORM.N(probabilidade ; média ; desv_padrão)

Como no caso acima, é o inverso da função geral DIST.NORM.N(), aplicável a qualquer VA Normal X, desde que conhecidos sua média e desvio padrão.

Função PADRONIZAR(x ; média ; desv_padrão)

Retorna o desvio padrão normalizado z, considerando os argumentos x, média e desvio padrao, utilizando a fórmula já apresentada:

$$z = \frac{x - \mu}{\sigma}$$

Exemplo do emprego da **PADRONIZAR(x ; média ; desv_padrão)**:

A altura dos alunos de uma escola é normalmente distribuída com média 1,60 m e desvio padrão 0,30 m. Calcule a probabilidade de um aluno medir entre 1,50 m e 1,80 m.

z1 = (1,50 – 1,60)/0,30 ► PADRONIZAR(1,5;1,6;0,3) ► **= -0,33**

z2 = (1,80 – 1,60)/0,30 ► PADRONIZAR(1,8;1,6;0,3) ► **= 0,67**

Logo: P(-0,33 < z < 0,67) ► DIST.NORMP.N(0,67;1)-DIST.NORMP.N(-0,33;1) ► **= 0,3779 ou 37,79%**

MODELOS PROBABILÍSTICOS PADRÕES

EXERCÍCIOS DISTRIBUIÇÃO NORMAL

Considere que o peso de um rolo de arame seja normalmente distribuído com média 100 e desvio-padrão 10, ou seja, N(100,10). Então o peso (massa) está em torno de 100 variando entre ±10.

a) Calcular qual a probabilidade que um rolo, retirado ao acaso da produção, possuir peso menor ou igual a 110.

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

Modelos Probabilísticos Padrões

Exercícios Distribuição Normal

Considere que o peso de um rolo de arame seja normalmente distribuído com média 100 e desvio-padrão 10, ou seja, N(100,10). Então o peso (massa) está em torno de 100 variando entre ±10.

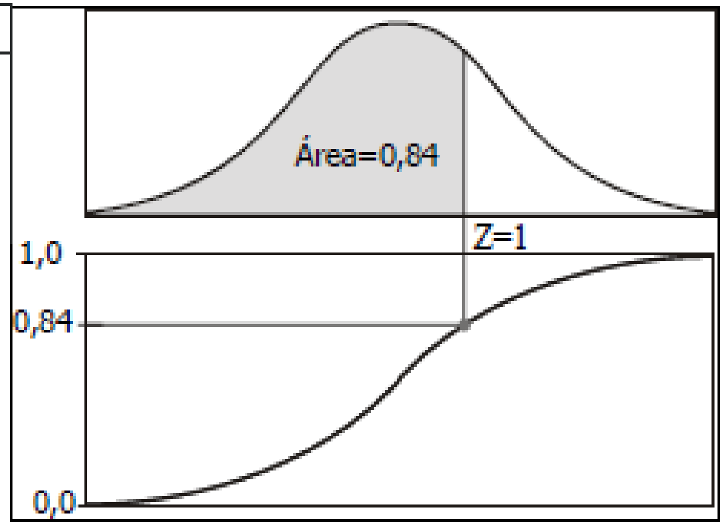
a) Calcular qual a probabilidade que um rolo, retirado ao acaso da produção, possuir peso menor ou igual a 110.

$$Z = \frac{X-\mu}{\sigma} = \frac{110-100}{10} = 1$$

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.0	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	0.9278	0.9292	0.9306	0.9319

Probabilidade de ocorrência de valores abaixo de Z

$$P(x \leq 110) = P(Z < 110) = 0{,}8413$$



Calculando a Prob. Com uso do EXCEL

Processo 1: DIST.NORM.N(110;100;10;1) = 84,13%

Processo 2: a) PADRONIZAR(110;100;10) = Z = 1

b) DIST.NORMP.N(1;1) = 84,13%

Modelos Probabilísticos Padrões

Exercícios Distribuição Normal

Considere que o peso de um rolo de arame seja normalmente distribuído com média 100 e desvio-padrão 10, ou seja, N(100,10). Então o peso (massa) está em torno de 100 variando entre ±10.

b) Calcular a probabilidade do peso do rolo ser maior que 111,6?

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

MODELOS PROBABILÍSTICOS PADRÕES

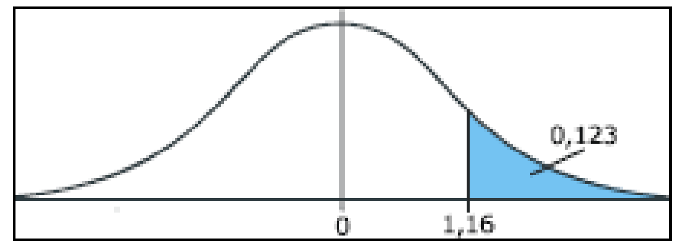
EXERCÍCIOS DISTRIBUIÇÃO NORMAL

Considere que o peso de um rolo de arame seja normalmente distribuído com média 100 e desvio-padrão 10, ou seja, N(100,10). Então o peso (massa) está em torno de 100 variando entre ±10.

b) Calcular a probabilidade do peso do rolo ser maior que 111,6?

$$z = \frac{x-\mu}{\sigma} = \frac{111{,}6-100}{10} = 1{,}16$$

	0.00	0.01	0.02	0.03	0 04	0.05	0.06	0 07	0.. 8	0 09
1.0	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	6 880
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	0.9278	0.9292	0. 30	0 9309



$$P(x > 111{,}6) = 1 - P(Z \le 111{,}6) = 1 - 0{,}8770 = 0{,}123$$

Calculando a Prob. Com uso do EXCEL

Processo 1: 1 - DIST.NORM.N(111,6;100;10;1) = 12,30%

Processo 2: a) PADRONIZAR(111,6;100;10) = Z = 1,16
b) 1 - DIST.NORMP.N(1,16;1) = 12,30%

Modelos Probabilísticos Padrões

Exercícios Distribuição Normal

Considere que o peso de um rolo de arame seja normalmente distribuído com média 100 e desvio-padrão 10, ou seja, N(100,10). Então o peso (massa) está em torno de 100 variando entre ±10.

c) Calcular a probabilidade do peso do rolo estar entre 120 e 130?

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

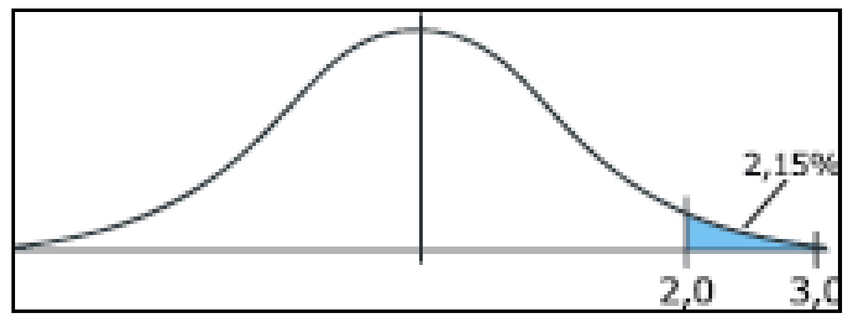
Modelos Probabilísticos Padrões

Exercícios Distribuição Normal

Considere que o peso de um rolo de arame seja normalmente distribuído com média 100 e desvio-padrão 10, ou seja, N(100,10). Então o peso (massa) está em torno de 100 variando entre ±10.

c) Calcular a probabilidade do peso do rolo estar entre 120 e 130?

$$P(120 < x < 130) = P(x < 130) - P(x > 120) = P(Z < 3) - P(Z > 2) = 0{,}9987 - 0{,}9772 = 0{,}0215$$



Calculando a Prob. Com uso do EXCEL

Processo 1: DIST.NORM.N(130;100;10;1)-DIST.NORM.N(120;100;10;1) = 2,15%

Processo 2: a) PADRONIZAR(130;100;10) = Z = 3
b) PADRONIZAR(120;100;10) = Z = 2
c) DIST.NORMP.N(3;1)-DIST.NORMP.N(2;1) = 2,15%

5. DISTRIBUIÇÃO QUI-QUADRADO

É um modelo de distribuição contínua muito importante para a teoria da inferência estatística.

Considere $\mathbf{x_1, x_2, x_3 ... x_p}$, "**n**" variáveis aleatórias **independentes**, **normalmente distribuídas** com **média zero** e **variância 1**, ou seja, **"n" variáveis tipo normal padrão**.

Define-se a variável aleatória com distribuição Qui-Quadrado como: $\chi_n^2 = x_1^2 + x_2^2 + x_3^2 + ... + x_n^2$ ou

$$\chi_n^2 = \sum_{i=1}^{n} z_i^2$$

onde "**n**" é um parâmetro da função densidade de probabilidade denominado **grau de liberdade** e geralmente denotado pela letra grega **φ** (*lê-se fi*), ou eventualmente por **gl**.

CARACTERÍSTICAS DA DISTRIBUIÇÃO QUI-QUADRADO

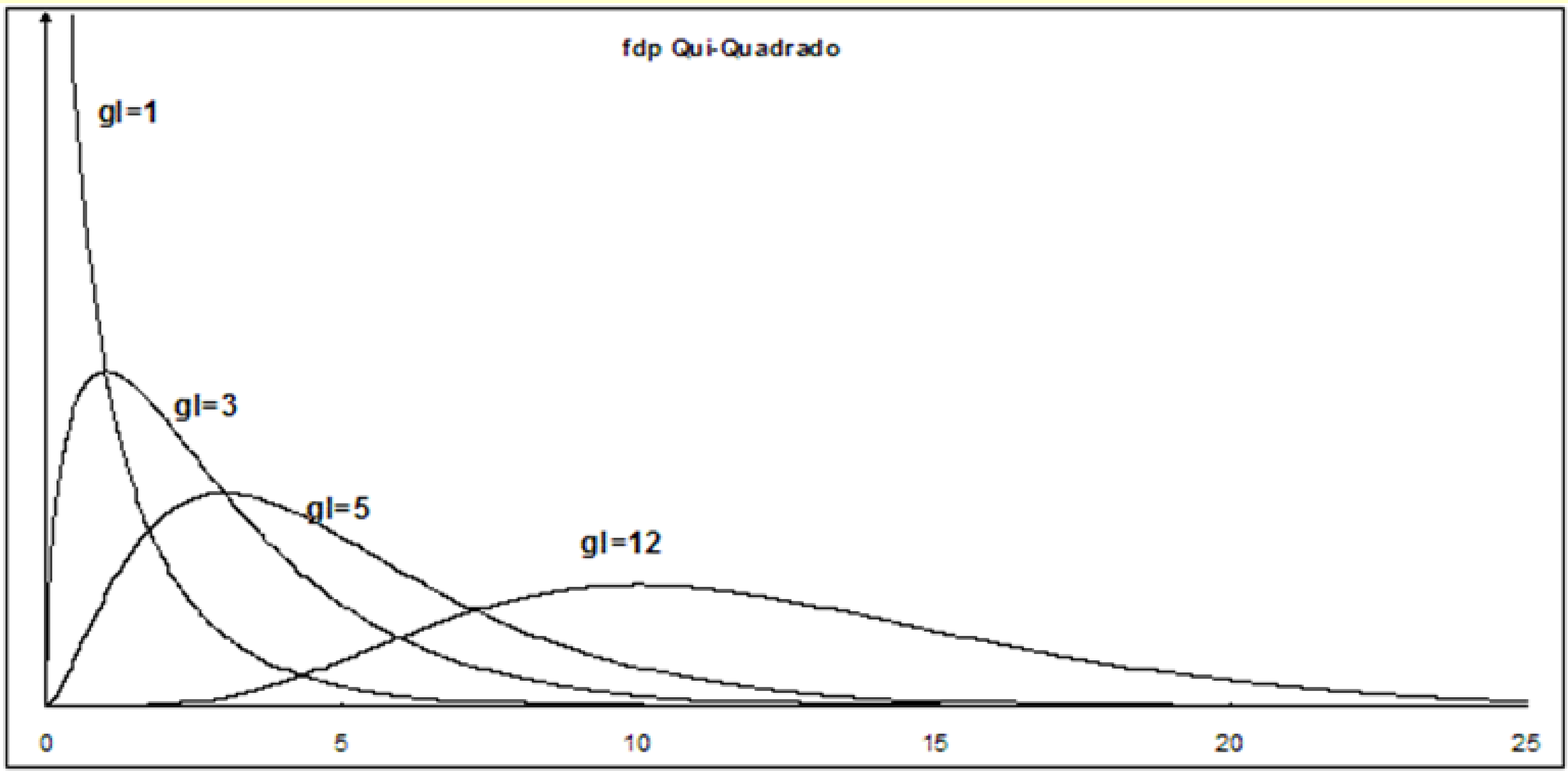
1. $\boldsymbol{\chi_n^2 \geq 0}$ **2. Média = n** **3**. **Variância = 2n**

4. A **função densidade de probabilidade** está representada graficamente para alguns valores de **n**:

MODELOS PROBABILÍSTICOS PADRÕES

5. DISTRIBUIÇÃO QUI-QUADRADO

Observe que à medida que **n cresce**, (*ou* ***gl***), a **função de densidade de probabilidade** tende à forma da **FUNÇÃO NORMAL**.



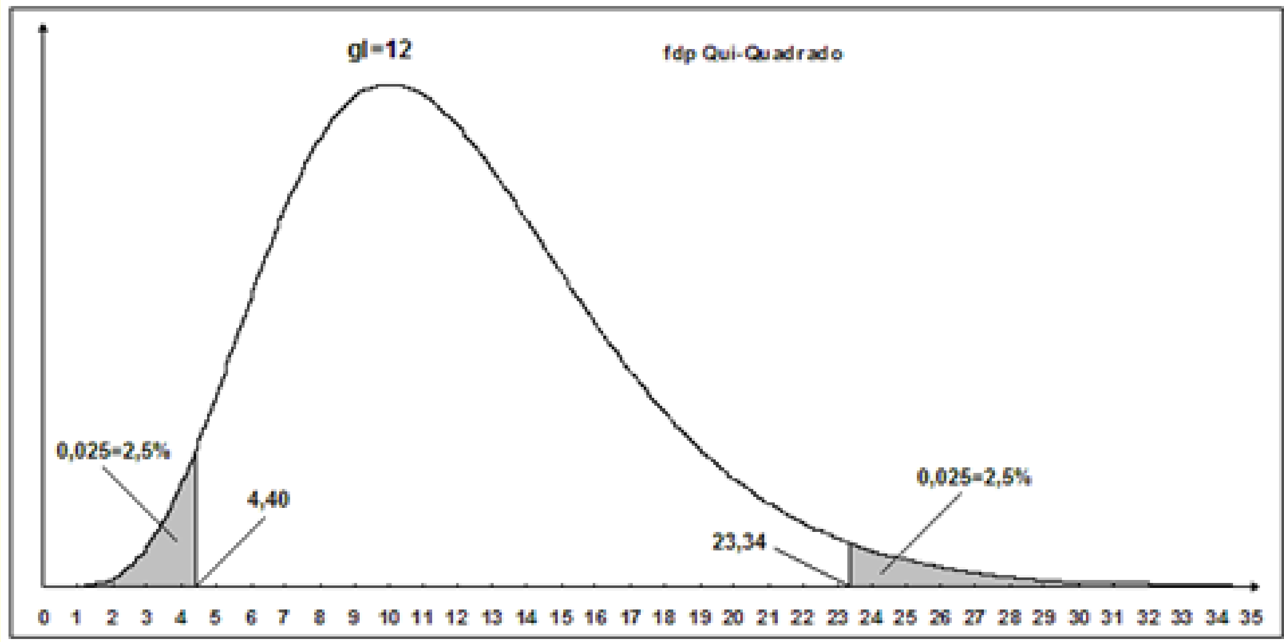
5. DISTRIBUIÇÃO QUI-QUADRADO

TABELA QUI-QUADRADO

A tabela do **Qui-Quadrado** em função do **grau de liberdade n** (ou **gl**), apresenta o **valor numérico da VA** que **deixa à sua direita** determinada área **α**, ou seja **α = P(X ≥ x)**

Para cálculo da probabilidade **P(X ≤ x)**, ou seja, área na cauda esquerda da distribuição, utiliza-se a propriedade **P(X ≤ x) = 1 – P(X ≥ x) = 1 – α**, conforme ilustrado abaixo.

1. O valor à **direita**, chamado, **qui-quadrado superior** é obtido na tabela com **n** =12 e **α** =0,025. Logo, x^2 = **23,34** ou **INV.QUIQUA.CD(0,025;12)**



2. O valor da abscissa à **esquerda**, Chamado **qui-quadrado inferior**, é obtido da tabela com **n** =12 e **α** =1 - 0,025, portanto **α** =0,975. Logo, x^2 = **4,40** ou **INV.QUIQUA.CD(0,975;12)**

5. DISTRIBUIÇÃO QUI-QUADRADO

EXEMPLOS

Qual a probabilidade P(X > 18), sendo n = 25 (*sendo n = gl*):
Resp: DIST.QUIQUA.CD(18;25) ► = **84,24%.**

Qual o valor da variável que gerou a probabilidade 0,8424 ?
Resp: INV.QUIQUA.CD(0,8424;25) ► = 1**8.**

6. DISTRIBUIÇÃO T DE STUDENT

É um modelo de distribuição contínua que **se assemelha à distribuição normal padrão** N(0 ; 1). É utilizada para inferências estatísticas, quando se tem **amostras com tamanhos inferiores a 30 elementos**.

A distribuição **t de Student**, com ***n*** **graus de liberdade** é dada por:

$$t = \frac{z}{\sqrt{\frac{\chi^2_n}{n}}}$$

Para valores de **n < 30**, a distribuição apresenta **maior dispersão** que **z N(0;1)**. À medida que **n aumenta**, ***t*** se aproxima cada vez mais de **z**.

MODELOS PROBABILÍSTICOS PADRÕES

6. DISTRIBUIÇÃO T DE STUDENT

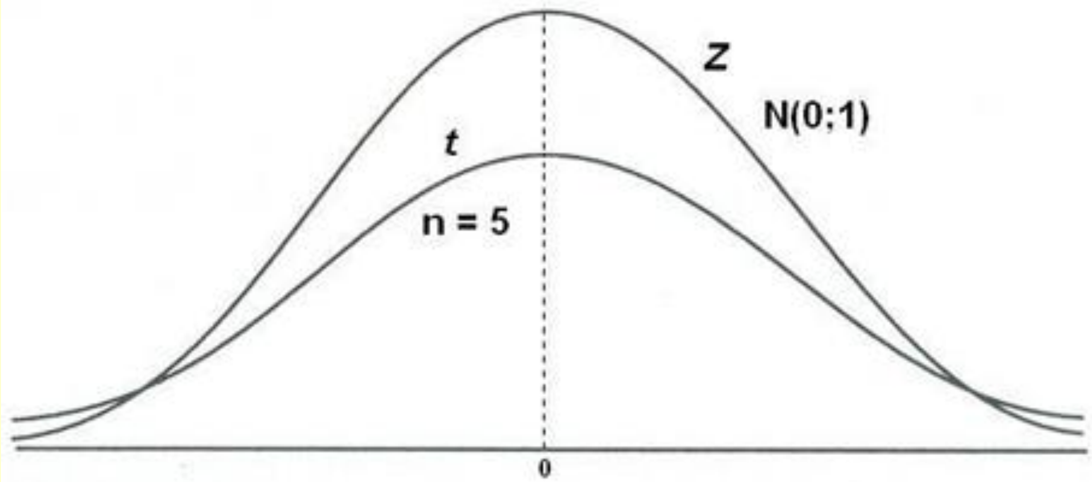
CARACTERÍSTICAS DA DISTRIBUIÇÃO *t* DE STUDENT

1. Média = 0

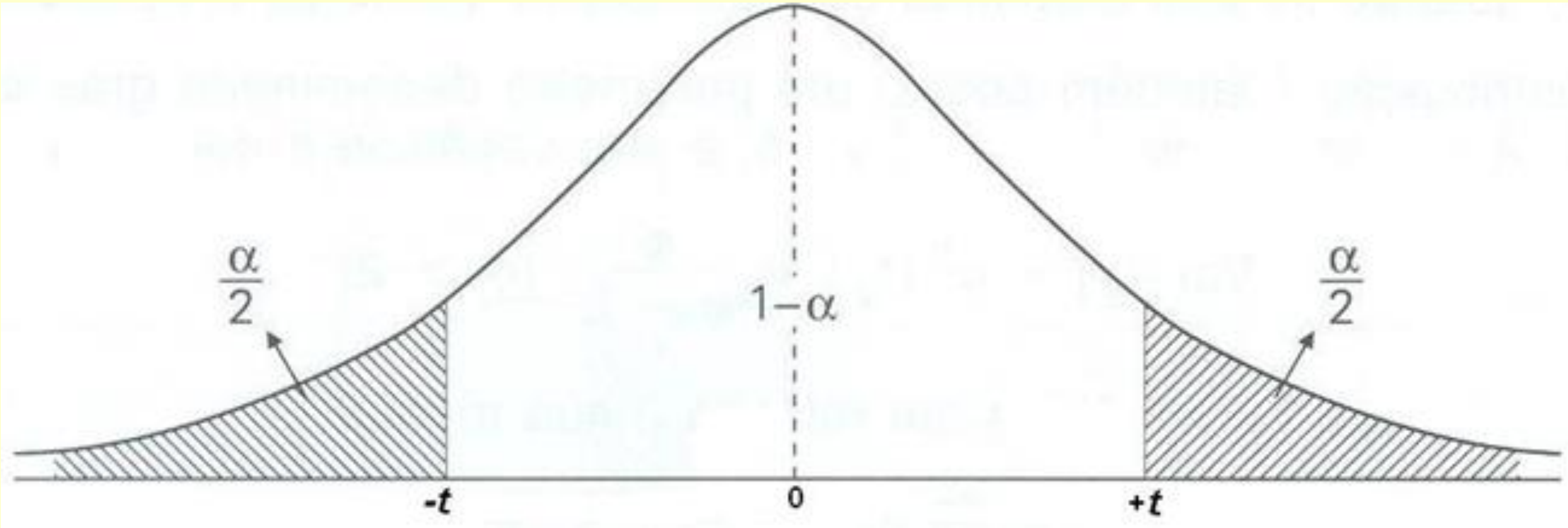
2. Variância =

3. A distribuição é **simétrica em relação à média**.

4. A comparação entre ***t*** e **z** é mostrada no gráfico ao lado.



Trata-se de uma distribuição bicaudal, conforme ilustrado abaixo:



6. DISTRIBUIÇÃO T DE STUDENT

EXEMPLOS

Qual o valor de "t" para: n = 10 ; gl = n – 1 = 9 e α = 5%. (*sendo gl = n -1*):
Resp: INV.T(0,05;10-1) ► **= -1,833** ; para distribuição **UNICAUDAL .**

Resp: INV.T.BC(0,05;10-1) ► **= 2,262** ; para distribuição **BICAUDAL.**

Se n = 10, qual o valor da probabilidade "p" para um t de 1,833 ou P(X = 1,833) ?
Resp: 1-DIST.T(1,833;9;1) ► = **0,05001** ou **5,0%** ; para distribuição **UNICAUDAL.**

Se n = 10, qual o valor da probabilidade "p" para um t de 2,262 ou P(X = 2,262) ?
Resp: 1-DIST.T.BC(2,262;9) ► = **0,05001** ou **5,0%** ; para distribuição **BICAUDAL.**

7. DISTRIBUIÇÃO F DE SNEDECOR

Considere duas distribuições , a primeira com ***n*** graus de liberdade e a segunda com ***d*** graus de liberdade.

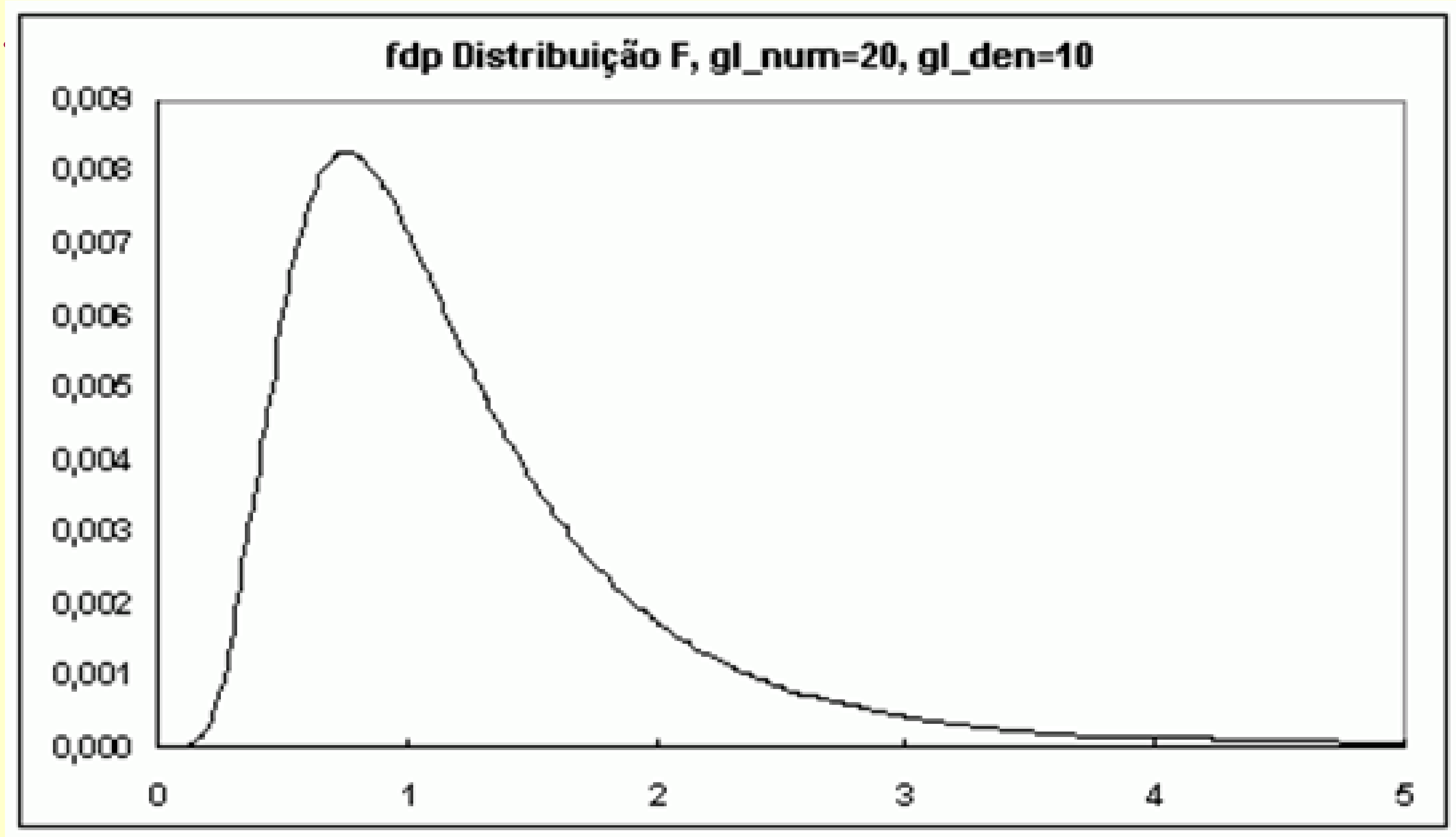
A distribuição **F de Snedecor**, também referida como **F de Fisher** (Snedecor foi o pseudônimo adotado por Fisher), é definida como:

$$F(n,d) = \frac{\dfrac{\chi_n^2}{n}}{\dfrac{\chi_d^2}{d}}$$

MODELOS PROBABILÍSTICOS PADRÕES

7. DISTRIBUIÇÃO F DE SNEDECOR

No caso de **n=20** e **d=10**, a Distribuição F possui a forma de sua **fdp** como mostrado abaixo:



Para fins de utilização da distribuição, na prática tabela-se, para **cada par (n,d)**, o valor de F que deixa ***determinada área à sua direita***.

7. DISTRIBUIÇÃO F DE SNEDECOR

EXEMPLOS

Qual o valor de **P(F>x)** para o gl_n=20, gl_d=10, sendo x=2,77:

Lembrando que: gl = n – 1, ou seja, (gl é igual ao tamanho da amostra menos uma unidade)

Resp: DIST.F(2,774;20;10;FALSO) ► **=0,0504** ou **5,04%.**

Qual o valor crítico "X", para o gl_n=20, gl_d=10, sendo a probabilidade p=0,05 ?

Resp: INV.F(1-0,05;20;10) ► = **2,774**

Fim