ESTATÍSTICA AULA 4

Modelos probabilísticos mais comuns – Unidade 2

Para variáveis aleatórias contínuas – 2ª parte

Professor Marcelo Menezes Reis

Aulas prévias

- Variável aleatória:
  - Conceito, variáveis discretas e contínuas, valor esperado e variância.
- Modelos probabilísticos para variáveis aleatórias discretas: Binomial, Poisson.
- Modelos probabilísticos para variáveis aleatórias contínuas – 1ª parte:
  - Uniforme, normal.

Conteúdo desta aula

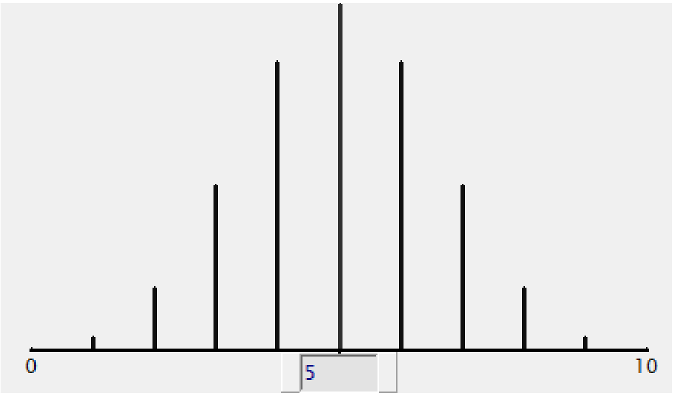
- Principais modelos probabilísticos para variáveis aleatórias contínuas:
  - Modelo normal como aproximação do binomial
  - t de Student;
  - Quiquadrado
- Como usá-los no cálculo de probabilidades.

Modelos para variáveis contínuas

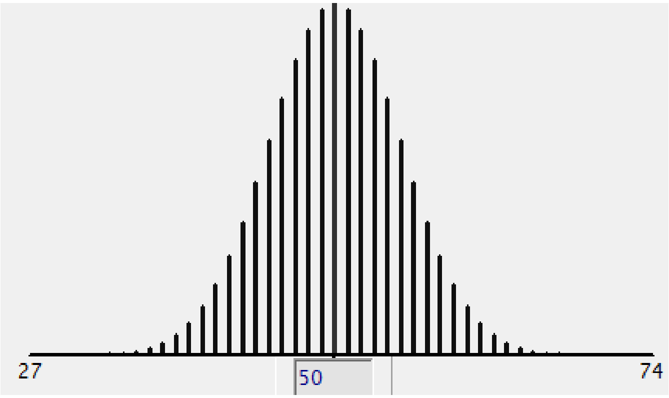
- Variáveis aleatórias contínuas: contradomínio infinito.
- Modelo uniforme.
- Modelo normal.
- Modelo t
- Modelo qui-quadrado.

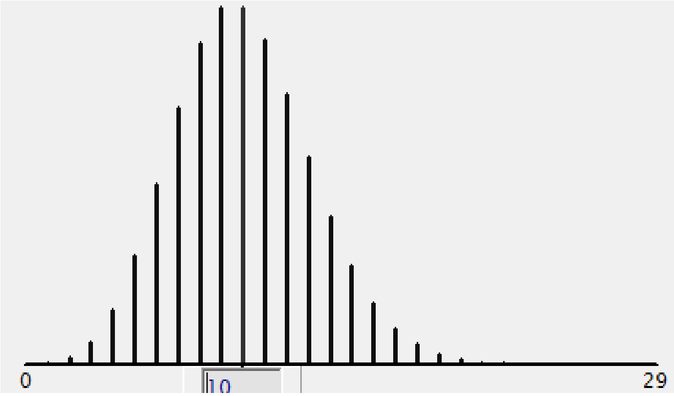
Aproximações pela normal

Binomial: n = 10 p = 0,5

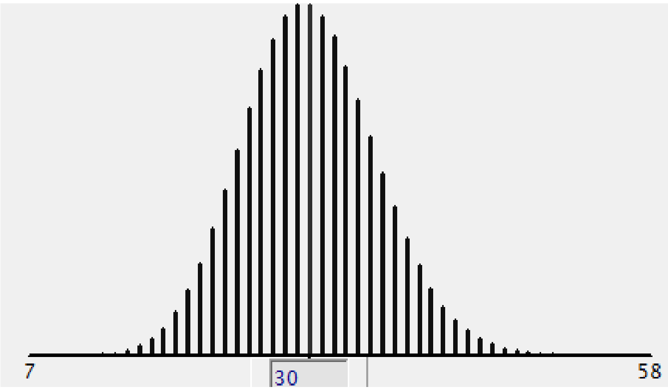


Binomial: n = 100 p = 0,5





Poisson: m = 10



Poisson: m = 30

Aproximação - princípios

- Binomial (modelo discreto) com parâmetros **n** e **p** pode ser aproximada por uma normal (modelo contínuo) com parâmetros μ e σ.

Aproximação - princípios

- Aproximação pela Normal - se p constante quando n aumenta e $\mathbf{n} \times \mathbf{p} \geq 5$ ***E*** $\mathbf{n} \times (1-\mathbf{p}) \geq 5$, deve-se usar uma Normal com:

$$\mu = n \times p \qquad \sigma = \sqrt{n \times p \times (1-p)}$$

Correção de continuidade

- Usando a distribuição normal (**contínua**) para aproximar uma binomial (**discreta**) é necessária uma **correção de continuidade**: associar um intervalo (na distribuição contínua) ao valor discreto.
- Usualmente a variável discreta tem um incremento igual a 1 entre seus valores (0, 1, etc.).

Correção de continuidade

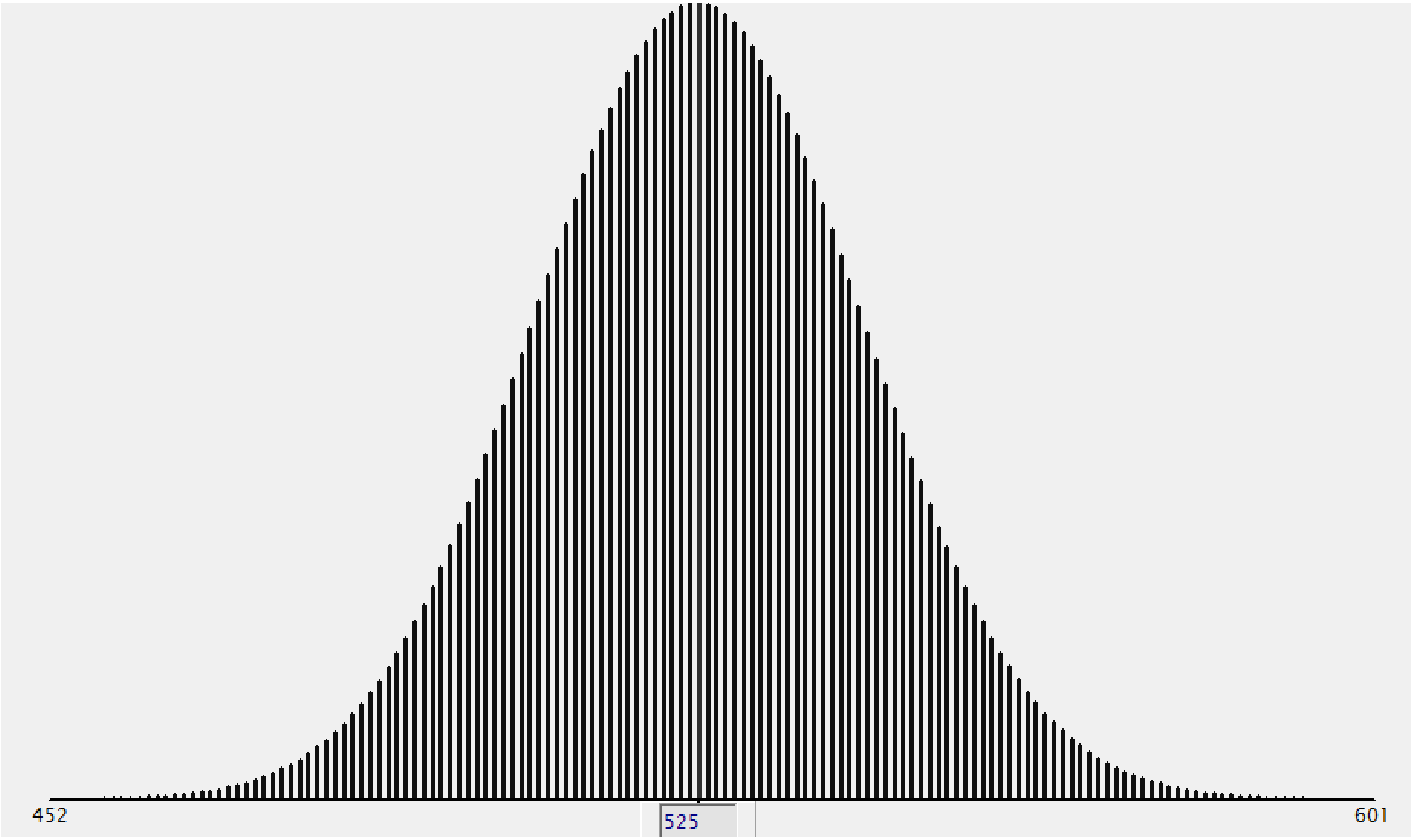
Binomial	Normal
$X = k$	$k - 0{,}5 \le X \le k + 0{,}5$
$X \le k$	$X \le k + 0{,}5$
$X \ge k$	$X \ge k - 0{,}5$
$X < k$	$X \le k - 0{,}5$
$X > k$	$X \ge k + 0{,}5$

Exemplo 1

- Exemplo 7 da Unidade 2. Um município tem 40000 eleitores. Para uma pesquisa de opinião eleitoral uma amostra aleatória de 1500 pessoas foi selecionada. Qual é a probabilidade de que pelo menos 500 dos eleitores sejam menores de 25 anos se 35% dos 40000 são menores do que 25 anos?

Exemplo 1

- Binomial: $n = 1500$, $p = 0{,}35$ (35%)
- $P(X \geq 500) = P(X = 500) + P(X = 501) + \ldots + P(X = 1500)$
- Aproximação pela normal:
  - $\mathbf{n} \times \mathbf{p} = 1500 \times 0{,}35 = 525 > 5$ ***E***
  - $\mathbf{n} \times (\mathbf{1} - \mathbf{p}) = 1500 \times 0{,}65 = 975 > 5$
  - Aproximação possível!


452
525
601

Exemplo 1

$$\mu = n \times p = 1500 \times 0,35 = 525$$

$$\sigma = \sqrt{n \times p \times (1 - p)} = \sqrt{1500 \times 0,35 \times 0,65} = 18,47$$

Binomial: $P(X \geq 500)$ => Normal: $P(X \geq 499,5)$

$P(X \geq 499,5) = P(Z > z_1)$ $z_1 = (499,5 - 525)/18,47 = -1,38$

$P(X \geq 500) \cong P(X \geq 499,5) = P(Z > -1,38) = 1 – P(Z > 1,38)$

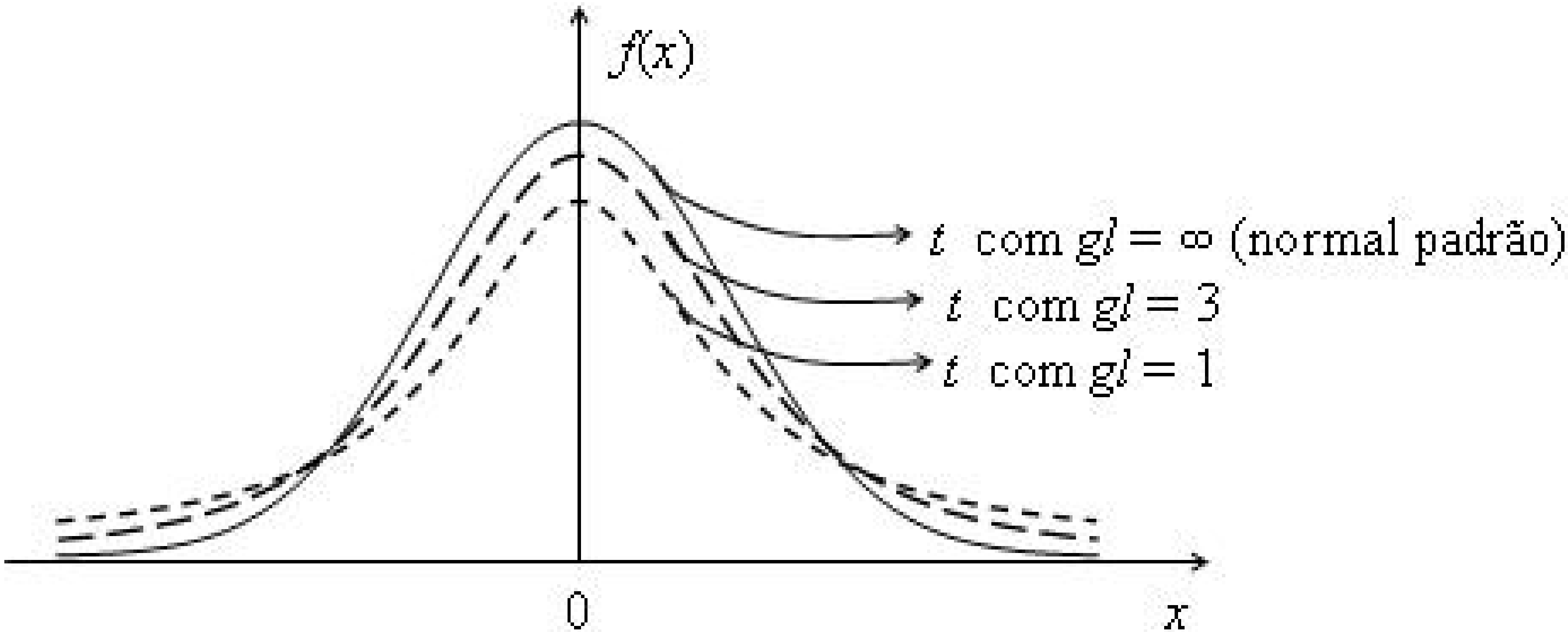
$P(X \geq 500) \cong P(X \geq 499,5) = 1 – 0,0838 = 0,9162$

Modelo t de Student

- Derivação do modelo normal, desenvolvido por William Gosset (Student).
- Mais apropriado para o processo de estimação de parâmetros usando pequenas amostras (Unidade 5).

Modelo t de Student

- O modelo tem média igual a zero (como a normal padrão).
- Mas sua variância é: n/(n-2).
  - Onde n é o tamanho da amostra que foi retirada para o processo de Estimação.
- O modelo t tem n – 1 graus de liberdade.



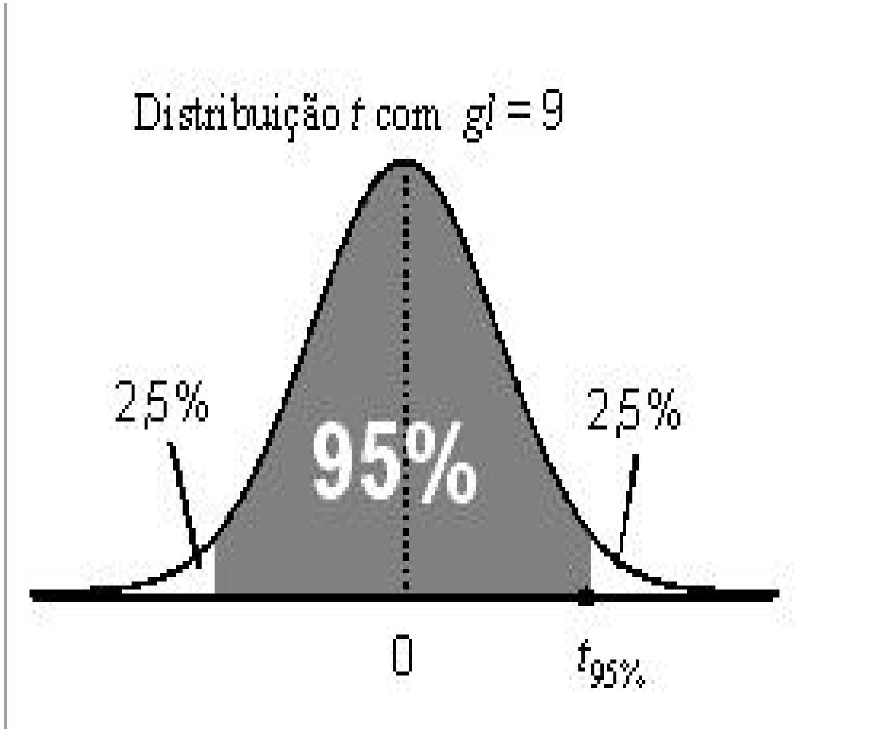
Para cada valor de n temos uma curva t.

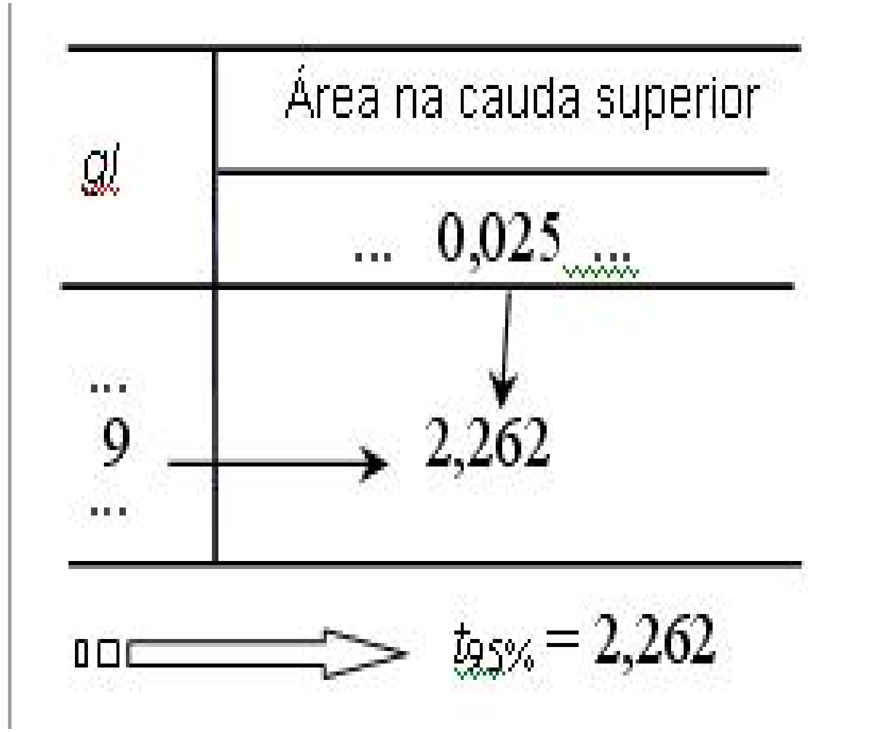
Exemplo 2

- Veja o Exemplo 10 da Unidade 2.
- Imagine a situação do Exemplo 1, obter os valores de t simétricos em relação à média que contêm 95% dos dados, supondo uma amostra de 10 elementos.

Exemplo 2

- Encontrar os valores de t_1 e t_2, simétricos em relação à média, que contém 95% dos dados.
- $t_1 = - t_2$ $P(t > t_2) = 0,025$
- n = 10, então t terá 9 graus de liberdade.





$t_2 = 2{,}262 \qquad t_1 = -\,t_2 = -2{,}262$

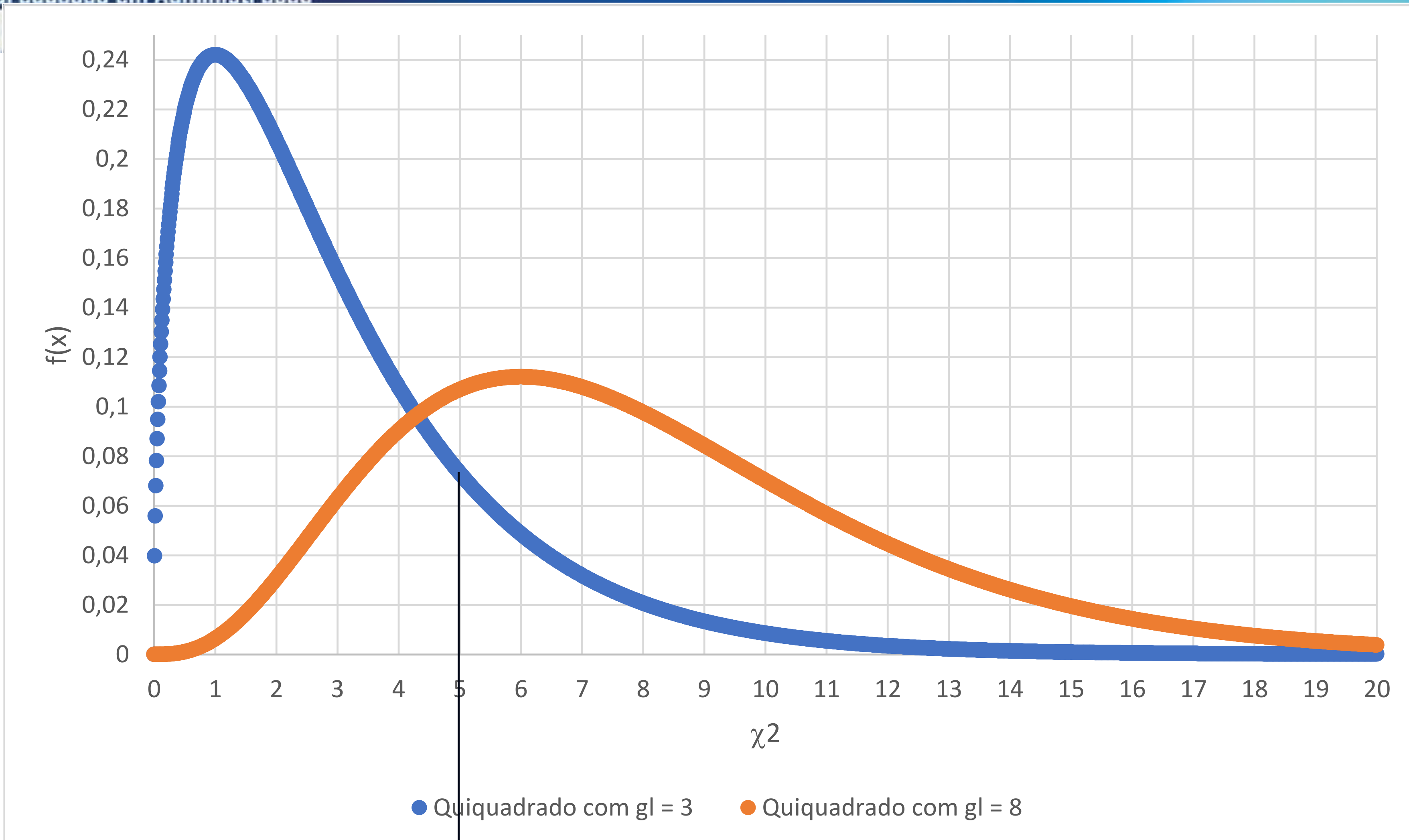
Exemplo 2

- No Excel:
  - =INVT(probabilidade; graus de liberdade)
  - =INVT(0,05;9) = 2,262

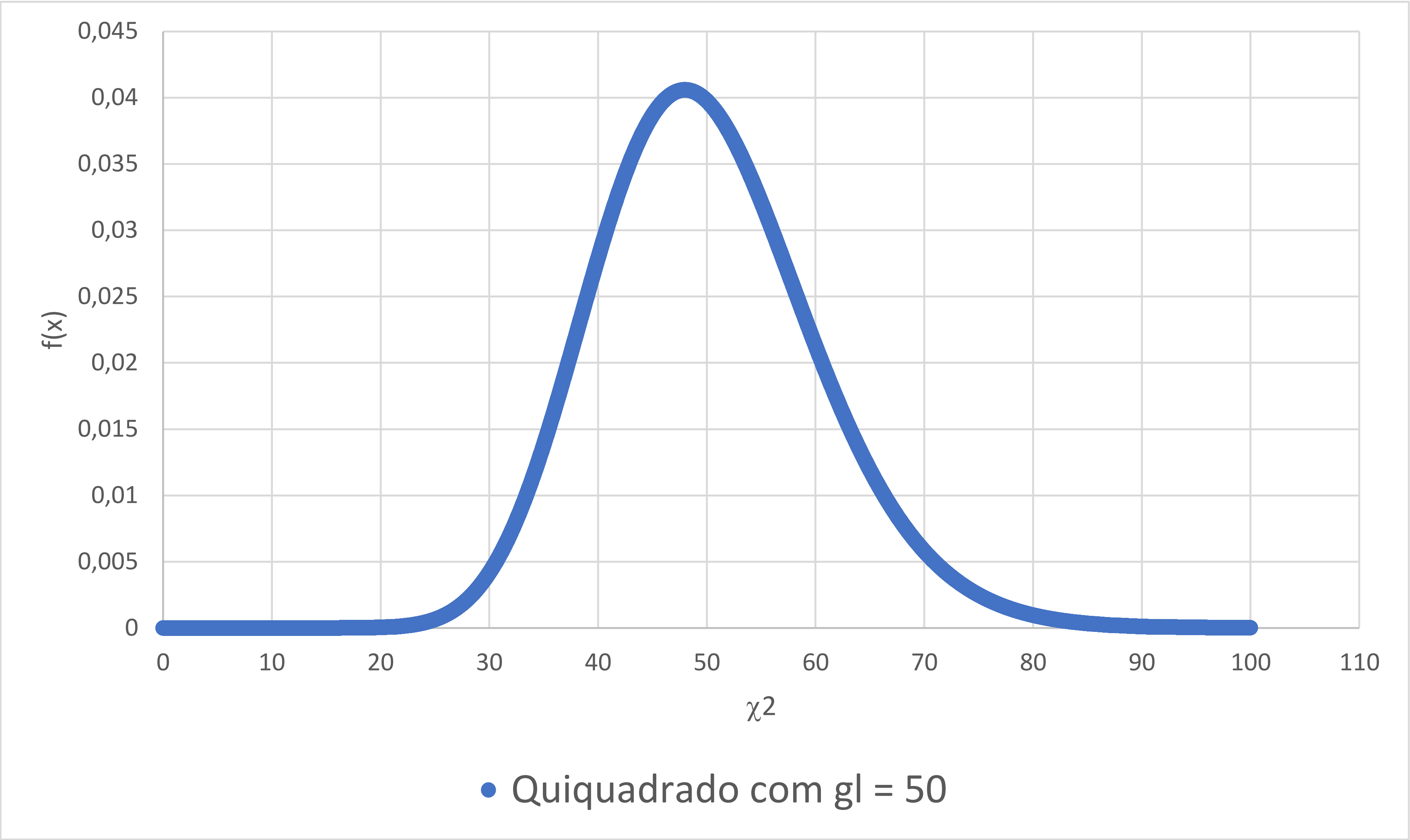
0,025 + 0,025   10 – 1

Distribuição Quiquadrado $\chi 2$

- Derivada da distribuição normal, sempre +.
- Definida pelo número de graus de liberdade
  - Média = graus de liberdade
  - Variância = 2 × graus de liberdade
  - Uso em inferência estatística relacionada à variância, e associação entre variáveis qualitativas (Unidade 6).



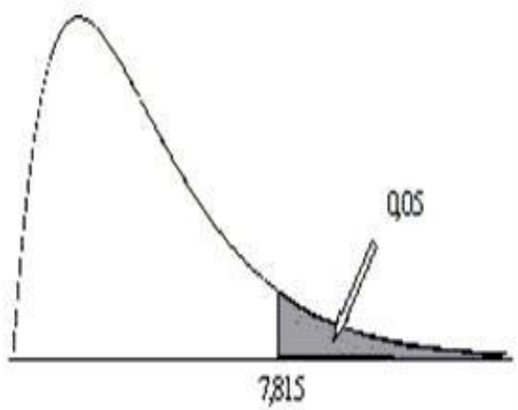
$$P(\chi^2{}_3 < 5) = DIST.QUIQUA(5;3;VERDADEIRO) = 0{,}8282$$

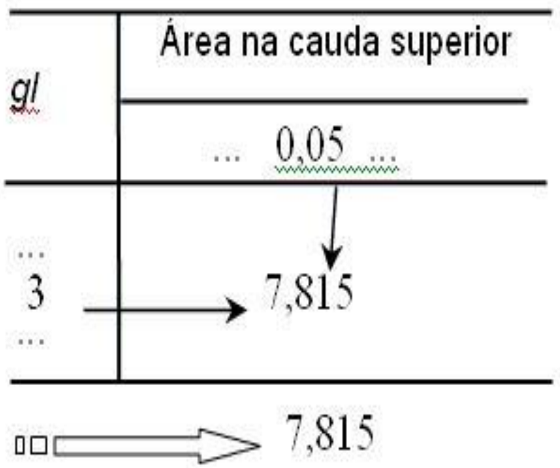

0,045
0,04
0,035
0,03
0,025
0,02
0,015
0,01
0,005
0
f(x)
0 10 20 30 40 50 60 70 80 90 100 110
χ2
Quiquadrado com gl = 50

Exemplo 3

- Exemplo 9 da Unidade 2. Imagine que queremos encontrar o valor da estatística quiquadrado, para 3 graus de liberdade, deixando uma área na cauda superior de 5%.

Exemplo 3





No Excel: = INV.QUI(0,05;3) = 7,815

Tô afim de saber..

- Sobre modelos probabilísticos para variáveis aleatórias contínuas:
  - BARBETTA, P.A., REIS, M.M., BORNIA, A.C. Estatística para Cursos de Engenharia e Informática. 3ª ed. São Paulo: Atlas, 2010, capítulo 6.
  - STEVENSON, Willian J. Estatística Aplicada à Administração. São Paulo: Ed. Harbra, 2001, capítulo 5.

Próxima aula

- Inferência estatística
  - Conceito, tipos, parâmetros e estatísticas.
  - Distribuições amostrais das principais estatísticas: média e proporção.