

# Métodos Estatísticos de Apoio à Decisão

## Aula 7

Mônica Barros, D.Sc.

Agosto de 2008

monica@ele.puc-rio.br

1

# Programa do Curso

Disciplina		Métodos Estatísticos de Apoio à Decisão - BI MASTER 2008	
Responsável		Mônica Barros	
Ferramentas		Excel, @Risk	
Aula	Tipo (T-P-C)	Tema	Descrição
1	T, P	Estatística Descritiva	Gráficos, tabelas e medidas numéricas
2	T	Probabilidade: Definições básicas	Definições básicas; probabilidade, espaço amostral, eventos, propriedades das probabilidades, Probabilidade Condicional, Independência, Teorema de Bayes
3	T	Probabilidade: Definições básicas	Variáveis Aleatórias Contínuas e Discretas, Função de Probabilidade, Função Densidade, Função de Distribuição, Momentos de uma v.a., Média, Variância e Desvio Padrão
4	T, P	Probabilidade: Definições básicas	Variáveis Discretas: Bernoulli, Binomial, Hipergeométrica, Geométrica, Binomial Negativa, Poisson
5	T, P	Probabilidade: v.a. Contínuas	Variáveis Contínuas: Uniforme, Exponencial, Normal
6	P	Prática 1	Aula de exercícios - As funções do Excel para cálculo de probabilidades para v.a. Contínuas e discretas
7	T, C	Probabilidade: v.a. Contínuas E CASE 1: Simulação - soma de v.a. e o teorema central do limite CASE 2: Otimização de um portfólio simulado - propriedades da média e variância e o uso do Solver	O teorema central do limite e a importância da distribuição Normal. O teorema central do limite na prática - soma de variáveis aleatórias e a convergência para a Normal. Distribuição da soma de v.a. e da média amostral. Propriedades da média e variância de combinações lineares de v.a. - o efeito da correlação. O uso do Solver do Excel
8	T, P	Distribuições Amostrais	Amostra aleatória simples, distribuição da média amostral, distribuição de $\chi^2$
9	T, P	Estatística - estimação pontual	Estimação da média da população com sigma conhecido e desconhecido e para proporções
10	T/P	Estatística - estimação por intervalos	Intervalos de confiança para amostras Normais e proporção Binomial - Exercícios - intervalos de confiança empregando o Excel
11	T/P	Estatística - testes de hipóteses	Teste de hipótese para amostras normais e Exercícios

monica@ele.puc-rio.br

2

# Aula 7

- Teorema Central do Limite
- Teorema de DeMoivre e Laplace (aplicação do TCL à Binomial)

monica@ele.puc-rio.br

3

# Teorema Central do Limite

- O teorema central do limite é um dos resultados fundamentais da teoria de probabilidades.
- Por que ele é tão importante?
- Em palavras, o teorema diz que **toda soma de variáveis aleatórias independentes é aproximadamente Normal**, desde que o número de termos na soma seja suficientemente grande.

monica@ele.puc-rio.br

4

## Teorema Central do Limite



- ❑ O Teorema Central do Limite é um resultado importantíssimo porque nos diz que:
  - ❑ Somas de variáveis independentes são aproximadamente Normais
  - ❑ Médias de variáveis independentes são aproximadamente Normais
- ❑ E o mais surpreendente é – **NÃO INTERESSA** qual a densidade das variáveis que estão sendo somadas!

monica@ele.puc-rio.br

5

## Teorema Central do Limite



- ❑ Para ver o TCL em ação, e se convencer de que ele funciona, veja os próximos gráficos.
- ❑ Simulamos diversas amostras de tamanho 10000 da Uniforme(0,1).
- ❑ Em seguida, criamos novas variáveis que são a soma destas Uniformes.
- ❑ Veja o resultado a seguir.

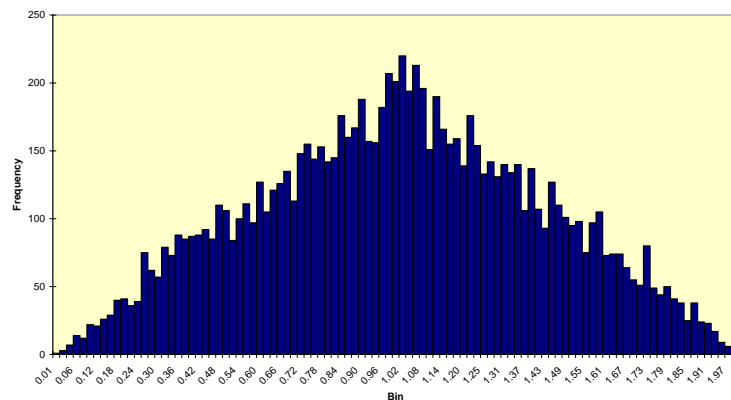
monica@ele.puc-rio.br

6

## Soma de DUAS Unif(0,1)



Histograma - Soma de 2 Unif(0,1) = Distribuição Triangular



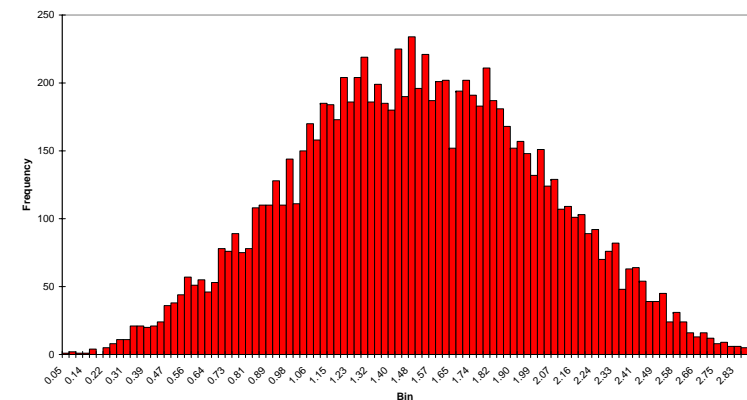
monica@ele.puc-rio.br

7

## Soma de TRÊS Unif(0,1)



Histograma - Soma de 3 Unif(0,1)



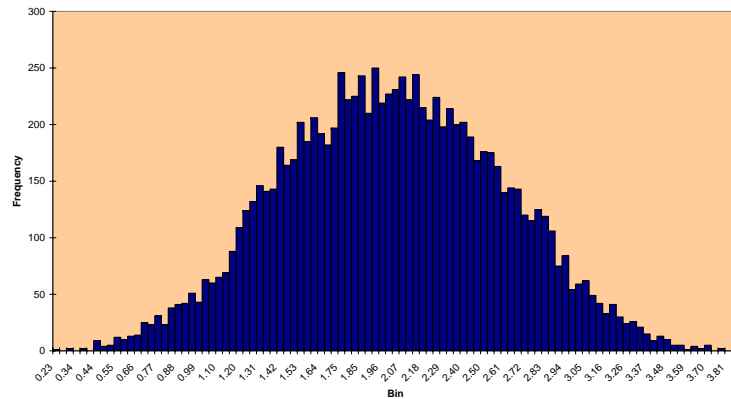
monica@ele.puc-rio.br

8

## Soma de QUATRO Unif(0,1)



Histograma - Soma de 4 Unif(0,1)



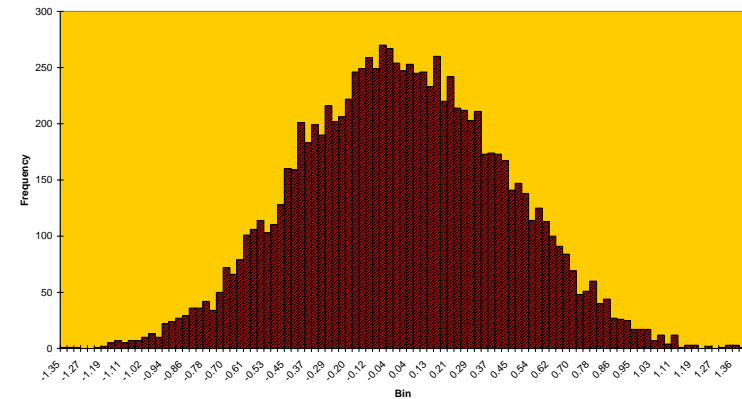
monica@ele.puc-rio.br

9

## Soma de SEIS Unif(0,1)



Histograma - Soma de 6 Unif(0,1) Padronizada - É aproximadamente N(0,1)



monica@ele.puc-rio.br

10

## Teorema Central do Limite



### □ O exemplo mostra que...

- Quanto mais termos eu uso na soma, mais “Normal” parece o histograma....
- E note que eu comecei com uma “coisa” absolutamente “não-Normal”, a Uniforme!
- De uma maneira MUITO informal, é isso mesmo o que o Teorema Central do Limite diz!

monica@ele.puc-rio.br

11

## Teorema Central do Limite



### □ Teorema

- Sejam  $X_1, X_2, \dots, X_n$  variáveis aleatórias independentes tais que  $E(X_i) = \mu_i$  e  $VAR(X_i) = \sigma_i^2$ , ambas finitas. Seja  $Y = X_1 + X_2 + \dots + X_n$ . Então, sob condições bastante gerais podemos afirmar que:

$$Z = \frac{Y - \sum_{i=1}^n \mu_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

- tem aproximadamente uma distribuição  $N(0,1)$ .
- Esta aproximação torna-se cada vez melhor à medida que  $n$  cresce.
- **A importância do teorema decorre do fato da distribuição dos  $X_i$ 's ser qualquer!!**

monica@ele.puc-rio.br

12

## Teorema Central do Limite



- O teorema central do limite pode ser escrito de maneira equivalente usando-se a média de variáveis aleatórias (ao invés da soma). Esta nova versão do teorema é especialmente importante quando os  $X_i$ 's são identicamente distribuídos, além de independentes, isto é, todos têm a mesma distribuição de probabilidade.
- Definição
- Dizemos que  $X_1, X_2, \dots, X_n$  são **identicamente distribuídos** quando todos têm a mesma distribuição de probabilidade. Neste caso, a média e a variância são as mesmas para todos os  $X_i$ , isto é,  $E(X_i) = \mu$  e  $VAR(X_i) = \sigma^2$  para  $i = 1, 2, \dots, n$ .

## Teorema Central do Limite



- Se, além de identicamente distribuídos, os  $X_i$ 's são independentes, dizemos que eles são **i.i.d.** (independentes e identicamente distribuídos).
- **Teorema Central do Limite (versão iid)**
- Sejam  $X_1, X_2, \dots, X_n$  variáveis aleatórias **independentes e identicamente distribuídas** (iid) tais que  $E(X_i) = \mu$  e  $VAR(X_i) = \sigma^2$ , ambas finitas. Seja  $Y = X_1 + X_2 + \dots + X_n$ . Então, sob condições bastante gerais:
$$Z = \frac{Y - n\mu}{\sqrt{n\sigma^2}}$$
- tem aproximadamente uma distribuição  $N(0,1)$ . Esta aproximação torna-se cada vez melhor à medida que  $n$  cresce.

## Teorema Central do Limite



- O teorema pode ser reescrito em termos da média amostral,  $\bar{X}$ , como a seguir.
- **Teorema Central do Limite (versão iid em termos da média amostral)**
- Sejam  $X_1, X_2, \dots, X_n$  variáveis aleatórias independentes e identicamente distribuídas (iid) tais que  $E(X_i) = \mu$  e  $VAR(X_i) = \sigma^2$ , ambas finitas. Seja  $\bar{X}$  a média amostral. Então:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} = \sqrt{n} \cdot \left( \frac{\bar{X} - \mu}{\sigma} \right) \text{ é aproximadamente } N(0,1)$$

- Como nos outros casos, a aproximação melhora com o aumento do tamanho da amostra.

## Teorema Central do Limite



- Pode-se provar que estas duas últimas versões do Teorema Central do Limite são totalmente equivalentes.
- **Uso do teorema central do limite na prática**
  - o número de termos na soma é relativamente grande ( $n$  maior que 20 ou 30)
  - a distribuição dos  $X_i$ 's é "bem comportada", isto é, não é muito assimétrica. Se a distribuição de cada  $X_i$  é muito diferente de uma Normal geralmente é necessário aumentar  $n$  para garantir uma aproximação razoável.

## Teorema Central do Limite



- Note que estamos definindo  $Y$  como a soma dos  $X_i$ 's, e o teorema central do limite é aplicado a uma transformação linear de  $Y$ , e não ao próprio  $Y$ .
- Para que serve esta transformação?
  - O seu objetivo é padronizar  $Y$ , transformando-o numa variável aleatória com média zero e variância 1. Esta nova variável tem a "cara" de uma  $N(0,1)$ .

## Teorema Central do Limite



- Como transformar uma variável aleatória com média  $\mu$  e variância  $\sigma^2$  numa variável com média 0 e variância 1?
- Suponha que  $W$  é uma variável aleatória qualquer com média  $E(W) = \mu$  e variância  $VAR(W) = \sigma^2$ . Então:

$$Z = \frac{W - E(W)}{\sqrt{VAR(W)}}$$

- é uma variável com média zero e variância 1.

## Teorema Central do Limite



- Este tipo de transformação foi usado para gerar uma variável  $Z$  com média zero e variância 1 a partir da soma (ou média) de  $X_1, X_2, \dots, X_n$ .
- Esta nova variável  $Z$  tem, quando  $n$  (o tamanho da amostra) é grande, aproximadamente a densidade  $N(0,1)$ .

## Teorema Central do Limite



- Exemplo
- Uma caixa de madeira contém 250 peças pequenas.
- O peso de cada peça é uma variável aleatória com média 0.20 Kg e desvio padrão 0.08 kg.
- 20 caixas são colocadas num guindaste, que tem a capacidade de agüentar 1010 kg sem problemas. Acima deste peso recomenda-se o uso de outro equipamento.
- Qual a probabilidade da substituição por outro guindaste ser necessária?

## Teorema Central do Limite



- ❑ Solução
- ❑ Esta é uma aplicação típica do teorema central do limite.
- ❑ A distribuição dos pesos das peças **não foi** especificada, sabemos apenas a sua média e variância.
- ❑ O teorema central do limite pode ser usado por que estamos interessados no peso total, isto é, na soma dos pesos individuais, que são supostos independentes e identicamente distribuídos.

## Teorema Central do Limite



- ❑ Seja  $X_i$  o peso em kg da  $i$ -ésima peça, onde  $i = 1, 2, \dots, 5000$ . Então  $Y = \sum X_i$  é o peso total das 20 caixas de peças (excluindo-se o peso das caixas propriamente ditas).
- ❑ Note que  $E(Y) = n.(0.2) = 5000(0.2) = 1000$  kg e  $VAR(Y) = n.(0.08)^2 = 5000.(0.08)^2$  e então o desvio padrão de  $Y$  é  $0.08\sqrt{5000} \approx \sqrt{32} = 5.65$  kg.

## Teorema Central do Limite



- ❑ Seja:

$$Z = \frac{Y - 5000(0.20)}{(0.08)\sqrt{5000}} = \frac{Y - 1000}{(0.08)\sqrt{5000}} = 0.1768(Y - 1000)$$

- ❑ Pelo Teorema Central do Limite,  $Z$  é aproximadamente  $N(0,1)$ . A probabilidade de interesse é:

$$\begin{aligned} \Pr(Y > 1010) &= \Pr(Z > 0.1768(1010 - 1000)) = \\ &= \Pr(Z > 1.768) = 1 - \Phi(1.768) = 1 - 0.9615 = 0.0385 = 3.85\% \end{aligned}$$

## Teorema Central do Limite (exemplo de simulação)



- ❑ O Teorema Central do Limite é usado na prática para gerar v.a. aproximadamente Normais a partir de v.a.  $Unif(0,1)$ , conforme o algoritmo descrito a seguir.
- ❑ Sejam  $X_1, X_2, \dots, X_n$  iid  $Unif(0,1)$ . Então cada  $X_i$  tem média 0.5 e variância 1/12.
- ❑ Seja  $Y$  a soma de 12 v.a.  $Unif(0,1)$ . Então  $Y$  tem uma distribuição (não Uniforme) com média  $12(0.5) = 6$  e variância  $(12)^*(1/12) = 1$ .

## Teorema Central do Limite (exemplo de simulação)



- ❑ Crie uma nova variável Z da seguinte maneira:

$$Z = (Y - 6) = \left( \sum_{i=1}^{12} X_i - 6 \right)$$

- ❑ Pelo CLT, a variável Z tem aproximadamente uma distribuição  $N(0,1)$ .
- ❑ **Será que isso funciona na prática? Veremos o resultado no Excel.**

monica@ele.puc-rio.br

25

## Teorema Central do Limite (exemplo de simulação)

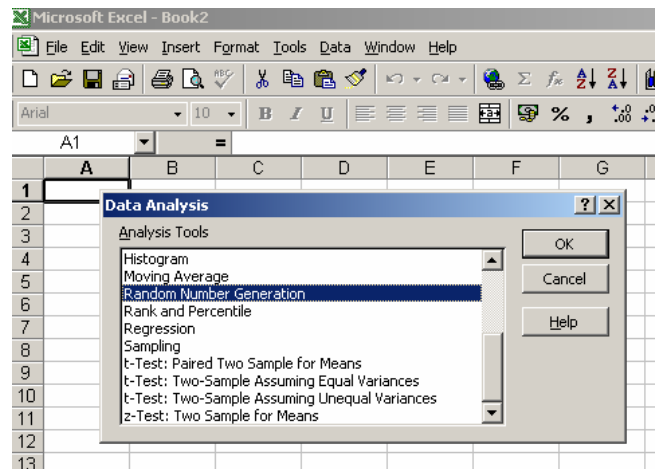


- ❑ A seguir geramos 12 colunas contendo, cada uma, 50 números aleatórios produzidos pelo procedimento de geração de números aleatórios existente no suplemento “Análise de dados”.
- ❑ A 13a. Coluna da planilha conterá a soma das 12 anteriores (isto é, a variável Y), e a 14a. Coluna será a variável Z (Y padronizado).
- ❑ Se tudo der certo, o histograma de Z será uma curva em formato de sino, simétrica em torno de zero.

monica@ele.puc-rio.br

26

## Teorema Central do Limite (exemplo de simulação)



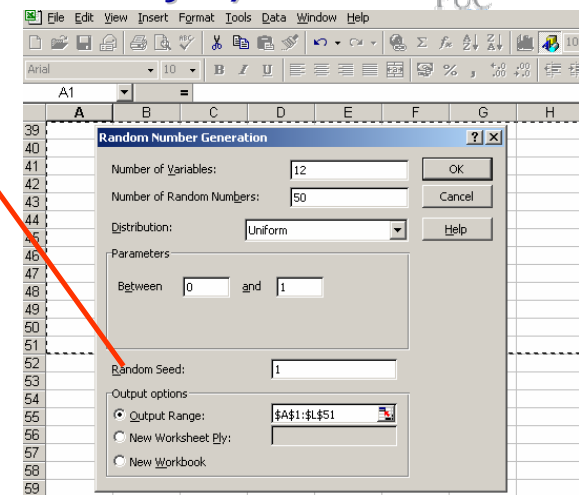
monica@ele.puc-rio.br

27

## Teorema Central do Limite (exemplo de simulação)



Coloquei a semente do gerador de números aleatórios igual a um número determinado (1 neste caso) para poder repetir a geração dos MESMOS números em outra ocasião



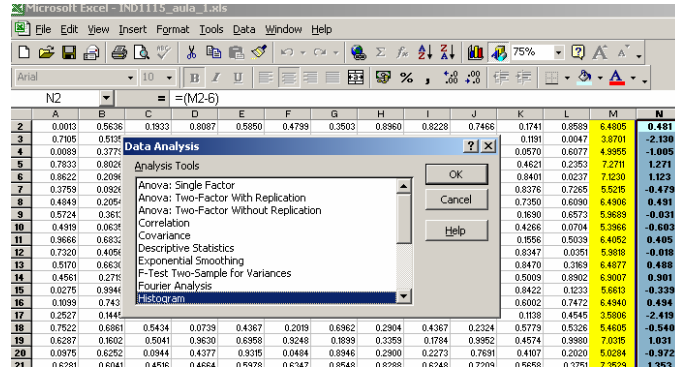
monica@ele.puc-rio.br

28

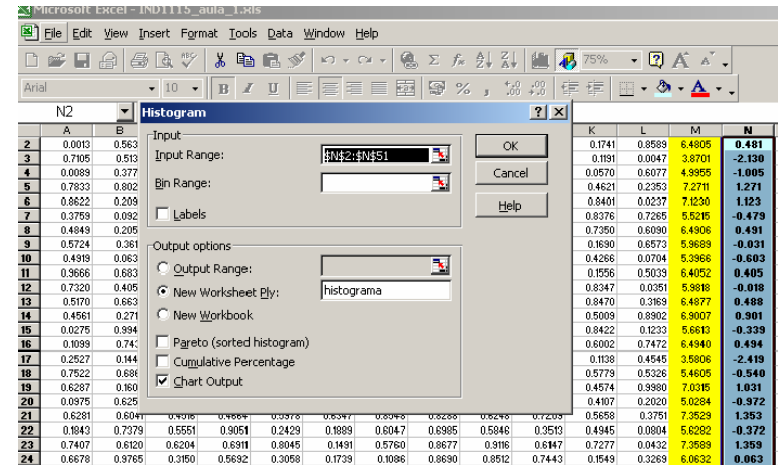
# Teorema Central do Limite (exemplo de simulação)



- Procedimento para obter o histograma (também requer a instalação prévia da ferramenta de análise de dados)



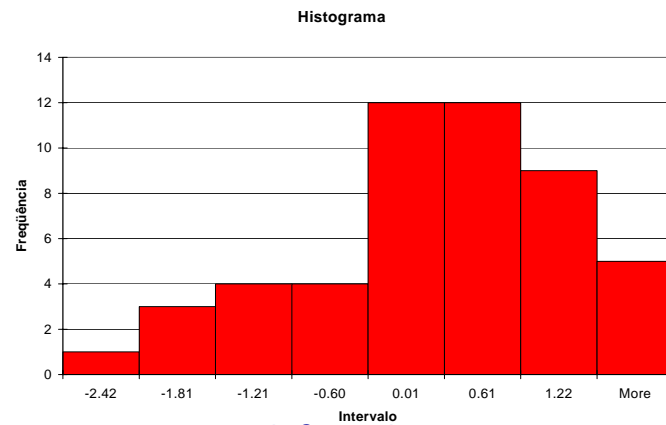
# Teorema Central do Limite (exemplo de simulação)



# Teorema Central do Limite (exemplo de simulação)



- Histograma de Z

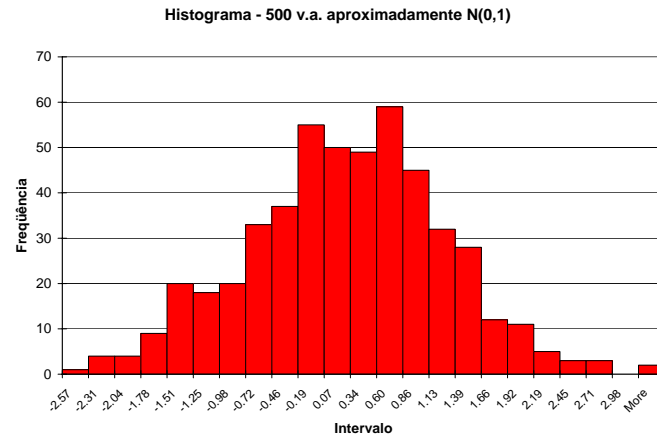


# Teorema Central do Limite (exemplo de simulação)



- O histograma não ficou perfeito, mas já tem uma “cara” de distribuição Normal (embora ainda um pouco assimétrico).
- Tente repetir este exemplo gerando 500 variáveis aproximadamente Normais. Como fica o histograma?
- No próximo gráfico está o resultado de uma simulação que eu fiz usando 500 v.a.

## Teorema Central do Limite (exemplo de simulação)



monica@ele.puc-rio.br

33

## Teorema Central do Limite

- ❑ O resultado do próximo exemplo pode parecer surpreendente à primeira vista (mas não é, se você entendeu o Teorema Central do Limite!).
- ❑ Imagine que você peça a 100 pessoas para jogarem dados 5 vezes, e que elas calculem a média nas 5 jogadas. Qual será a “cara” da média?
- ❑ Como você já sabe, a média é aproximadamente Normal (pelo TCL).

monica@ele.puc-rio.br

34

## Teorema Central do Limite

- ❑ Abaixo estão parte dos resultados de uma simulação...

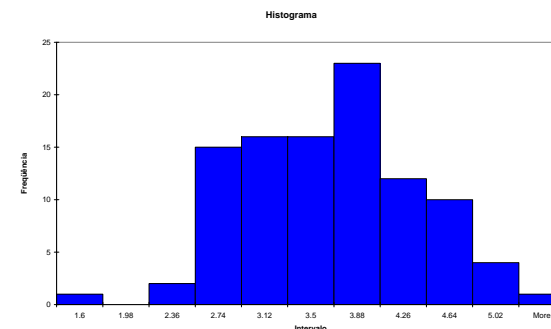
PESSOA	Jogada de 5 dados por 100 pessoas					média
1	4	1	4	2	5	3.2
2	3	6	3	6	1	3.8
3	2	2	5	3	2	2.8
4	5	2	6	6	6	5.0
5	6	1	1	2	4	2.8
6	6	1	5	2	3	3.4
7	2	5	3	2	6	3.6
8	3	1	1	3	5	2.6
9	4	3	4	6	4	4.2
10	4	1	2	5	3	3.0
11	6	6	1	6	3	4.4
12	6	3	6	2	1	3.6
13	3	4	4	4	3	3.6
14	6	6	3	1	6	4.4
15	6	6	4	2	4	4.4
16	1	4	6	6	1	3.6
17	2	6	4	5	1	3.6
18	2	2	3	2	6	3.0
19	5	3	1	4	1	2.8
20	4	5	2	3	4	3.6

monica@ele.puc-rio.br

35

## Teorema Central do Limite

- ❑ Se você fizer o histograma da coluna marcada “Média” você encontrará...



Pode-se até argumentar que o gráfico não parece “muito” Normal, mas as observações originais (jogadas do dado) eram Uniformes nos números 1, 2, 3, 4, 5, 6, o que é bem diferente disso!

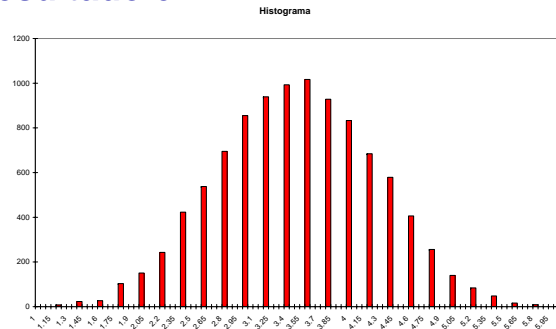
monica@ele.puc-rio.br

36

## Teorema Central do Limite



- Se você agora simula 10 mil pessoas jogando dados e calculando a média, o resultado é...



monica@ele.puc-rio.br

37

## Teorema Central do Limite (para casa)



- Um computador, ao adicionar números, arredonda cada número para o inteiro mais próximo. Suponha que os erros de arredondamento são independentes e identicamente distribuídos com densidade Uniforme no intervalo  $(-0.5, +0.5)$ .
- Suponha que 144 números são somados. Qual a probabilidade de que o módulo da soma dos erros ultrapasse 2?
- Dica:* Use o teorema central do limite. Qual é a média e a variância de cada um dos 144 erros de arredondamento?

monica@ele.puc-rio.br

38

## Teorema Central do Limite (para casa)



- O saldo devedor médio dos usuários de um certo cartão de crédito é uma variável aleatória com média R\$ 112 e desvio padrão R\$ 56.
- Toma-se uma amostra de 50 portadores do cartão. Qual a probabilidade do saldo devedor médio na amostra estar entre R\$ 100 e R\$ 130 ?

monica@ele.puc-rio.br

39

## Teorema Central do Limite (para casa)



- Você está encarregado da compra de champagne para uma festa de reveillon. Sabe-se que, na média, uma pessoa consome 360 ml de champagne, com um desvio padrão de 90ml. Existem 50 convidados na sua festa.
- a) Você comprou 22 garrafas de champagne, cada uma com 750 ml. Qual a probabilidade da champagne não ser suficiente, isto é, do consumo total exceder as 22 garrafas?
- b) Qual a quantidade (em ml) de champagne que você deve comprar para ter certeza que o consumo total só vai exceder esta quantidade 5% do tempo? E 10% do tempo?

monica@ele.puc-rio.br

40

## Teorema de DeMoivre e Laplace



- ❑ O Teorema de DeMoivre e Laplace - aproximação da Binomial pela Normal
- ❑ Este teorema é **apenas um caso particular do teorema central do limite**, pois uma variável aleatória com distribuição Binomial pode ser encarada como a soma de  $n$  variáveis Bernoulli( $p$ ) independentes.
- ❑ Na verdade, o teorema de DeMoivre e Laplace foi formulado muito antes do teorema central do limite.

## Teorema de DeMoivre e Laplace



- ❑ A distribuição **Binomial** com parâmetros  $n$  e  $p$  pode ser aproximada pela distribuição **Poisson** quando  $n$  é grande e  $p$  é pequeno.
- ❑ A aproximação da Binomial pela Normal (Teorema de DeMoivre e Laplace) "funciona" quando  $n$  é grande e  $p$  não é necessariamente pequeno (por exemplo,  $p$  próximo de  $1/2$ ).

## Teorema de DeMoivre e Laplace



- ❑ O fato da distribuição **Binomial** ser **discreta** e da Normal ser contínua cria a **necessidade de corrigir os valores** das probabilidades encontrados via DeMoivre e Laplace.
- ❑ Esta correção será chamada de **correção de continuidade**.

## Teorema de DeMoivre e Laplace



- ❑ A **correção de continuidade** permite uma **melhora significativa** na aproximação dada pelo teorema de DeMoivre e Laplace.
- ❑ Teorema
- ❑ Seja  $Y \sim \text{Bin}(n, p)$  onde  $n$  é "grande" e  $p$  não está próximo de zero. Então:

$$Z = \frac{Y - E(Y)}{\sqrt{\text{VAR}(Y)}} = \frac{Y - np}{\sqrt{np(1-p)}}$$

- ❑ tem aproximadamente uma distribuição  $N(0,1)$ .

## Teorema de DeMoivre e Laplace



- Em particular, podemos calcular probabilidades de valores específicos da variável Y usando a tabela da distribuição Normal padrão.
- Este resultado segue diretamente do teorema central do limite, pois Y pode ser encarado como a soma de n variáveis aleatórias iid, cada uma com distribuição Bernoulli(p).

monica@ele.puc-rio.br

45

## Teorema de DeMoivre e Laplace



- Nota**
- A aproximação Normal fornecida pelo Teorema de DeMoivre e Laplace funciona bem se  $n > 10$  e  $p$  está próximo de  $1/2$ .
- Para outros valores de  $p$  é necessário aumentar o valor de  $n$ .
- Em geral o teorema de DeMoivre e Laplace produz resultados aceitáveis se  $n.p > 5$  quando  $p < 1/2$  e se  $n.(1-p) > 5$  quando  $p > 1/2$ .

monica@ele.puc-rio.br

46

## Teorema de DeMoivre e Laplace



- Exemplo**
- Suponha que desejamos calcular  $\Pr(Y \leq y)$ , onde  $Y \sim \text{Bin}(n, p)$ . Como poderemos fazê-lo usando o teorema de DeMoivre e Laplace?

$$\Pr(Y \leq y) = \Pr\left(\frac{Y - np}{\sqrt{npq}} \leq \frac{y - np}{\sqrt{npq}}\right) = \Pr\left(Z \leq \frac{y - np}{\sqrt{npq}}\right) \approx \Phi\left(\frac{y - np}{\sqrt{npq}}\right)$$

- O resultado desta aproximação pode ser sensivelmente melhorado com a correção de continuidade indicada a seguir.

monica@ele.puc-rio.br

47

## Teorema de DeMoivre e Laplace



- Correção de Continuidade**

Quantidade desejada na distribuição Binomial	Quantidade Calculada através da correção de continuidade	Expressão aproximada usando a densidade Normal
$\Pr(Y = y)$	$\Pr(y - 0.5 \leq Y \leq y + 0.5)$	$\Phi\left(\frac{y+0.5-np}{\sqrt{npq}}\right) - \Phi\left(\frac{y-0.5-np}{\sqrt{npq}}\right)$
$\Pr(Y \leq y)$	$\Pr(Y \leq y + 0.5)$	$\Phi\left(\frac{y+0.5-np}{\sqrt{npq}}\right)$
$\Pr(Y < y) = \Pr(Y \leq y-1)$	$\Pr(Y \leq y - 1 + 0.5)$	$\Phi\left(\frac{y-1+0.5-np}{\sqrt{npq}}\right) = \Phi\left(\frac{y-0.5-np}{\sqrt{npq}}\right)$
$\Pr(Y \geq y)$	$\Pr(Y \geq y - 0.5)$	$1 - \Phi\left(\frac{y-0.5-np}{\sqrt{npq}}\right)$
$\Pr(Y > y) = \Pr(Y \geq y + 1)$	$\Pr(Y \geq y + 1 - 0.5)$	$1 - \Phi\left(\frac{y+1-0.5-np}{\sqrt{npq}}\right)$
$\Pr(a \leq Y \leq b)$	$\Pr(a - 0.5 \leq Y \leq b + 0.5)$	$\Phi\left(\frac{b+0.5-np}{\sqrt{npq}}\right) - \Phi\left(\frac{a-0.5-np}{\sqrt{npq}}\right)$

monica@ele.puc-rio.br

48

## Teorema de DeMoivre e Laplace



- Exemplo
- Uma fábrica produz amortecedores para carros, dos quais 20% são defeituosos.
- Uma amostra de 100 amortecedores é selecionada diariamente. Seja  $X$  o número de amortecedores defeituosos na amostra.
- Calcule a probabilidade de  $X$  ser menor ou igual a 15 usando a aproximação Normal com correção de continuidade.

monica@ele.puc-rio.br

49

## Teorema de DeMoivre e Laplace



- Solução
- Neste caso  $X \sim \text{Bin}(100, 0.2)$  e então  $E(X) = np = 20$ ,  $\text{VAR}(X) = npq = 20(0.8) = 16$ . Usando a correção de continuidade:

$$\begin{aligned}\Pr(X \leq 15) &\approx \Phi\left(\frac{15 + 0.5 - np}{\sqrt{npq}}\right) = \Phi\left(\frac{15.5 - 20}{\sqrt{16}}\right) = \Phi\left(\frac{-4.5}{4}\right) = \\ &= \Phi(-1.125) = 0.1292\end{aligned}$$

- A aproximação sem a correção de continuidade seria:

$$\Pr(X \leq 15) \approx \Phi\left(\frac{15 - 20}{\sqrt{16}}\right) = \Phi\left(\frac{-5}{4}\right) = \Phi(-1.25) = 0.1056$$

monica@ele.puc-rio.br

50

## Teorema de DeMoivre e Laplace



- O valor exato desta probabilidade seria calculado através da fórmula:

$$\Pr(X \leq 15) = \sum_{x=0}^{15} \binom{100}{x} (0.2)^x \cdot (0.8)^{100-x} = 0.1285$$

- O resultado com a correção de continuidade é muito mais próximo do exato do que o cálculo sem a correção.
- A correção de continuidade serve para melhorar a "performance" da aproximação da Binomial pela Normal.

monica@ele.puc-rio.br

51

## Teorema de DeMoivre e Laplace



- Além disso, se desejamos aproximar através de DeMoivre e Laplace a probabilidade da variável Binomial ser igual a um número qualquer, isto só pode ser feito usando a correção de continuidade.
- Por que? A distribuição Normal é contínua, e portanto a probabilidade de qualquer valor é zero, e assim precisamos encontrar a probabilidade de um intervalo em torno do ponto desejado.

monica@ele.puc-rio.br

52

## Teorema de DeMoivre e Laplace



- Exemplo
- Seja  $X \sim \text{Bin}(20, 1/2)$ . Calcule a probabilidade dos valores 5, 6, ...10 diretamente (usando a própria distribuição Binomial) e usando a aproximação Normal com correção de continuidade. Compare os resultados.
- Solução
- O cálculo exato é dado pela fórmula:

$$\Pr(X = x) = \binom{20}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{20-x} = \binom{20}{x} \left(\frac{1}{2}\right)^{20}$$

monica@ele.puc-rio.br

53

## Teorema de DeMoivre e Laplace



- O cálculo aproximado com correção de continuidade é:

$$\begin{aligned} \Pr(X = x) &\approx \Phi\left(\frac{x + 0.5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{x - 0.5 - np}{\sqrt{npq}}\right) = \\ &= \Phi\left(\frac{x + 0.5 - 10}{\sqrt{5}}\right) - \Phi\left(\frac{x - 0.5 - 10}{\sqrt{5}}\right) \end{aligned}$$

- A próxima tabela compara os resultados exato e aproximado:

x	Pr(X = x) exata	Pr(X = x) aproximada
5	0.01479	0.01513
6	0.03696	0.03668
7	0.07393	0.07301
8	0.12013	0.11939
9	0.16018	0.16036
10	0.17620	0.17694

monica@ele.puc-rio.br

54

## Teorema de DeMoivre e Laplace



- Exemplo
- No exemplo anterior calcule  $\Pr(5 \leq X \leq 9)$  de 3 maneiras:
  - Exatamente,
  - Aproximadamente com correção de continuidade,
  - Aproximadamente sem correção de continuidade.

- Solução
- Cálculo exato

$$\Pr(5 \leq X \leq 9) = \Pr(X = 5) + \Pr(X = 6) + \dots + \Pr(X = 9) = 0.40599$$

monica@ele.puc-rio.br

55

## Teorema de DeMoivre e Laplace



- Aproximação com correção de continuidade

$$\begin{aligned} \Pr(5 \leq X \leq 9) &\approx \Pr(4.5 \leq X \leq 9.5) = \Phi\left(\frac{9.5 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{4.5 - np}{\sqrt{npq}}\right) = \\ &= \Phi\left(\frac{9.5 - 10}{\sqrt{5}}\right) - \Phi\left(\frac{4.5 - 10}{\sqrt{5}}\right) = 0.41153 - 0.00695 = 0.40458 \end{aligned}$$

- Aproximação sem correção de continuidade

$$\begin{aligned} \Pr(5 \leq X \leq 9) &\approx \Phi\left(\frac{9 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{5 - np}{\sqrt{npq}}\right) = \\ &= \Phi\left(\frac{9 - 10}{\sqrt{5}}\right) - \Phi\left(\frac{5 - 10}{\sqrt{5}}\right) = 0.32736 - 0.01267 = 0.31469 \end{aligned}$$

- Que é um resultado bem pior que a aproximação com correção de continuidade.

monica@ele.puc-rio.br

56

## Teorema de DeMoivre e Laplace (para casa)



- A probabilidade de uma pessoa com mais de 65 anos pegar uma gripe no outono é 75%. Toma-se uma amostra de 60 pessoas na "3a. idade" e seja  $X$  o número destas pessoas com gripe na amostra. Calcule as seguintes probabilidades:
  - $\Pr\{X \geq 40\}$  aproximadamente pelo teorema de DeMoivre e Laplace *com* correção de continuidade.
  - $\Pr\{X \geq 40\}$  exatamente (só se você tiver acesso a um computador, pois do contrário será bastante trabalhosos!)
  - $\Pr\{45 \leq X \leq 50\}$  aproximadamente pelo teorema de DeMoivre e Laplace *com* correção de continuidade.

## Teorema de DeMoivre e Laplace (para casa)



- Seja  $Y \sim \text{Bin}(12, 1/2)$ .
- Calcule  $\Pr(Y \geq 9)$  exatamente.
- Calcule  $\Pr(Y \geq 9)$  aproximadamente pelo teorema de DeMoivre e Laplace com correção de continuidade.
- Calcule  $\Pr(Y = 7)$  exatamente.
- Calcule  $\Pr(Y = 7)$  aproximadamente pelo teorema de DeMoivre e Laplace com correção de continuidade.