Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística

# Introdução às Técnicas de Simulação em Estatística

*Luiz H.Duczmal, Lupércio F. Bessegato,*
*Marcos A. C. Santos e Sabino J. Ferreira Neto*

**Relatório Técnico**
**RTE-04/2003**
**Série Ensino**

# ÍNDICE

# INTRODUÇÃO

Um forma de descrever fenômenos reais é através de modelos probabilísticos. Existe porém uma condição inevitável, quanto mais realista é o modelo usado mais difícil é a sua análise matemática. Uma maneira alternativa de se analisar um modelo é através de simulação computacional. Com a disponibilidade de computadores cada vez mais poderosos e baratos o uso de técnicas computacionais tem sido cada vez mais extensivo.

O objetivo destas notas é apresentar através de exemplos práticos diversas técnicas de simulação em Estatística. Presume-se que o leitor tenha um conhecimento dos conceitos básicos de Probabilidade como variáveis aleatórias contínuas e discretas e seus modelos principais. No capítulo 1 apresentamos a base de qualquer simulação que envolva fenômenos aleatórios, ou seja, como gerar em um computador um número aleatório (na verdade pseudo-aleatório como discutiremos no capítulo 1). Este número aleatório na verdade é uma simulação dos valores que uma variável aleatória contínua uniforme assume no intervalo (0,1). Usando estes números aleatórios mostramos como simular variáveis aleatórias contínuas e discretas em geral e em particular os principais modelos: binomial, de Poisson, multinomial, normal e exponencial. No capítulo 2 apresentamos diversas aplicações de simulação estocástica na solução de problemas usando o EXCEL com seu pacote de rotinas estatísticas para geração de números aleatórios e de sua configuração como planilha eletrônica de cálculos. Apresentamos os procedimentos para a geração de números aleatórios e de algumas variáveis aleatórias através do EXCEL, para cada uma das variáveis aleatórias apresentamos alguns exemplos comentados de sua utilização e propomos alguns exercícios.

# 1. GERAÇÃO DE VARIÁVEIS ALEATÓRIAS

## 1.1 Geração de Números Aleatórios Uniformemente Distribuídos

Como veremos a seguir nestas notas a base de qualquer geração de variáveis aleatórias em computadores é um gerador de números "aleatórios" que representem os valores que uma variável aleatória uniformemente distribuída no intervalo (0,1) pode assumir. Na verdade os números gerados constituem uma sequência de valores que são determinísticamente gerados de acordo com uma determinada regra iterativa (algoritmo) e neste sentido poderíamos classifica-los como números *pseudo-aleatórios* , porém em termos práticos os valores gerados tem todas as características de valores sucessivos e independentes uma variável aleatória uniforme (0,1).

Vamos apresentar aqui o algoritmo denominado o *método congruencial multiplicativo (MCM)* por sua simplicidade e uso extensivo. Em [1] pode ser encontrada uma discussão e referências mais completas sobre este e outros algoritmos. O *MCM* gera uma sequência de números pseudo-aleatórios partindo de um valor inicial $x_o$, um número inteiro denominado *semente*, e calculando recursivamente os valores sucessivos $x_n,\ n \geq 1$, como:

$$x_n = a * x_{n-1} \quad \text{modulo}\, m \tag{1.1}$$

onde $a$ e $m$ são números inteiros positivos convenientemente escolhidos (ver comentário abaixo) e a operação módulo significa dividir por $m$ o produto $a * x_{n-1}$ e tomar o resto da divisão como o valor de $x_n$. Ou seja, cada $x_n$ pode assumir um dos possíveis valores $\{0,1,...,m-1\}$ e o número real $x_n/m$ é o nosso número pseudo-aleatório, uma aproximação para os possíveis valores assumidos por uma variável aleatória uniforme no intervalo (0,1). Desde que os números gerados $x_n$ assumem um dos valores possíveis em $\{0,1,...,m-1\}$ isto significa que após um número finito passos (no máximo $m$) e todos os valores distintos possíveis terem sido gerados o algoritmo começará a gerar valores repetidos para $x_n$. Os valores para as constantes $a$ e $m$ são escolhidos de tal maneira que, para qualquer valor da semente $x_o$, o número de valores gerados antes da geração do primeiro valor repetido seja o maior possível. Neste sentido o valor de $m$ é escolhido como o maior número primo que possa ser representado por uma *palavra básica* do computador utilizado. Por exemplo, para um computador de 32-*bits* e considerando que o primeiro bit é reservado para o sinal do número representado, pode-se mostrar [2], que os seguintes valores de $a$ e $m$ tem as propriedades desejadas:

$$m = 2^{31} - 1 \quad \text{e} \quad a = 7^5 = 16807$$

Como veremos a seguir, em pelo menos um dos passos de todos os algoritmos apresentados para simulação de variáveis aleatórias necessitaremos de um número aleatório. A notação utilizada nos pseudo-algoritmos será MAKE U = RAND que significa gerar um valor que uma variável aleatória uniformente distribuída no intervalo (0,1) poderia assumir. No texto vamos denotar por RAND um número pseudo-aleatório que aproxima uma variável aleatória uniforme (0,1), ao qual nos referimos simplesmente com número aleatório.

## 1.2 Estimação de Integrais usando Números Aleatórios

Suponha que $f(x)$ seja uma função real e que desejamos calcular a integral $I$ :

$$I = \int_0^1 f(x)\,dx \tag{1.2}$$

Se a variável aleatória $U$ é uniformente distribuída no intervalo (0,1), ou seja $U \sim U(0,1)$, então podemos reexpressar a integral $I$ como:

$$I = E[f(U)] \tag{1.3}$$

onde $E[f(U)]$ é a esperança matemática da variável aleatória $Y = f(U)$. Se $U_1, U_2, \ldots, U_k$ formam uma sequência de variáveis aleatórias independentes, idênticamente distribuídas (i.i.d.), no caso uniformemente distribuídas no intervalo (0,1), as variáveis $Y_1 = f(U_1), \ldots, Y_k = f(U_k)$ também formam uma sequência i.i.d. de variáveis aleatórias com esperança igual à $I$. Pela Lei dos Grandes Números [3] temos que, com probabilidade 1:

$$\hat{I} = \sum_{i=1}^{k} \frac{f(U_i)}{k} \;\rightarrow\; E[f(U)] = I \quad \text{no limite quando } k \rightarrow \infty \tag{1.4}$$

Este método para estimação de integrais, ou seja, geramos um grande número $k$ de números aleatórios $u_i$ e tomamos a média aritmética $\sum_{i=1}^{k} f(u_i)/k$ como uma estimativa $\hat{I}$ do valor exato da integral é denominado método de *Monte Carlo*. Abaixo apresentamos um pseudo-algoritmo para implementar o método:

```
AlGORITMO 1 – MONTECARLO
P1:  INPUT  K
P2:  S = 0
P3:  FOR  I =1 TO  K
P4:           MAKE U = RAND
P5:           S  =  S  +  f(U)
P6:  NEXT
P7:  PRINT   S/K
```

Considerando a Lei do Grandes Números pode-se mostrar que o erro de estimação (a diferença entre o valor exato e o valor estimado) deste método converge para zero com uma amplitude da ordem de $1/\sqrt{k}$ .

Para estimar integrais da forma $I = \int_a^b f(x)\,dx$ devemos adotar uma transformação de variáveis que faça com que possamos re-escrever $I$ como uma integral cujos limites inferior e superior sejam respectivamente 0 e 1 como na expressão (1.2) Neste caso basta fazer $y = (x-a)/(b-a)$ e $g(y) = (b-a)\, f(a + [b-a]y)$ e teremos:

$$I = \int_0^1 g(y)\,dy \tag{1.5}$$

Para integrais do tipo $I = \int_0^\infty f(x)\,dx$ fazemos as transformações:

$$y = \frac{1}{x+1} \qquad \text{e} \qquad g(y) = \frac{f(1/y - 1)}{y^2} \tag{1.6}$$

para obter $I$ como em (1.5).

As idéias acima podem ser extendidas para a estimação de integrais multidimensionais do tipo:

$$I = \int_0^1 \int_0^1 \cdots \int_0^1 f(x_1,\ldots,x_n)\; dx_1 dx_2 \ldots dx_n \tag{1.7}$$

Para estimar integral acima via o método de Monte Carlo vamos re-escrever $I$ como:

$$I = E[f(U_1,\ldots,U_n)] \tag{1.8}$$

onde $U_1,\ldots,U_n$ são $n$ variáveis aleatórias i.i.d. uniformemente distribuídas no intervalo (0,1). Então podemos escrever um algoritmo para estimar $I$ como:

```
AlGORITMO 2 – MONTECARLO MULTIDIMENSIONAL
P1:   INPUT  K, N
P2:   S = 0
P3:   FOR  I = 1 TO  K
P4:            FOR J = 1 TO N
P5:                 MAKE U_J  =  RAND          // gere n números aleatórios
P6:            NEXT
P7:            S =  S + f(U_1, U_2, …,U_N)
```

```
P8:   NEXT
P9:   PRINT   S/K
```

Para estimar integrais de baixa dimensionalidade existem outros métodos numéricos cuja taxa de convergência para zero do erro de estimação os torna mais eficientes que o método de Monte Carlo, porém para 4 ou mais dimensões o método de Monte Carlo é o único disponível e portanto o mais eficiente [1].

# Exercícios

**1.** Considere integrais da forma $I = \int_a^b f(x)\,dx$ onde a função $f(x)$ seja uma função analítica e bem comportada no intervalo $(a,b)$ para a qual sabemos resolver exatamente a integral dada, por exemplo $f(x) = e^x$. Utilizando o Algoritmo 1 apresentado anteriormente, defina a função erro de estimação $e(K) = |\hat{I}(K) - I|$ onde $K$ representa o número de iterações do algoritmo e $\hat{I}(K)$ representa a estimativa obtida de $I$ após $K$ iterações do algoritmo. Para diversos valores de $K$ (considere pelo menos 3 décadas de variação) calcule o erro de estimação $e(K)$. Faça um gráfico de $e(K)$ x $K$ para ter uma idéia da taxa de convergência do método de Monte Carlo.
Obs.: vide comentário após o Algoritmo 1.

**2.** Estime pelo método de Monte Carlo a seguinte integral:

$$I = \int_0^{\infty}\int_0^{x} e^{-(x+y)}\,dy\,dx$$

## 1.3 Geração de Variáveis Aleatórias Discretas e Contínuas

Existem dois métodos básicos para simulação de variáveis aleatórias discretas e contínuas:

— O algoritmo da transformação inversa

— O algoritmo da aceitação/rejeição

### 1.3.1 Algoritmo da Transformação Inversa

Apresentamos a seguir uma proposição que valida o algoritmo da transformação inversa para a geração de variáveis aleatórias contínuas.

**Proposição** Seja $U \sim U(0,1)$ uma variável aleatória contínua uniformemente distribuída no intervalo $(0,1)$. Para qualquer função de distribuição $F(.)$ a variável aleatória $X$ definida por

$$X = F^{-1}(U) \tag{1.9}$$

tem função de distribuição $F$. ($F^{-1}(u)$ é definida como o valor $x$ tal que $F(x) = u$.)

**Prova** Seja $F_X$ a função de distribuição da variável aleatória $X = F^{-1}(U)$. Então

$$\begin{aligned} F_X(x) &= P(X \le x) \\ &= P(F^{-1}(U) \le x) \end{aligned} \tag{1.10}$$

Sendo $F$ uma função de distribuição, decorre que $F(x)$ é uma função monótona crescente de $x$ e que a desigualdade $a \le b$ seja equivalente à $F(a) \le F(b)$. Usando este resultado em (1.10) temos que

$$\begin{aligned} F_X(x) &= P(\,F(F^{-1}(U) \le F(x)\,) \\ &= P(\,U \le F(x)\,) \\ &= F(x) \qquad \text{pois } U \text{ é uniforme em } (0,1) \bullet \end{aligned}$$

Neste caso, o algoritmo para gerar valores de uma variável aleatória $X$ com função de distribuição contínua $F(x)$ será simplesmente:

| AlGORITMO 3 – TRANSFORMAÇÃO INVERSA (CASO CONTÍNUO) |
|---|
| P1: MAKE U = RAND |
| P2: SET X = $F^{-1}$(U) |

Óbviamente este algoritmo sómente é válido quando sabemos como calcular analíticamente a função inversa $x = F^{-1}(u)$ da função de distribuição $u = F(x)$.

**Exemplo 1** Se $X$ é uma variável aleatória exponencial com parâmetro $\lambda = 1$, então sua função de distribuição é:

$$F(x) = 1 - e^{-x}$$

Neste caso podemos obter a inversa $x = F^{-1}(u)$. Seja

$$u = F(x) = 1 - e^{-x}$$

então

$$\begin{aligned} 1 - u &= e^{-x} \\ x &= -\log(1-u) \end{aligned}$$

Ou seja, para gerar um valor de uma v.a. exponencial com parâmetro 1 basta gerar U = RAND e fazer

$$X = F^{-1}(U) = -\log(1-U) \tag{1.11}$$

Para fazer uma pequena economia no custo computacional do algoritmo podemos substituir a operação subtração $1-U$ por simplesmente $U$ (pois assim como $U$ a v.a. $1-U$ também é uniformente distribuída no intervalo (0,1) ). Se o parâmetro $\lambda$ da distribuição exponencial é diferente de 1 é facíl ver que em vez da expressão (1.11) teremos:

$$X = -\lambda^{-1} \log U \tag{1.12}$$

Apresentamos a seguir versão do algoritmo da transformação inversa para a geração de valores de uma variável aleatória discreta $X$ que tem uma função de probabilidade tal que:

$$P(X = x_j) = p_j, \qquad j = 0,1,... \text{ e } \sum_j p_j = 1$$

| AlGORITMO 4 – TRANSFORMAÇÃO INVERSA (CASO DISCRETO) | |
|---|---|
| P1: | MAKE U = RAND |
| P2: | IF U < $p_0$, SET X = $x_0$, STOP |
| P3: | IF U < $p_0 + p_1$, SET X = $x_1$, STOP |
| P4: | IF U < $p_0 + p_1 + p_2$, SET X = $x_2$, STOP |

$$\vdots$$

$$\text{P(J+1):} \qquad \text{IF} \quad \sum_{i=1}^{j-1} p_i \le \text{U} < \sum_{i=1}^{j} p_i, \qquad \text{SET X} = x_j, \ \text{STOP}$$

$$\vdots$$

Se $U \sim U(0,1)$ e $0 < a < b < 1$ então $P(a \le U < b) = b - a$. Usando este resultado pode-se mostrar que o (J+1)-ésimo passo do algoritmo 4 gera o valor $x_j$ da variável aleatória $X$ com a probabilidade correta, ou seja

$$P\left(\sum_{i=1}^{j-1} p_i \le U < \sum_{i=1}^{j} p_i\right) = p_j = P(X = x_j)$$

**Exemplo 2** Queremos simular valores de variável aleatória discreta e uniforme $X$ que assume os valores $j = 1,2,\ldots,n$ com igual probabilidade, ou seja, $P(X = j) = 1/n$, $j = 1,\ldots n$. De acordo com o algoritmo 4 o primeiro passo é gerar um RAND, então

$$X = j \qquad \text{se} \qquad \frac{j-1}{n} \le U \le \frac{j}{n} \tag{1.13}$$

as desigualdades acima podem ser re-escritas como $j - 1 \le nU \le j$, o que nos permite escrever

$$X = \mathrm{Int}(nU) + 1.$$

o valor da função $\mathrm{Int}(x)$ é o maior número inteiro menor ou igual à $x$.

## **Exemplo 3** Vamos gerar valores de uma variável aleatória binomial com parâmetros $n$ e $p$, ou seja, $X \sim B(n,p)$ e

$$P(X = k) = \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \qquad \text{sendo } k = 0,1,\ldots,n$$

Para construir o algoritmo da transformação inversa vamos utilizar a relação recursiva para probabilidades binomiais:

$$P(X = k+1) = \frac{n-k}{k+1}\,\frac{p}{1-p}\,P(X = k) \qquad k \ge 0$$

| AlGORITMO 5 – TRANSFORMAÇÃO INVERSA (VARIÁVEL BINOMIAL) |
| --- |
| P1: INPUT $n$ e $p$ |
| P2: MAKE U = RAND |
| P3: SET $c = p/(1-p)$, $\mathrm{I} = 0$, $\mathrm{Pr} = (1-p)^n$, $\mathrm{F} = \mathrm{Pr}$. |
| P4: IF $\mathrm{U} < \mathrm{F}$, SET X = I, STOP |
| P5: $\mathrm{Pr} = [c(n-\mathrm{I})/(\mathrm{I}+1)]\ \mathrm{Pr}$, $\mathrm{F} = \mathrm{F} + \mathrm{Pr}$, $\mathrm{I} = \mathrm{I} + 1$. |
| P6: GO TO P3 |

Existe outra maneira de simular valores para uma variável binomial $X \sim B(n,p)$. A variável $X$ que representa o números de "sucessos" que ocorrem, cada um com probabilidade $p$, ao realizarmos $n$ repetições independentes de um experimento aleatório pode também ser representada como uma soma de $n$ variáveis independentes de Bernoulli. Cada uma destas variáveis independentemente assume ou o valor 1 ("sucesso") com probabilidade $p$ ou o valor 0 com probabilidade $1-p$. Podemos simular uma variável de Bernoulli $Y$ gerando um RAND U e fazendo $Y=1$ se $U<p$, caso contrário $Y=0$. Para simular a variável binomial $X$ simulamos independentemente $n$ variáveis de Bernoulli $Y_i$, $\mathrm{i}=1,\ldots 10$ e fazemos $X=\sum_{i=1}^{n} Y_i$.

| AlGORITMO 6 – VARIÁVEL BINOMIAL |
| --- |
| P1: INPUT $n$ e $p$ |
| P2: SET X = 0 |
| P3: FOR I = 1 TO $n$ |
| P4: MAKE U = RAND |
| P5: IF U < $p$, SET X = X+1 |
| P6: NEXT |
| P7: PRINT X, STOP |

**Exemplo 4** Um extensão simples do algoritmo 6 pode ser usada para simular uma variável multinomial. Neste caso realizamos $n$ repetições independentes de um experimento aleatório. A cada repetição um dos $k$ resultados possíveis $A_i$, $i=1,\ldots,k$ do experimento ocorre, cada um com sua respectiva probabilidade $p_1, p_2, \ldots, p_k$, tal que $\sum_{i=1}^{k} p_i = 1$. Denotando por $X_i$ o número de vezes que o resultado $A_i$ ocorre dentre as $n$ repetições do experimento aleatório podemos escrever a variável aleatória $k$-dimensional multinomial como $\mathbf{X} = (X_1, X_2, \ldots, X_k) \sim M(n,\mathbf{p})$ onde $\mathbf{p} = (p_1, p_2, \ldots, p_k)$. A função de probabilidade conjunta (distribuição multinomial) é

$$P(X_1 = n_1, X_2 = n_2, \ldots, X_k = n_k) = \frac{n!}{n_1! n_2! \ldots n_k!}\ p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k} \qquad (1.14)$$

$$\text{onde } \sum_{i=1}^{k} n_i = n.$$

O algoritmo a seguir simula uma variável multinomial $\mathbf{X} = (X_1, X_2, \ldots, X_k) \sim M(n, \mathbf{p})$

| AlGORITMO 7 – VARIÁVEL MULTINOMIAL | |
|---|---|
| P1: | INPUT $n$, $k$, $p_1, p_2, \ldots, p_k$ |
| P2: | SET $X_1 = X_2 = \ldots = X_k = 0$ |
| P3: | FOR I = 1 TO $n$ |
| P4: | MAKE U = RAND |
| P5: | SET SOMA = 0, J = 0 |
| P6: | DO |
| P7: | J = J+1 |
| P8: | SOMA = SOMA + $p_j$ |
| P9: | WHILE SOMA < U |
| P10: | $X_J = X_J + 1$ |
| P11: | NEXT |
| P12: | PRINT $X_1, X_2, X_3, \ldots, X_k$ , STOP |

**Exemplo 5** Vamos gerar valores de uma variável aleatória de Poisson $X$ com parâmetro $\lambda$, ou seja, $X$ tem uma distribuição de probabilidade

$$P(X = k) = e^{-\lambda} \frac{\lambda^{k}}{k!} \qquad \text{sendo } k = 0, 1, \ldots, n$$

Para construir o algoritmo da transformação inversa vamos utilizar uma relação recursiva para probabilidades de Poisson:

$$P(X = k+1) = \frac{\lambda}{k+1} P(X = k) \qquad k \geq 0$$

A seguir o algoritmo da transformação inversa para simular valores de uma variável de Poisson

| AlGORITMO 9 – TRANSFORMAÇÃO INVERSA (VARIÁVEL DE POISSON) | |
|---|---|
| P1: | INPUT $\lambda$ |
| P1: | MAKE U = RAND |
| P2: | SET I = 0, $\Pr = e^{-\lambda}$, F = Pr |
| P3: | IF U < F, SET X = I, STOP |

| |
|---|
| P4: $\Pr = [\lambda \ / (\mathrm{I}+1)] \Pr$, $\mathrm{F} = \mathrm{F} + \Pr$, $\mathrm{I} = \mathrm{I}+1$ |
| P5: GO TO P3 |

# Exercícios

1. **1.** Construa um algoritmo para gerar valores de variável aleatória geométrica $X$ com parâmetro $p$.
2. **2.** Construa um algoritmo para estimar o valor de $\bar{f} = \frac{1}{n}\sum_{i=1}^{n} f(i)$ onde $n$ é um inteiro grande e o cálculo de $f(i)$ para $i = 1,\ldots,n$ é trabalhoso.
   Sugestão: considere $X$ uma variável aleatório uniforme discreta sobre os inteiros $1,\ldots,n$. Faça um algoritmo para estimar a esperança $E[f(X)]$.

## 1.3.2 Algoritmo da Aceitação-Rejeição

A idéia do algoritmo da aceitação-rejeição (ou simplesmente rejeição) para gerar valores de uma variável aleatória, seja esta discreta ou contínua, é usar os valores gerados por um segundo algoritmo de uma variável aleatória auxiliar respectivamente discreta ou contínua que assuma o mesmo conjunto de valores da variável original. A mecânica do algoritmo é simples: um valor inicial é gerado para a variável auxiliar, caso satisfaça determinada condição o valor gerado será aceito para a variável original, caso contrário um novo valor para a variável auxiliar será gerado, testado e assim sucessivamente até que o valor gerado satisfaça a condição desejada e seja aceito como um valor da variável original.

Primeiro vamos apresentar a versão discreta do algoritmo da rejeição e em seguida a versão contínua junto com os respectivos exemplos.

ALGORITMO DA REJEIÇÃO (CASO DISCRETO) - Suponha que temos um algoritmo eficiente para simular valores de uma variável aleatória discreta $Y$ que tem distribuição de probabilidade $\{j \geq 0,\ q_j = P(Y = j)\}$. Nosso objetivo é o de simular valores de uma segunda variável aleatória discreta $X$ que assume o mesmo conjunto de valores $\{j \geq 0\}$ de $Y$ porém com uma função de probabilidade $p_j = P(X = j)$. Seja $c$ uma constante tal que

$$\frac{p_j}{q_j} \leq c \qquad \text{para todo } j \text{ tal que } p_j > 0 \tag{1.15}$$

neste caso podemos escrever o algoritmo da rejeição como

| AlGORITMO 10 – MÉTODO DA REJEIÇÃO (CASO DISCRETO) |
| --- |
| P1: GENERATE $Y$ |
| P2: MAKE U = RAND |
| P3: IF U < $p_Y / c\, q_Y$, SET X = Y, STOP |
| P4 RETURN TO P1 |

Obs.: os passos P2 e P3 significam aceitar o valor gerado para $X$ com probabilidade $p_Y / c\, q_Y$.
O Teorema a seguir comprova a validade do método da rejeição.

**Teorema 1** O algoritmo da aceitação-rejeição gera valores de uma variável aleatória $X$ tal que

$$P(X = j) = p_j, \qquad j = 0, \ldots$$

Além disto, o número de iterações do algoritmo necessárias para gerar um valor para $X$ é uma variável aleatória geométrica com média $c$.

**Prova** Vamos calcular a probabilidade de que em uma iteração do algoritmo o valor gerado para a variável $Y$ seja $j$ e que este valor seja aceito para a variável $X$. Definindo o evento A = {o valor gerado para $Y$ é aceito para $X$ } temos que

$$\begin{aligned} P(\{Y = j\} \cap \mathrm{A}) &= P(Y = j)\, P(\mathrm{A} \mid Y = j) \\ &= q_j \frac{p_j}{c\, q_j} = \frac{p_j}{c} \end{aligned}$$

Somando a probabilidade acima com respeito a todos os valores $j$ que a variável $Y$ pode assumir teremos a probabilidade de um valor qualquer gerado para $Y$ ser aceito para $X$

$$P(\mathrm{A}) = \sum_{j} \frac{p_j}{c} = \frac{1}{c}$$

Ou seja, independentemente para cada iteração do algoritmo, um valor qualquer gerado é aceito com probabilidade $1/c$. Neste caso o número de iterações até que o valor gerado seja aceito (sucesso) é uma variável aleatória geométrica com média $c$. Além disso

$$P(X = j) = \sum_{n} P(\text{valor } j \text{ aceito na } n\text{-ésima interação})$$

$$= \sum_{n} (1-1/c)^{n-1} \frac{p_j}{c} = p_j \qquad \bullet$$

**Exemplo 6** Suponha que queremos simular valores da variável aleatória discreta $X$ que tem a seguinte distribuição de probabilidade

| $X$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $p(x_i)$ | 0.11 | 0.12 | 0.09 | 0.08 | 0.12 | 0.10 | 0.09 | 0.09 | 0.10 | 0.10 |

Poderíamos usar o método da transformação inversa para fazer a simulação porém vamos usar o método da rejeição a partir de uma variável uniforme discreta que assume cada um dos valores acima com probabilidade 0.10. De acordo com (1.15) precisamos determinar a constante $c$ tal que

$$\frac{p(x_i)}{0.10} \le c \qquad \text{para todo } i \in \{1,2,\ldots,10\}$$

ou seja, $c = 1.2$. Neste caso o algoritmo será:

| AlGORITMO 11 – EXEMPLO DO MÉTODO DA REJEIÇÃO (CASO DISCRETO) |
|---|
| P1: MAKE $U_1$ = RAND |
| P2: SET $Y = \mathrm{Int}(10\,U_1) + 1$ |
| P3: MAKE $U_2$ = RAND |
| P4: IF $U_2 \le p(Y)/0.12$, SET $X = Y$, STOP |
| P5: GO TO P1 |

ALGORITMO DA REJEIÇÃO (CASO CONTÍNUO) - De maneira análoga ao caso discreto suponha que temos um algoritmo para gerar valores de uma variável aleatória contínua $Y$ que tem função densidade de probabilidade $g(y)$. Nosso objetivo é gerar valores de uma segunda variável aleatória contínua $X$ que pode assumir os mesmos valores que $Y$ porém tem função densidade de probabilidade $f(x)$. Seja $c$ uma constante tal que

$$\frac{f(y)}{g(y)} \leq c \qquad \text{para todo } y \text{ tal que } g(y) > 0 \tag{1.16}$$

neste caso podemos escrever o algoritmo da rejeição como

| AlGORITMO 12 – MÉTODO DA REJEIÇÃO (CASO CONTÍNUO) |
|---|
| P1: GENERATE $Y$ |
| P2: MAKE U = RAND |
| P3: IF U < $f(Y)/c\,g(Y)$, SET X = Y, STOP |
| P4 RETURN TO P1 |

Como o teorema 1 existe um teorema análogo no caso contínuo (ver prova nas pags. 459/460 em [4]) que valida o algoritmo, ou seja, os valores gerados são de uma variável aleatória que tem densidade de probabilidade $f(x)$.

**Exemplo 7** Vamos usar o método da rejeição para simular uma variável aleatória contínua que tem densidade de probabilidade

$$f(x) = 20x(1-x)^3 \qquad 0 < x < 1$$

como esta variável aleatória tem sua probabilidade concentrada no intervalo (0,1) vamos utilizar como variável auxiliar $Y$ uma variável uniforme em (0,1), ou seja, $Y \sim U(0,1)$ sendo a densidade de probabilidade $g(y) = 1$ para $0 < y < 1$. Obter a constante $c$ tal que a condição (1.16) seja satisfeita significa determinar o valor de $x$ que maximiza

$$\frac{f(x)}{g(x)} = 20\,x\,(1-x)^3 \tag{1.17}$$

tomando a derivada em relação à $x$ de (1.17) e igualando-a à zero verificamos que o máximo é obtido em $x = 1/4$, o que implica em

$$\frac{f(x)}{g(x)} \leq 20\left(\frac{1}{4}\right)\left(\frac{3}{4}\right)^3 = \frac{135}{64} \equiv c$$

e que

$$\frac{f(x)}{c\ g(x)} = \frac{256}{27} x\,(1-x)^3$$

Neste caso escrevemos o algoritmo como

| AlGORITMO 13 – EXEMPLO MÉTODO DA REJEIÇÃO (CASO CONTÍNUO) |
| --- |
| P1: MAKE $\mathrm{U}_1$ = RAND, $\mathrm{U}_2$ = RAND |
| P2: IF $\mathrm{U}_2 < 256\,\mathrm{U}_1(1-\mathrm{U}_1)^3/27$, SET $\mathrm{X} = \mathrm{U}_1$, STOP |
| P3: GO TO P1 |

**Exemplo 8** Vamos apresentar agora um algoritmo baseado no método da rejeição para simular uma variável normal padrão $Z \sim N(0,1)$. Primeiro vamos simular a variável que representa o valor absoluto de $Z$. É fácil mostrar que $X = |Z|$ tem densidade de probabilidade

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2} \qquad \text{para} \quad 0 < x < \infty \tag{1.18}$$

Para simular a variável $X$ com um algoritmo baseado no método da rejeição utilizaremos como variável auxiliar $Y$ uma variável exponencial com parâmetro $\lambda = 1$

$$g(y) = e^{-x} \qquad \text{para} \quad 0 < y < \infty$$

Então

$$\frac{f(x)}{g(x)} = \sqrt{\frac{2}{\pi}}\ \exp[x - x^2/2]$$

O máximo da razão $f(x)/g(x)$ ocorre para o valor de $x$ que maximiza a função que é o argumento da exponencial. Derivando e igualando a zero a derivada em relação à $x$ a função $x - x^2/2$, obtemos que o máximo ocorre quando $x = 1$. Ou seja,

$$c = \text{Max}\,\frac{f(x)}{g(x)} = \frac{f(1)}{g(1)} = \sqrt{\frac{2e}{\pi}}$$

Neste caso

$$\frac{f(x)}{c\,g(x)} = \exp\{-\frac{(x-1)^2}{2}\}$$

e poderíamos escrever o algoritmo como

| ALGORITMO 14 |
|---|
| P1: MAKE $U_1$ = RAND, SET $Y = -\log U_1$ //gera uma variável exponencial $Y$ |
| P2: MAKE $U_2$ = RAND |
| P3: IF $U_2 < \exp\{-(Y-1)^2/2\}$, SET $X = Y$, STOP |
| P4: GO TO P1 |

Esta ainda não é a forma final do algoritmo. O valor gerado para $X$ é o valor absoluto de uma variável normal padrão $Z$. Dado $X$ determinamos o valor de $Z$ fazendo $P(Z=-X) = P(Z=X) = 1/2$. Podemos otimizar o algoritmo 14 re-escrevendo a condição do passo 3 (P3) como $-\log U_2 \geq (Y-1)^2/2$. No exemplo 1 mostramos que $-\log U$ é o valor gerado para uma variável exponencial com parâmetro 1. Neste caso podemos escrever a forma completa do algoritmo para gerar valores de uma variável normal padrão como

| ALGORITMO 15 GERAÇÃO DE UMA VARIÁVEL NORMAL PADRÃO |
|---|
| P1: MAKE $U_1$ = RAND, SET $Y_1 = -\log U_1$ //gera uma variável exponencial $Y_1$ |
| P2: MAKE $U_2$ = RAND, SET $Y_2 = -\log U_2$ //gera uma variável exponencial $Y_2$ |
| P3: IF $Y_2 < (Y_1-1)^2/2$, GO TO P1 |
| P4: MAKE $U_3$ = RAND |
| P5: IF $U_3 < 1/2$, SET $Z = Y_1$, STOP |
| P6: SET $Z = -Y_1$, STOP |

Se quisermos gerar o valor de uma variável normal $X \sim N(\mu, \sigma^2)$ bastar gerar um valor para $Z$ de acordo com o algoritmo 15 e fazer $X = \mu + \sigma\, Z$. Existe um segundo método para simular uma variável normal padrão, o chamado método polar (assim nomeado devido ao uso de coordenadas polares) que utiliza as transformações de Box-Muller [1].

# Exercícios

**1.** Construa um algoritmo para gerar uma variável aleatória que tem densidade de probabilidade

**a)** $f(x)=\begin{cases} e^x/(e-1) & 0\le x\le 1 \\ 0 & \text{caso contrário} \end{cases}$

**b)** $f(x)=\begin{cases} (x-2)/2 & 2\le x\le 3 \\ (2-x/3)/2 & 3<x\le 6 \\ 0 & \text{caso contrário} \end{cases}$

**c)** $f(x)=\begin{cases} e^{2x} & -\infty<x<0 \\ e^{-2x} & 0<x<\infty \end{cases}$

**d)** $f(x)=\begin{cases} 60x^3(1-x)^2 & 0<\mathrm{x}<1 \\ 0 & \text{caso contrário} \end{cases}$

**e)** $f(x)=\begin{cases} 30(x^2-2x^3+x^4) & 0\le x\le 1 \\ 0 & \text{caso contrário} \end{cases}$

## REFERÊNCIAS 1


1. Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; Flannery, B.P. (2002); *Numerical Methods in C++, The Art of Scientific Computing*, 2nd ed.; Cambridge University Press.
2. Knuth, D. (2000); *The Art of Computer Programming*, Vol. 2, 2nd ed., *Seminumerical Algorithms*; Addison-Wesley.
3. Meyer, P.L. (1983); *Probabilidade - Aplicações à Estatística*, 2ª ed, LTC – Livros Técnicos e Científicos Editora S.A.
4. Ross, S.M. (2002); *A First Course in Probability*, 6th ed.; Prentice Hall, Inc..

# 2. SIMULAÇÃO: APLICAÇÕES EM EXCEL

## 2.1 INTRODUÇÃO:

Simulação é a representação, durante determinado período de tempo, de fenômeno, situação ou processo do mundo real. A simulação envolve a geração de uma história artificial de um sistema, podendo-se, a partir dela, estimar como o sistema funcionaria em situação real. O comportamento do sistema é estudado a partir da construção do um Modelo de Simulação que tem a forma de um conjunto de considerações relacionadas à sua operação e que, uma vez construído e validado, pode ser usado para estudar questões relacionadas com o sistema do mundo real.
Métodos de simulação têm sido usados, por exemplo, para estudar o fluxo de tráfego em rodovias, para elaborar esquemas de produção, para resolver problemas de estoques e para estudar muitas outras situações que envolvem planejamento e organização globais. Com a simulação, elimina-se o custo decorrente da construção e operação de equipamentos dispendiosos. Esses métodos podem ser aplicados quando a experimentação direta é impossível. Salienta-se que nem sempre é possível elaborar simulações que são reproduções fiéis de experimentos reais.
O objetivo do presente capítulo é mostrar algumas das possibilidades de simulação estocástica através do Excel, utilizando seu próprio gerador de números aleatórios. Há amplas alternativas de utilização em modelos simples, permitindo seu uso de uma maneira introdutória ou em situações de menor esforço computacional.

## 2.2 USO DO EXCEL:

| |
|---|
| Esta ferramenta de análise preenche um intervalo com números aleatórios independentes retirados de uma dentre várias distribuições. Pode-se assim caracterizar indivíduos em uma população com uma distribuição de probabilidade. Por exemplo, você pode usar uma distribuição normal para caracterizar a população de alturas dos indivíduos ou pode usar uma distribuição de Bernoulli para caracterizar a população de resultados tipo "sucesso" ou "fracasso". |

Para selecionarmos variáveis aleatórias de uma distribuição de probabilidade, utilizamos o pacote de ferramentas de análise do Excel, seguindo os seguintes passos:
Selecione o menu **Ferramentas** da barra de menus. Escolha **'Análise de dados'**, a partir do menu que se abre;
Selecione **'Geração de número aleatório'**, a partir da lista da **'Análise de dados'** nessa caixa de diálogo;
Aberta a caixa de diálogo Geração de Número Aleatório, devemos introduzir um valor em cada um dos dois primeiros campos: **'Número de variáveis'** e **'Número de números aleatórios' .** O número total de números aleatórios que o Excel vai extrair é o *produto* dos números inseridos nesses dois campos.
O terceiro campo na caixa de diálogo é definido como **'Distribuição'**, no qual escolhe-se a espécie de variável aleatória que se deseja simular.
As distribuições disponíveis são: Uniforme, Normal, Bernoulli, Binomial, Poisson e Discreta.
A distribuição escolhida determina os campos que irão aparecer na caixa de diálogo, e estaremos verificando em detalhes a utilização de cada uma delas.
Clique no botão de rádio ao lado de **'Intervalo de saída:'** e, em seguida, clique no ponteiro do mouse no campo adjacente, de tal modo que um campo de inserção surja. Digite ou selecione, com o mouse, o intervalo de células nos quais deseja armazenar os números gerados.
Pode-se optar pelo uso de uma semente para geração dos números aleatórios, podendo voltar a usar este valor para produzir os mesmos números aleatórios posteriormente.

## 2.3 UNIFORME CONTÍNUA:

As variáveis uniformes contínuas caracterizam-se por possuírem mesma probabilidade para todos os valores em um determinado intervalo, determinado por seus limites inferior e superior. Um aplicativo comum usa uma distribuição uniforme no intervalo de 0 a 1

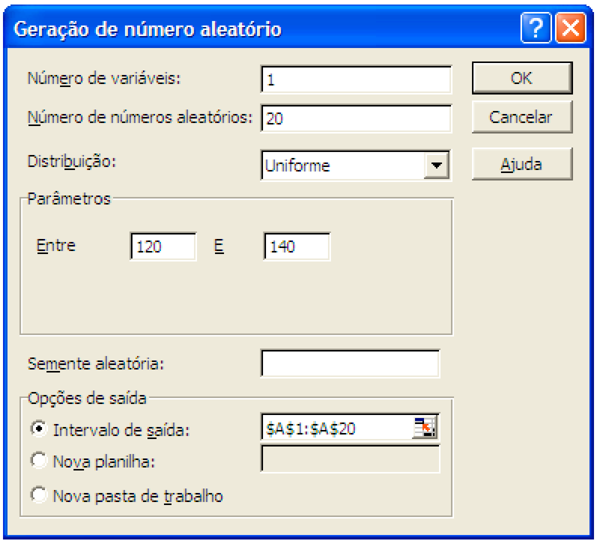
Para gerar números aleatórios de acordo com uma distribuição uniforme contínua de probabilidades, siga os passos relacionados no item 2.2 e selecione 'Uniforme' para 'Distribuição'. Em 'Parâmetros', insira os limites inferiores e superiores do intervalo desejado para a geração dos números aleatórios.

<u>Exemplo de Aplicação</u>:

*Considere a variável aleatória* ***X****, que representa o tempo de vôo de um avião voando entre os aeroportos A e B. Suponha que dados suficientes de vôos reais indiquem que o tempo de vôo pode ser qualquer valor no intervalo de 120 a 140 minutos. Simule o tempo de vôo de 20 aviões entre A e B?*

– Geração dos números aleatórios para solução:

Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo.




Resultados da Simulação:
Foram gerados os seguintes valores:

| | | | | |
|---|---|---|---|---|
| 127,6400 | 122,0136 | 131,9297 | 137,9821 | 137,6922 |
| 139,1693 | 120,2899 | 128,1484 | 137,2649 | 122,7717 |
| 124,9007 | 120,9095 | 120,6476 | 123,2826 | 124,3922 |
| 120,3418 | 125,7009 | 126,8618 | 131,0727 | 127,1474 |

Obtendo-se as seguintes estatísticas para esta simulação

| | |
|---|---|
| Mediana: | 126,28 |
| Média | 127,51 |
| Desvio-padrão: | 6,33 |
| Mínimo: | 120,29 |
| Máximo: | 139,17 |

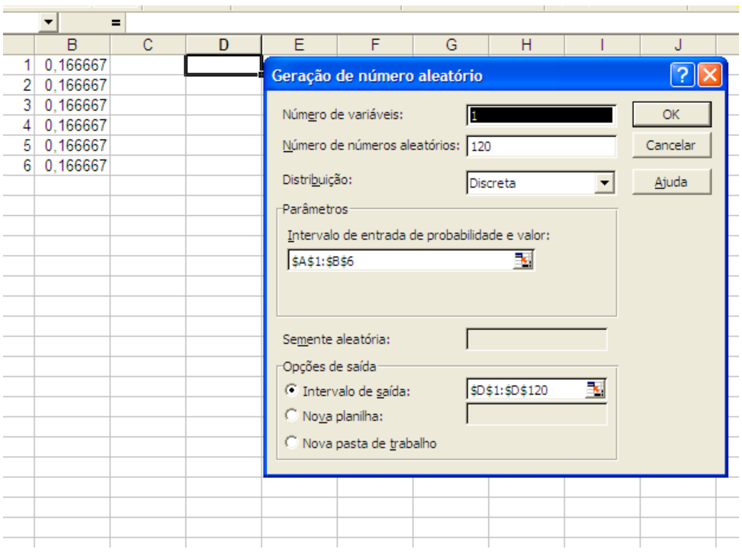
– Comentário:

- Apesar de termos gerado uma amostra pequena, verificamos que as estimativas oferecidas pela média e pelo desvio-padrão amostral estão próximas dos verdadeiros valores destes parâmetros.

## 2.4 UNIFORME DISCRETA:

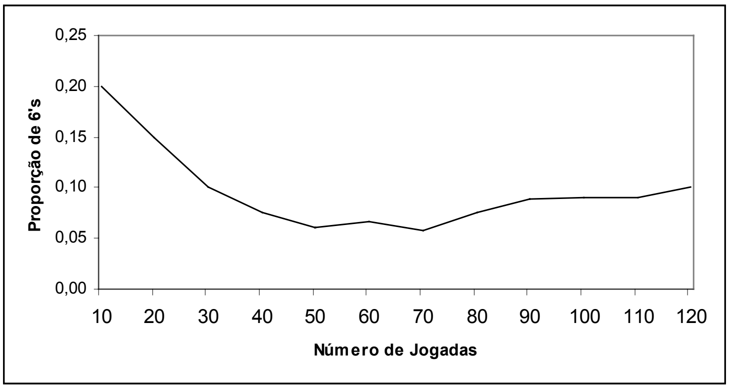
As variáveis uniformes discreta caracterizam-se por possuírem a mesma probabilidade em um conjunto finito de valores.
Para gerar números aleatórios de acordo com uma distribuição discreta de probabilidades, siga os passos relacionados no item 2.2 e selecione 'Discreta' para 'Distribuição'. Monte uma tabela com os valores da variável aleatória, associando a cada um deles a mesma probabilidade.

Exemplo de Aplicação:
Representando as seis faces de um dado pelos algarismos 1, 2, 3, 4, 5 e 6, simule 120 jogadas de um dado equilibrado. Determine, também, a proporção acumulada de 6's após cada décima jogada, e faça um diagrama entre a proporção de 6's e o número de jogadas simuladas. O diagrama parece confirmar a Lei dos Grandes Números?
Geração dos números aleatórios para solução:
Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo.




– Resultados da Simulação:

Da maneira descrita no parágrafo anterior, foram gerados valores de 1 a 6, em 120 linhas de uma dada coluna. Através do comando CONT.SE(Contar.se) contamos a quantidade de 6's gerados pela simulação a cada 10ª. jogada. Poderíamos ter feito o mesmo tipo de análise para cada um dos números gerados. Os resultados estão apresentados abaixo:




| Jogadas | Qte. | Proporção |
|---|---|---|
| 10 | 2 | 0,2000 |
| 20 | 3 | 0,1500 |
| 30 | 3 | 0,1000 |
| 50 | 3 | 0,0600 |
| 40 | 3 | 0,0750 |
| 50 | 3 | 0,0600 |
| 60 | 4 | 0,0667 |
| 70 | 4 | 0,0571 |
| 80 | 6 | 0,0750 |
| 90 | 8 | 0,0889 |
| 100 | 9 | 0,0900 |
| 110 | 10 | 0,0909 |
| 120 | 12 | 0,1000 |

– Comentário:

- Verifica-se que à medida que aumenta-se o número de jogadas, a proporção de 6's se aproxima de seu valor esperado, ou seja, de 1/6., o que confirmaria a Lei dos Grandes Números, entretanto necessitaríamos de uma quantidade maior de jogadas simuladas para que esta convergência se tornasse mais clara.

## 2.5 BERNOULLI:

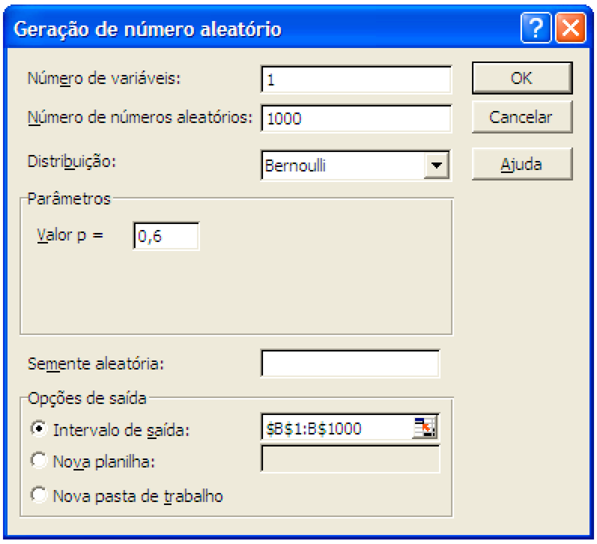
As variáveis aleatórias de Bernoulli são caracterizadas pela probabilidade de sucesso (valor p) em uma determinada tentativa e têm o valor 0 ou 1. Por exemplo, você pode formular uma variável aleatória uniforme no intervalo de 0 a 1. Se a variável for menor ou igual à probabilidade de sucesso, atribui-se o valor 1 à variável aleatória Bernoulli; caso contrário, será atribuído o valor 0.

Para gerar números aleatórios de acordo com uma distribuição Bernoulli de probabilidades, siga os passos relacionados no item 2.2 e selecione 'Bernoulli' para 'Distribuição'. Em 'Parâmetros', insira o valor da probabilidade de sucesso (*p*) desejado.

<u>Exemplo de Aplicação</u>:
Suponha que uma moeda é viciada, de tal sorte que favoreça mais cara que coroa. Para estimar a probabilidade de cara você pode lançá-la um certo número de vezes. Suponha que a probabilidade de ocorrer cara seja p = 0,60. Simule o lançamento da moeda e verifique sua estimativa para vários valores de n (quantidade de lançamentos da moeda):
Construa um gráfico dos valores assumidos por seu estimador de p, para vários valores de n;
Altere algumas vezes o valor da probabilidade de ocorrer cara (p) e verifique o comportamento do seu estimador de p

– Geração dos números aleatórios:

Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo. O resultado será a simulação do lançamento de uma moeda 1000 vezes, com o valor 1 representando cara e 0, coroa.



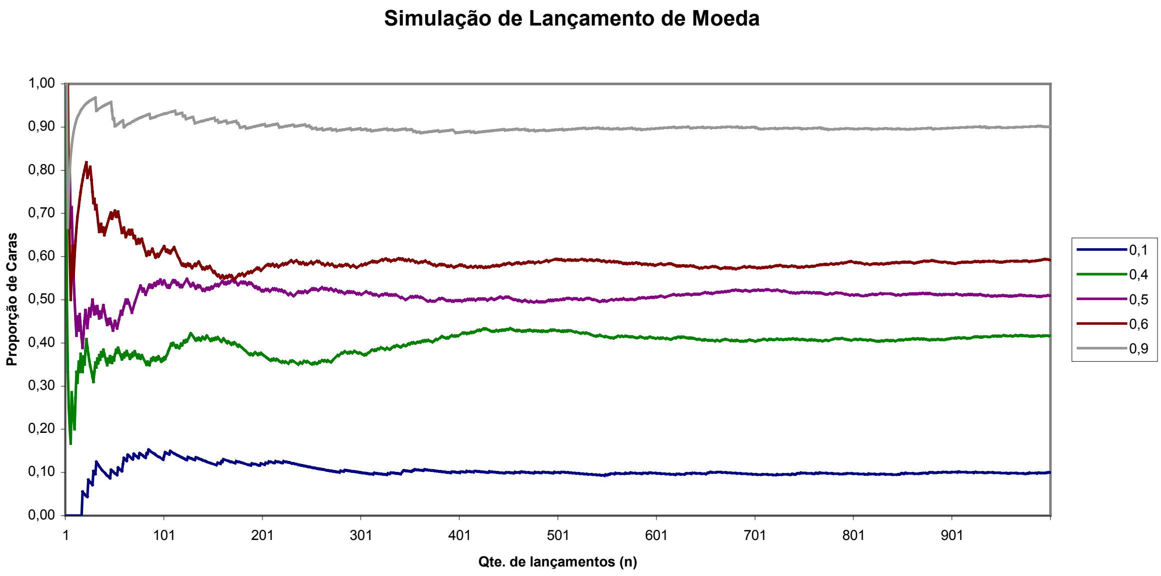

– Resultados da Simulação:

Da maneira descrita no parágrafo anterior, foram gerados 1.000 números aleatórios de uma distribuição de Bernoulli, com probabilidade de sucesso dados por *p = 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8 e 0,9* e verificada a estimativa de probabilidade após cada lançamento. Estes valores estimados para cada uma das amostras estão apresentados na tabela abaixo. Para simplificar a apresentação, consideramos apenas as seqüências referentes às probabilidades

*p = 0,1; 0,4; 0,5; 0,6; e 0,9.*

Resultado de simulação
(1.000 lançamentos)

| Probabilidade Cara | Quantidade Cara | Estimativa Probabilidade |
|---|---|---|
| 0,1 | 100 | 0,10 |
| 0,2 | 208 | 0,21 |
| 0,3 | 297 | 0,30 |
| 0,4 | 417 | 0,42 |
| 0,5 | 510 | 0,51 |
| 0,6 | 592 | 0,59 |
| 0,7 | 700 | 0,70 |
| 0,8 | 805 | 0,81 |
| 0,9 | 901 | 0,90 |



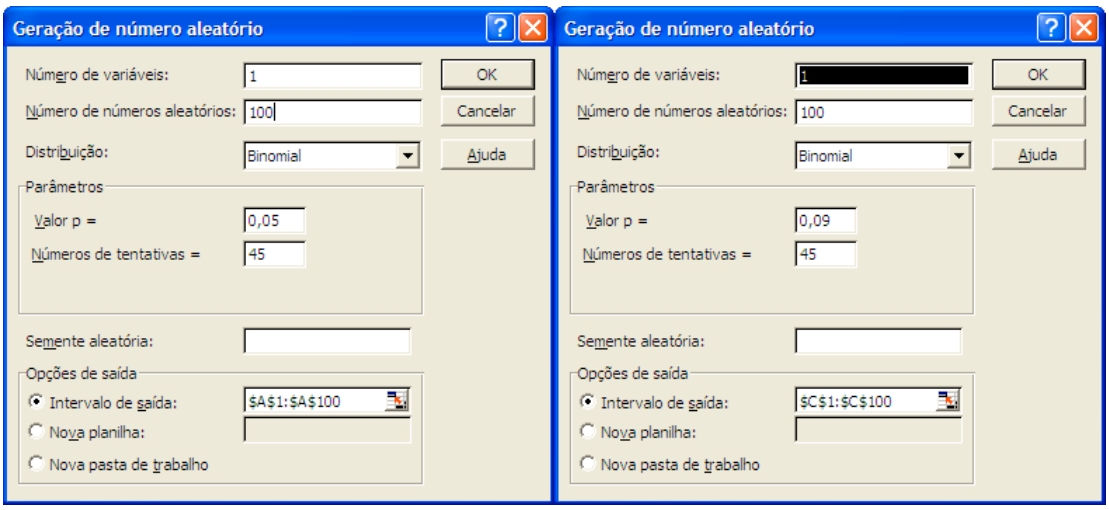

– Comentários:

- Verifica-se, na tabela acima, que, para 1.000 lançamentos, os resultados das estimativas de *p (probabilidade de sucesso)* convergem para seu verdadeiro, em todos os casos estudados.
- Entretanto, percebemos que as 'velocidades' destas convergências são diferentes. Os valores de *p* mais próximos de 0,5 *(p=0,4; 0,5 e 0,6)* aparentam convergir mais lentamente que aqueles valores mais próximos de 0 *(p=0,1)* ou 1 *(p=0,9)*. Há uma relação entre a taxa de convergência de estimadores e sua variabilidade e, no caso da Bernoulli, a variância é máxima para *p=0,5*, sendo elevada em suas proximidades.

## 2.6 BINOMIAL:

Caracterizada por uma probabilidade de sucesso (*p*) para várias tentativas (*n*). Uma variável aleatória binomial é a soma do número de tentativas de variáveis aleatórias de Bernoulli.
Para gerar números aleatórios de acordo com uma distribuição Binomial de probabilidades, siga os passos relacionados no item 2.2 e selecione 'Binomial' para 'Distribuição'. Em 'Parâmetros', insira os valores referentes à probabilidade de sucesso (*p*) e número de tentativas (*n*), para caracterizar a distribuição.
<u>Exemplo de Aplicação</u>:
A especificação para a fabricação de um determinado tipo de lâmpada aceita até 5 por cento de lâmpadas defeituosas no decorrer de seu processo de fabricação. Um inspetor de controle de qualidade extrai uma amostra de 45 lâmpadas de um lote produzido recentemente. Se nenhuma dessas lâmpadas apresentar defeito, o lote será aprovado pela inspeção e liberado para transporte.
Simule a inspeção de 100 lotes de lâmpadas originárias de um processo de produção que esteja conforme o especificado. Verifique quantos lotes serão reprovados pela inspeção e, portanto impedidos de serem transportados.
Simule a inspeção de 100 lotes de lâmpadas originárias de um processo de fabricação que produza nove por cento das lâmpadas com defeitos. Verifique quantos lotes serão aprovados pela inspeção e, portanto autorizados a serem embarcados.

Geração dos números aleatórios para solução:
Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, preencha os dados necessários à simulação, conforme ilustrado na figura abaixo. Repetiremos este procedimento para gerarmos o resultado da inspeção do processo referente ao item 0, modificando a probabilidade de fabricação de lâmpadas defeituosas para 0,09



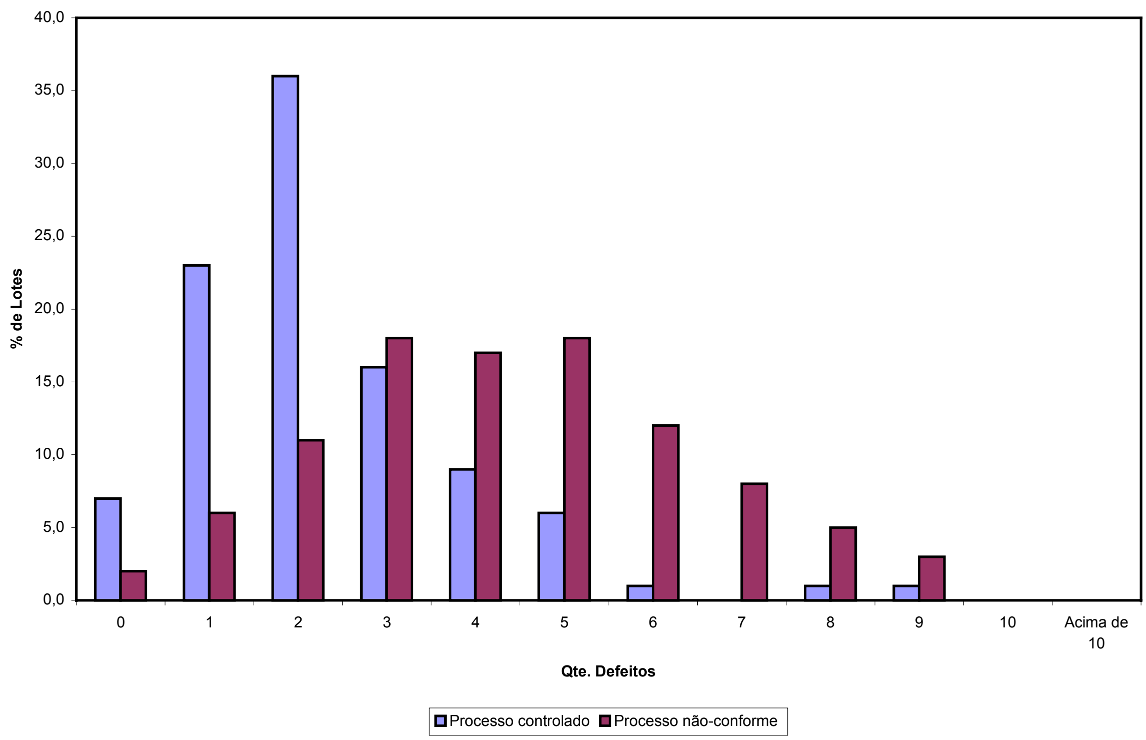
– Resultados da Simulação:

Foi simulada a inspeção de 100 lotes de lâmpadas de um processo sob controle, ou seja, com probabilidade de defeitos *p=0,5*, e 100 lotes de um processo não-conforme, com probabilidades de defeitos *p=0,09*, de acordo com o procedimento descrito no item anterior.

Os resultados estão apresentados a seguir:

| | Processo controlado *(p=0,05)* | | Processo não-conforme *(p=0,09)* | |
|---|---|---|---|---|
| Qte. defeitos por lote | Quantidade de Lotes | % | Quantidade de Lotes | % |
| 0 | 7 | 7,0 | 2 | 2,0 |
| 1 | 23 | 23,0 | 6 | 6,0 |
| 2 | 36 | 36,0 | 11 | 11,0 |
| 3 | 16 | 16,0 | 18 | 18,0 |
| 4 | 9 | 9,0 | 17 | 17,0 |
| 5 | 6 | 6,0 | 18 | 18,0 |
| 6 | 1 | 1,0 | 12 | 12,0 |
| 7 | 0 | 0,0 | 8 | 8,0 |
| 8 | 1 | 1,0 | 5 | 5,0 |
| 9 | 1 | 1,0 | 3 | 3,0 |
| 10 | 0 | 0,0 | 0 | 0,0 |
| Acima de 10 | 0 | 0,0 | 0 | 0,0 |
| *Total* | *100* | *100,0* | *100* | *100,0* |

**Quantidade de Defeitos por Lote**



– Comentários:

- Verifica-se que o critério de aceitação do lote é muito rigoroso, pois seriam aceitos apenas 7% dos lotes provenientes de um processo de fabricação sob controle.
- Pelo histograma seria razoável imaginar um critério de aceitação mais tolerante de maneira a incorporar uma quantidade maior de lotes oriundos de processos controlados, embora esta alternativa apresente um aumento na proporção de lotes que seriam aceitos, casos fossem provenientes de um processo de fabricação não-conforme. A regra a ser adotada será aquela que fornecer o balanço mais razoável entre estes dois tipos de erros de decisão.
- Este tipo de problema pode ser facilmente equacionado, utilizando-se das propriedades da distribuição exponencial, embora este tipo de simulação facilite o entendimento da situação real.

## 2.7 POISSON:

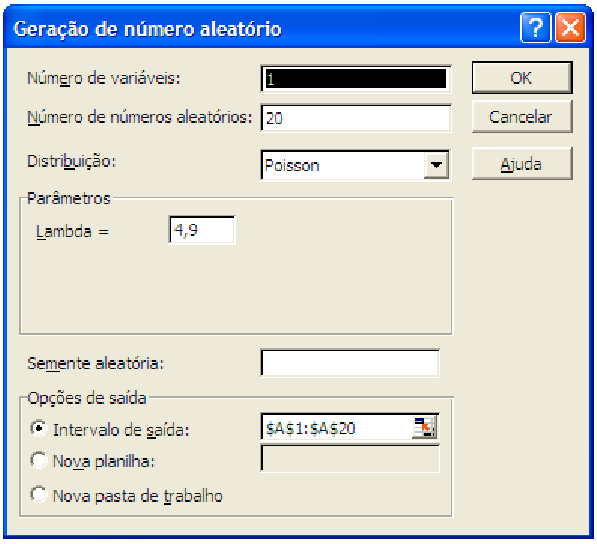
Caracterizada por um valor lambda equivalente a sua média. A distribuição Poisson é usada com freqüência para caracterizar o número de eventos que ocorre por unidade de tempo — por exemplo, uma taxa média na qual os carros chegam a um posto de pedágio.

Para gerar números aleatórios de acordo com uma distribuição Poisson de probabilidades, siga os passos relacionados no item 2.2 e selecione 'Poisson' para 'Distribuição'. Em 'Parâmetros', insira o valor referente à média (taxa) da distribuição desejada.

<u>Exemplo de Aplicação</u>:

Considere que a chegada de carros a um posto de pedágio de uma ponte durante um período qualquer de um minuto na parte da tarde é uma variável aleatória com distribuição de Poisson com média 4,9. Simule a chegada de carros a este posto de pedágio para o período de 1 minuto. Repita este procedimento 20 vezes.

Geração dos números aleatórios para solução:
Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo. Obteremos então a simulação da chegada de carros em 20 intervalos de um minuto.




– Resultados da Simulação:

Da maneira descrita no parágrafo anterior, simulamos a chegada de veículos em um posto de pedágio durante 20 intervalos de 1 minuto, de acordo a uma distribuição de Poisson com parâmetro $\lambda = 4,9$, e calculamos a média e a variância desta amostra.

Os resultados estão apresentados a seguir:

| | | | | |
|---|---|---|---|---|
| 6 | 3 | 8 | 5 | 5 |
| 5 | 3 | 2 | 5 | 7 |
| 4 | 4 | 4 | 3 | 4 |
| 9 | 4 | 7 | 2 | 3 |

Média amostral: 4,65
Variância amostral: 3,71

– Comentário:

- Apesar de termos gerado uma amostra pequena, verificamos que as estimativa oferecidas pela média amostral está próxima do verdadeiro valor deste parâmetro. Entretanto, a estimativa oferecida pela variância amostral distancia-se do valor esperado do parâmetro, que é o mesmo valor da média, dada que a amostra provém de uma distribuição de Poisson. Este fato pode ser explicado pelo tamanho relativamente pequeno da amostra.

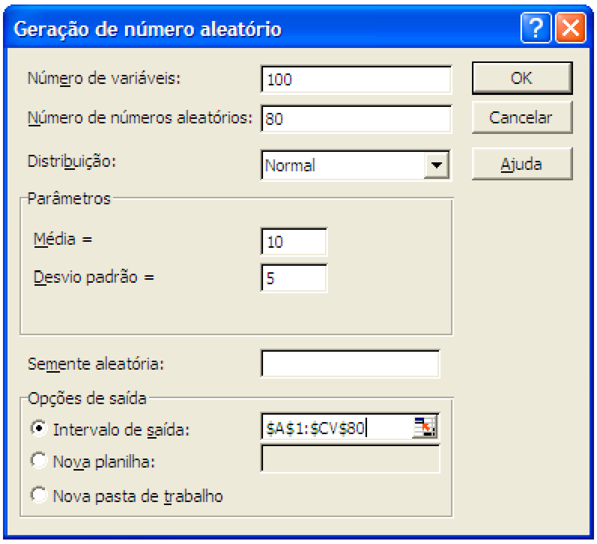
## 2.8 NORMAL:

A distribuição normal é caracterizada por uma média e um desvio padrão. Para gerar números aleatórios de acordo a uma determinada distribuição Normal de probabilidades, siga os passos relacionados no item 2.2 e selecione 'Normal'para 'Distribuição'. Em 'Parâmetros', insira os valores referentes à média e ao desvio-padrão da distribuição desejada.

<u>Exemplo de Aplicação</u>:

Simule a coleta de 80 observações de uma variável normal (10,25) e obtenha o intervalo de 92% de confiança para a média. Repita esse procedimento 100 vezes, isto é, obtenha 100 intervalos de confiança. Verifique quantos deles contêm a verdadeira média.

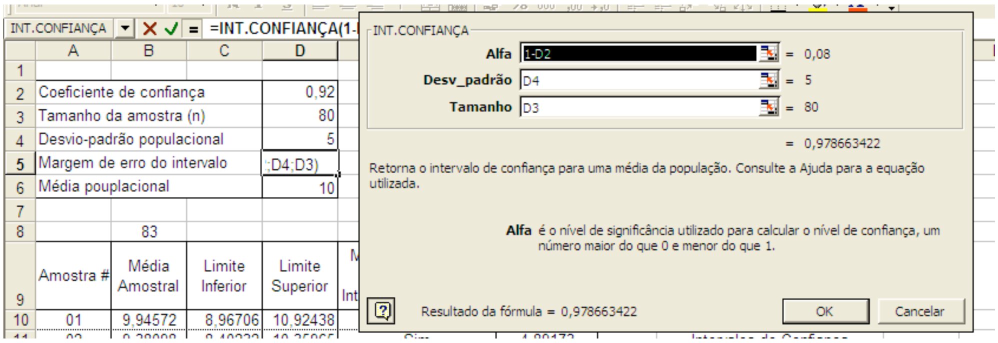
Geração dos números aleatórios para solução:
Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo. O resultado será a simulação de 100 amostras aleatórias com as características solicitadas.




– Resultados da Simulação:

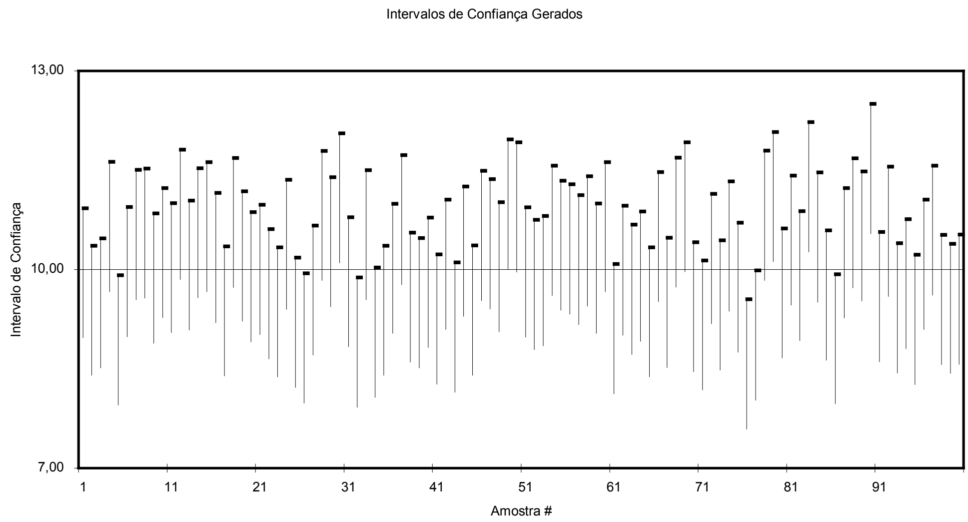
Seguimos o procedimento descrito no item anterior e simulamos a coleta de 100 amostras com distribuição Normal, de acordo com os parâmetros solicitados, com 80 elementos cada amostra. Calculamos assim a média e o desvio-padrão de todas as colunas da planilha, já que dispusemos cada amostra em uma coluna. Para facilidade, copiamos estes resultados em outra planilha, transpondo as médias amostrais em uma única coluna, a partir da qual estimamos os limites inferior e superior de cada intervalo. Para determinação da margem de erro do intervalo utilizamos o comando Excel INT.CONFIANÇA, conforme figura abaixo, obtendo-se o valor de

0,978663422. Salienta-se que neste comando deve-se utilizar o nível de significância do Intervalo de Confiança ($\alpha = 1 -$ coeficiente de confiança).




Apresentamos a seguir uma visão da planilha efetuada e um resumo dos resultados obtidos.

| | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 9 | Amostra # | Média Amostral | Limite Inferior | Limite Superior | Média Populacional pertencente ao Intervalo de Confiança | Desvio-padrão Amostral |
| 10 | 01 | 9,94572 | 8,96706 | 10,92438 | Sim | 4,78629 |
| 11 | 02 | 9,38098 | 8,40232 | 10,35965 | Sim | 4,89173 |
| 12 | 03 | 9,49181 | 8,51315 | 10,47047 | Sim | 4,62398 |
| 13 | 04 | 10,64575 | 9,66708 | 11,62441 | Sim | 6,03825 |
| 14 | 05 | 8,93383 | 7,95517 | 9,91250 | Não | 5,80473 |
| 15 | 06 | 9,96337 | 8,98471 | 10,94203 | Sim | 4,54152 |
| 16 | 07 | 10,52570 | 9,54704 | 11,50437 | Sim | 5,18053 |
| 17 | 08 | 10,54599 | 9,56732 | 11,52465 | Sim | 5,31770 |
| 18 | 09 | 9,86625 | 8,88759 | 10,84491 | Sim | 4,99802 |
| 19 | 10 | 10,25131 | 9,27264 | 11,22997 | Sim | 4,40854 |
| 20 | 11 | 10,02501 | 9,04635 | 11,00367 | Sim | 5,56427 |
| 21 | 12 | 10,82862 | 9,84995 | 11,80728 | Sim | 5,13877 |
| 22 | 13 | 10,06346 | 9,08480 | 11,04212 | Sim | 4,41167 |
| 23 | 14 | 10,55164 | 9,57297 | 11,53030 | Sim | 4,57619 |
| 24 | 15 | 10,64432 | 9,66566 | 11,62298 | Sim | 4,48810 |
| 25 | 16 | 10,17665 | 9,19798 | 11,15531 | Sim | 5,46493 |
| 26 | 17 | 9,37103 | 8,39236 | 10,34969 | Sim | 5,50090 |
| 27 | 18 | 10,70659 | 9,72792 | 11,68525 | Sim | 5,60830 |
| 28 | 19 | 10,20011 | 9,22145 | 11,17878 | Sim | 5,23805 |
| 29 | 20 | 9,88516 | 8,90650 | 10,86383 | Sim | 5,85942 |
| 30 | 21 | 9,99652 | 9,01786 | 10,97519 | Sim | 5,19551 |
| 31 | 22 | 9,62919 | 8,65052 | 10,60785 | Sim | 5,19721 |
| 32 | 23 | 9,35649 | 8,37783 | 10,33515 | Sim | 4,47654 |
| 33 | 24 | 10,37742 | 9,39875 | 11,35608 | Sim | 4,98636 |
| 34 | 25 | 9,20009 | 8,22142 | 10,17875 | Sim | 5,43149 |
| 35 | 26 | 8,96281 | 7,98415 | 9,94147 | Não | 5,19195 |




Intervalos de Confiança
contendo a Média Populacional

| | Qte. | % |
|---|---|---|
| Sim | 89 | 89,0 |
| Não | 11 | 11,0 |
| | 100 | |

– Comentário:

- Pelo gráfico pode-se verificar a posição dos 100 intervalos de confiança obtidos a partir das amostras simuladas. Destes, 11 deles não contém a verdadeira média, valor bastante próximo do valor esperado (8 intervalos de confiança não contendo a média populacional), dado que o grau de confiança adotado foi 92%.

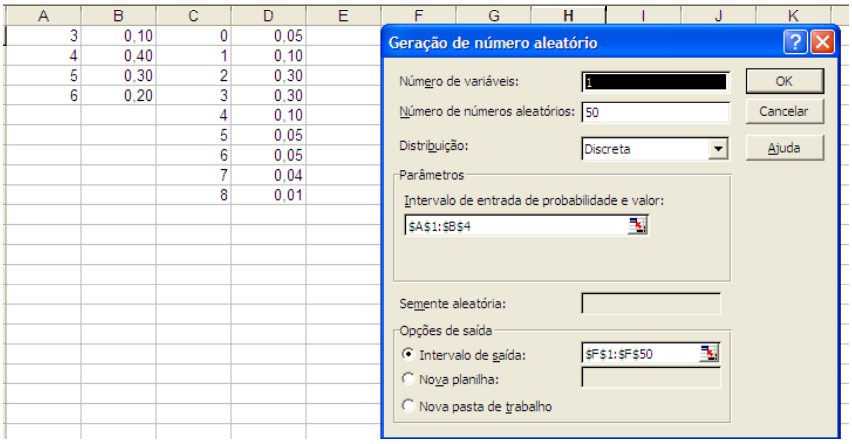
## 2.9 DISCRETA:

Caracterizada por uma distribuição de probabilidade discreta. O intervalo deve conter duas colunas: a coluna da esquerda contém valores e a da direita deve conter probabilidades associadas ao valor nesta linha. A soma das probabilidades deve ser 1.

Para gerar números aleatórios de acordo a uma determinada distribuição discreta de probabilidades, siga os passos relacionados no item 2.2 e selecione '**Discreta' para 'Distribuição**'. Em '**Intervalo de entrada de probabilidade e valor**', insira o endereço do intervalo de células que contém a distribuição de probabilidades.

<u>Exemplo de Aplicação</u>:
Dependendo da disponibilidade de peças, uma companhia pode fabricar semanalmente 3, 4, 5 ou 6 unidades de determinado artigo, com as probabilidades de 0,10; 0,40; 0,30 e 0,20, respectivamente. As probabilidades de haver uma procura semanal por 0, 1, 2, 3, ..., ou 8 unidades são respectivamente: 0,05 0,10 0,30 0,30 0,10 0,05 0,05 0,04 0,01.
Se uma unidade é vendida na semana em que é fabricada, gera um lucro de $100; este lucro é reduzido de $20 cada semana que uma unidade permanece estocada. Simule as operações desta companhia durante 50 semanas consecutivas, e estime seu lucro semanal esperado.

Geração dos números aleatórios para solução:
Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo. Repita a procedimento considerando os valores e suas correspondentes probabilidades, compreendidos no intervalo entre as células C1:D8. O resultado será a simulação da operação da empresa durante 50 semanas consecutivas.




– Resultados da Simulação:

Apresentamos abaixo uma planilha com os resultados da análise da simulação efetuada. Os valores das colunas 'Fabricação Semanal' e 'Demanda Semanal' foram gerados de acordo com os procedimentos descritos acima. Consideramos como Venda Semanal como o mínimo entre a demanda e o nível de estoque da semana, considerando a produção semanal em questão. Assim, pressupõe-se que as unidades vendidas inicialmente são aquelas produzidas durante a semana.

| Semana | Fabricação Semanal | Demanda Semanal | Venda Semanal | Estoque Semanal | Lucro na Semana | Custo Estoque | Lucro Final |
|---|---|---|---|---|---|---|---|
| 1 | 4 | 5 | 4 | 0 | 400 | 0,00 | 400,00 |
| 2 | 4 | 3 | 3 | 1 | 300 | (20,00) | 280,00 |
| 3 | 4 | 0 | 0 | 5 | 0 | (100,00) | (100,00) |
| 4 | 4 | 3 | 3 | 6 | 300 | (120,00) | 180,00 |
| 5 | 6 | 1 | 1 | 11 | 100 | (220,00) | (120,00) |
| 6 | 5 | 3 | 3 | 13 | 300 | (260,00) | 40,00 |
| 7 | 5 | 2 | 2 | 16 | 200 | (320,00) | (120,00) |
| 8 | 4 | 2 | 2 | 18 | 200 | (360,00) | (160,00) |
| 9 | 4 | 3 | 3 | 19 | 300 | (380,00) | (80,00) |
| 10 | 4 | 1 | 1 | 22 | 100 | (440,00) | (340,00) |
| 11 | 6 | 3 | 3 | 25 | 300 | (500,00) | (200,00) |
| 12 | 3 | 3 | 3 | 25 | 300 | (500,00) | (200,00) |
| 13 | 6 | 1 | 1 | 30 | 100 | (600,00) | (500,00) |
| 14 | 5 | 2 | 2 | 33 | 200 | (660,00) | (460,00) |
| 15 | 4 | 4 | 4 | 33 | 400 | (660,00) | (260,00) |
| 16 | 4 | 3 | 3 | 34 | 300 | (680,00) | (380,00) |
| 17 | 3 | 2 | 2 | 35 | 200 | (700,00) | (500,00) |
| 18 | 3 | 2 | 2 | 36 | 200 | (720,00) | (520,00) |
| 19 | 4 | 3 | 3 | 37 | 300 | (740,00) | (440,00) |
| 20 | 6 | 2 | 2 | 41 | 200 | (820,00) | (620,00) |
| 21 | 4 | 3 | 3 | 42 | 300 | (840,00) | (540,00) |
| 22 | 5 | 2 | 2 | 45 | 200 | (900,00) | (700,00) |
| 23 | 4 | 3 | 3 | 46 | 300 | (920,00) | (620,00) |
| 24 | 6 | 7 | 7 | 45 | 700 | (900,00) | (200,00) |
| 25 | 3 | 3 | 3 | 45 | 300 | (900,00) | (600,00) |
| 26 | 5 | 5 | 5 | 45 | 500 | (900,00) | (400,00) |
| 27 | 6 | 2 | 2 | 49 | 200 | (980,00) | (780,00) |
| 28 | 4 | 3 | 3 | 50 | 300 | (1.000,00) | (700,00) |
| 29 | 4 | 1 | 1 | 53 | 100 | (1.060,00) | (960,00) |
| 30 | 5 | 3 | 3 | 55 | 300 | (1.100,00) | (800,00) |
| 31 | 4 | 2 | 2 | 57 | 200 | (1.140,00) | (940,00) |
| 32 | 4 | 4 | 4 | 57 | 400 | (1.140,00) | (740,00) |
| 33 | 4 | 3 | 3 | 58 | 300 | (1.160,00) | (860,00) |
| 34 | 4 | 3 | 3 | 59 | 300 | (1.180,00) | (880,00) |
| 35 | 3 | 7 | 7 | 55 | 700 | (1.100,00) | (400,00) |
| 36 | 6 | 3 | 3 | 58 | 300 | (1.160,00) | (860,00) |
| 37 | 4 | 1 | 1 | 61 | 100 | (1.220,00) | (1.120,00) |
| 38 | 4 | 2 | 2 | 63 | 200 | (1.260,00) | (1.060,00) |
| 39 | 5 | 3 | 3 | 65 | 300 | (1.300,00) | (1.000,00) |
| 40 | 4 | 4 | 4 | 65 | 400 | (1.300,00) | (900,00) |
| 41 | 4 | 3 | 3 | 66 | 300 | (1.320,00) | (1.020,00) |
| 42 | 5 | 2 | 2 | 69 | 200 | (1.380,00) | (1.180,00) |
| 43 | 6 | 4 | 4 | 71 | 400 | (1.420,00) | (1.020,00) |
| 44 | 4 | 2 | 2 | 73 | 200 | (1.460,00) | (1.260,00) |
| 45 | 4 | 3 | 3 | 74 | 300 | (1.480,00) | (1.180,00) |
| 46 | 4 | 0 | 0 | 78 | 0 | (1.560,00) | (1.560,00) |
| 47 | 6 | 3 | 3 | 81 | 300 | (1.620,00) | (1.320,00) |
| 48 | 4 | 4 | 4 | 81 | 400 | (1.620,00) | (1.220,00) |
| 49 | 4 | 2 | 2 | 83 | 200 | (1.660,00) | (1.460,00) |
| 50 | 4 | 2 | 2 | 85 | 200 | (1.700,00) | (1.500,00) |
| | | | | | *Total* | | *(31.880,00)* |

– Comentários:

- Verifica-se a forte tendência de crescimento dos prejuízos acumulados semanalmente devido às diferenças entre as distribuições de probabilidade da demanda e da fabricação. Poderíamos repetir a simulação para testar quais seriam os parâmetros destas distribuições que poderiam suprimir o prejuízo verificado.
- Salienta-se que deveríamos repetir muitas vezes esta simulação no sentido de obtermos estimativas melhores dos parâmetros e das probabilidades da distribuição do resultado final desta empresa, permitindo uma análise mais acurada de suas operações.

## 2.10 APLICAÇÃO:

Apresentamos dois exemplos de aplicação buscando mostrar a facilidade em modelarmos situações reais ao utilizarmos as facilidades que uma planilha de cálculo pode oferecer em termos de montagem e utilização. As planilhas podem ser montadas de maneira a permitir análises envolvendo a variação de outros parâmetros que não sejam aqueles de caráter aleatório, permitindo montar vários cenários, necessários, por exemplo, em um processo de tomada de decisão. A complexidade da estrutura do modelo estará diretamente relacionada com a complexidade da situação em análise.

Salientamos que, de maneira geral, simulações devem ser efetuadas uma grande quantidade de vezes, de maneira a permitir estimativas razoáveis com relação aos parâmetros das distribuições de probabilidades das variáveis aleatórias envolvidas nos modelos. As estimativas mais úteis seriam as da esperança, variância e das probabilidades associadas aos valores das variáveis aleatórias.

Aplicação 1:
*Uma grande máquina industrial tem 3 rolamentos diferentes que quebram de tempos em tempos. A probabilidade da vida útil (em horas de operação) de um rolamento está dada na tabela abaixo:*

| ***Vida do Rolamento (horas)*** | ***Probabilidade*** | ***Vida do Rolamento (horas)*** | ***Probabilidade*** |
|---|---|---|---|
| *1.000* | *0,10* | *1.500* | *0,12* |
| *1.100* | *0,13* | *1.600* | *0,02* |
| *1.200* | *0,25* | *1.700* | *0,06* |
| *1.300* | *0,13* | *1.800* | *0,05* |
| *1.400* | *0,09* | *1.900* | *0,05* |

*Quando um rolamento quebra, a máquina pára e um mecânico é chamado para instalar um novo rolamento no lugar do que quebrou. O tempo que o mecânico demora para chegar ao rolamento quebrado também é uma variável aleatória, com a distribuição dada na tabela abaixo:*

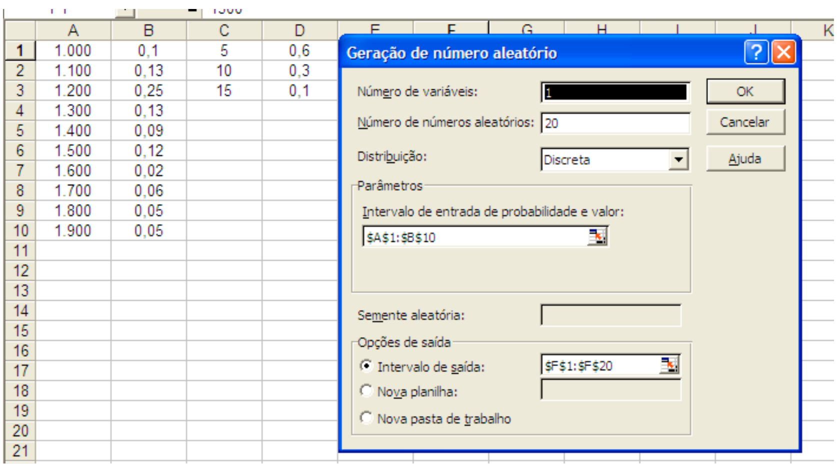
| ***Tempo de Espera (minutos)*** | ***Probabilidade*** |
|---|---|
| *5* | *0,60* |
| *10* | *0,30* |

| | |
|---|---|
| *15* | *0,10* |

*Cada minuto que a máquina fica parada custa $5 e o custo do mecânico é de $0,20/ minuto trabalhado substituindo o rolamento. O mecânico demora 20 minutos para trocar um rolamento, 30 minutos para trocar 2 e 40 minutos para trocar os 3. Cada rolamento novo custa $16. Alguém sugeriu que ao quebrar um dos rolamentos, se fizesse logo a troca dos 3. Deseja-se avaliar a situação do ponto de vista econômico. Analise a situação durante um longo período de tempo, fazendo uma simulação de 20.000 horas (pouco mais de 2 anos).*

– Geração dos números aleatórios para solução:

Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo. Repita a mesma operação nas colunas G e H para obter-se os tempos de vida dos outros dois rolamentos. Faça o mesmo procedimento quatro outras vezes, considerando agora os valores e suas correspondentes probabilidades, compreendidos no intervalo entre as células C1:D8, para obter-se o tempo de espera para conserto de cada um destes rolamentos, além do tempo de espera para a troca dos três rolamentos conjuntamente. O resultado será a simulação da operação desta fábrica durante 20.000 horas. Para facilitar a montagem da estrutura da planilha, não consideramos a ocorrência de manutenção em mais de uma máquina ao mesmo tempo.




– Resultados da Simulação:

Seguimos os procedimentos citados anteriormente e geramos 3 colunas ‘Vida útil’ para cada um dos rolamentos, de acordo com a primeira distribuição de probabilidades. A coluna ‘Espera’ para cada um dos rolamentos recebeu os valores gerados a partir da segunda distribuição de probabilidades. Os resultados estão apresentados abaixo:

| Rolamento A | | | | | Rolamento B | | | | | Rolamento C | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ctr | Vida util | Tempo da Falha | Espera | Custo Troca | Ctr | Vida util | Tempo da Falha | Espera | Custo Troca | Ctr | Vida util | Tempo da Falha | Espera | Custo Troca |
| 1 | 1300 | 1300 | 15 | 198,00 | 1 | 1400 | 1400 | 15 | 198,00 | 1 | 1300 | 1300 | 10 | 172,00 |
| 1 | 1200 | 2500 | 5 | 146,00 | 1 | 1700 | 3100 | 5 | 146,00 | 1 | 1900 | 3200 | 10 | 172,00 |
| 1 | 1600 | 4100 | 5 | 146,00 | 1 | 1300 | 4400 | 5 | 146,00 | 1 | 1400 | 4600 | 10 | 172,00 |
| 1 | 1000 | 5100 | 5 | 146,00 | 1 | 1100 | 5500 | 5 | 146,00 | 1 | 1200 | 5800 | 5 | 146,00 |
| 1 | 1400 | 6500 | 5 | 146,00 | 1 | 1100 | 6600 | 5 | 146,00 | 1 | 1100 | 6900 | 5 | 146,00 |
| 1 | 1200 | 7700 | 10 | 172,00 | 1 | 1700 | 8300 | 5 | 146,00 | 1 | 1900 | 8800 | 5 | 146,00 |
| 1 | 1800 | 9500 | 5 | 146,00 | 1 | 1300 | 9600 | 5 | 146,00 | 1 | 1100 | 9900 | 10 | 172,00 |
| 1 | 1700 | 11200 | 10 | 172,00 | 1 | 1400 | 11000 | 10 | 172,00 | 1 | 1300 | 11200 | 10 | 172,00 |
| 1 | 1200 | 12400 | 5 | 146,00 | 1 | 1300 | 12300 | 10 | 172,00 | 1 | 1000 | 12200 | 15 | 198,00 |
| 1 | 1000 | 13400 | 5 | 146,00 | 1 | 1200 | 13500 | 15 | 198,00 | 1 | 1300 | 13500 | 5 | 146,00 |
| 1 | 1000 | 14400 | 5 | 146,00 | 1 | 1200 | 14700 | 10 | 172,00 | 1 | 1200 | 14700 | 5 | 146,00 |
| 1 | 1200 | 15600 | 5 | 146,00 | 1 | 1400 | 16100 | 5 | 146,00 | 1 | 1200 | 15900 | 5 | 146,00 |
| 1 | 1800 | 17400 | 10 | 172,00 | 1 | 1000 | 17100 | 5 | 146,00 | 1 | 1000 | 16900 | 5 | 146,00 |
| 1 | 1100 | 18500 | 5 | 146,00 | 1 | 1200 | 18300 | 5 | 146,00 | 1 | 1500 | 18400 | 5 | 146,00 |
| 1 | 1100 | 19600 | 5 | 146,00 | 1 | 1200 | 19500 | 10 | 172,00 | 1 | 1200 | 19600 | 5 | 146,00 |
| 1 | 1300 | 20900 | 5 | 146,00 | 1 | 1800 | 21300 | 5 | 146,00 | 1 | 1300 | 20900 | 10 | 172,00 |
| 0 | 1200 | 22100 | 5 | 0,00 | 0 | 1900 | 23200 | 10 | 0,00 | 0 | 1100 | 22000 | 5 | 0,00 |
| 0 | 1200 | 23300 | 5 | 0,00 | 0 | 1200 | 24400 | 10 | 0,00 | 0 | 1400 | 23400 | 15 | 0,00 |
| 0 | 1900 | 25200 | 5 | 0,00 | 0 | 1200 | 25600 | 5 | 0,00 | 0 | 1000 | 24400 | 5 | 0,00 |
| 0 | 1000 | 26200 | 5 | 0,00 | 0 | 1600 | 27200 | 10 | 0,00 | 0 | 1100 | 25500 | 5 | 0,00 |
| | | | *Total* | *2.466,00* | | | | *Total* | *2.544,00* | | | | *Total* | *2.544,00* |

Para a análise da alternativa de troca conjunta de todos os rolamentos da máquina aproveitamos os valores gerados anteriormente e escolhemos o menor valor dentre os valores de mesma ordem e geramos novamente os valores referentes ao tempo de espera para troca dos rolamentos. Abaixo, apresentamos um resumo destes resultados:

| Todos Rolamentos | | | | |
|---|---|---|---|---|
| Ctr | Vida util | Tempo da Falha | Espera | Custo Troca |
| 1 | 1300 | 1300 | 5 | 282,00 |
| 1 | 1200 | 2500 | 5 | 282,00 |
| 1 | 1300 | 3800 | 5 | 282,00 |
| 1 | 1000 | 4800 | 10 | 308,00 |
| 1 | 1100 | 5900 | 10 | 308,00 |
| 1 | 1200 | 7100 | 5 | 282,00 |
| 1 | 1100 | 8200 | 5 | 282,00 |
| 1 | 1300 | 9500 | 5 | 282,00 |
| 1 | 1000 | 10500 | 5 | 282,00 |
| 1 | 1000 | 11500 | 5 | 282,00 |
| 1 | 1000 | 12500 | 15 | 334,00 |
| 1 | 1200 | 13700 | 5 | 282,00 |
| 1 | 1000 | 14700 | 5 | 282,00 |
| 1 | 1100 | 15800 | 5 | 282,00 |
| 1 | 1100 | 16900 | 15 | 334,00 |
| 1 | 1300 | 18200 | 5 | 282,00 |
| 1 | 1100 | 19300 | 5 | 282,00 |
| 1 | 1200 | 20500 | 5 | 282,00 |
| 0 | 1000 | 21500 | 10 | 0,00 |
| 0 | 1000 | 22500 | 15 | 0,00 |
| | | | *Total* | *5.232,00* |

Alertamos que a coluna 'Ctr' serve para controlar os valores que a serem considerados, já que truncamos a análise no primeiro tempo de falha maior que 20.000 horas.

- Comentários:

- Através desta simulação verificamos que a alternativa de troca conjunta de todos os rolamentos mostra-se mais vantajosa economicamente, importando um total de $5.232 contra

$7.554 referentes à alternativa de troca dos rolamentos à medida que suas falhas forem ocorrendo.

- Salienta-se que para efeito de uma análise mais acurada, deveríamos repetir esta simulação várias vezes, de maneira a obtermos estimativas da esperança, da variância e das probabilidades da distribuição da variável aleatória em estudo.

Aplicação 2:

*Uma loja tem somente 1 atendente. Os fregueses chegam aleatoriamente com intervalo, entre eles, variando de 1 a 8 minutos. Cada valor possível do intervalo entre chegadas tem a mesma probabilidade de ocorrência, como mostrado na tabela a seguir:*

| ***Tempo entre Chegadas (minutos)*** | ***Probabilidade*** | ***Tempo entre Chegadas (minutos)*** | ***Probabilidade*** |
|---|---|---|---|
| *1* | *0,125* | *5* | *0,125* |
| *2* | *0,125* | *6* | *0,125* |
| *3* | *0,125* | *7* | *0,125* |
| *4* | *0,125* | *8* | *0,125* |

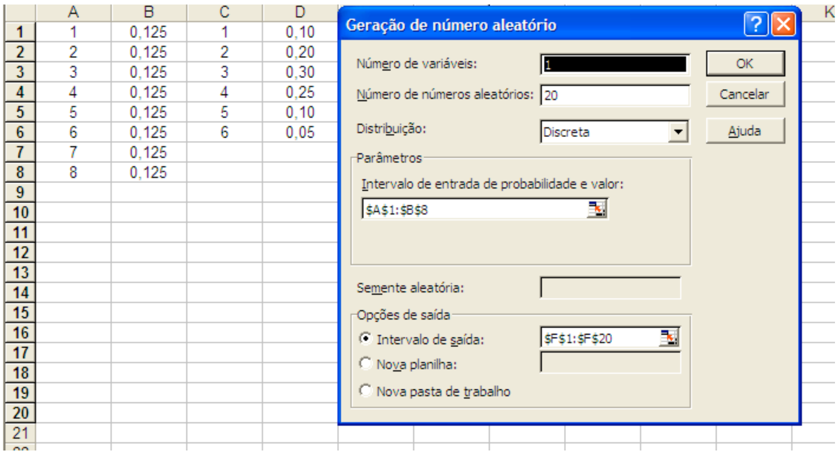
*A duração do atendimento aos clientes varia de 1 a 6 minutos, com probabilidades mostradas na tabela a seguir:*

| ***Duração do serviço (minutos)*** | ***Probabilidade*** | ***Duração do serviço (minutos)*** | ***Probabilidade*** |
|---|---|---|---|
| *1* | *0,10* | *4* | *0,25* |
| *2* | *0,20* | *5* | *0,10* |
| *3* | *0,30* | *6* | *0,05* |

*Simule o atendimento de 20 clientes e verifique: o tempo de espera na fila, o tempo total na loja, o tempo ocioso do atendente.*

- Geração dos números aleatórios para solução:

Escolha **Análise de dados** a partir do menu **Ferramentas** e selecione então **Geração de número aleatório**, em cuja caixa de diálogo preenchem-se os dados necessários à simulação, conforme ilustrado na figura abaixo. Faça o mesmo procedimento novamente, considerando agora os valores e suas correspondentes probabilidades, compreendidos no intervalo entre as células C1:D6, para obter-se a duração do serviço. O resultado será a simulação da operação desta loja durante o atendimento de 20 clientes.




– Resultados da Simulação:

Os resultados estão apresentados a seguir, sendo que os valores das colunas ‘Tempo entre chegadas’ e ‘Duração do serviço’ foram gerados aleatoriamente de acordo às distribuições dadas:

| Cliente # | Tempo entre chegadas (min) | Instante chegada | Atendimento | | Duração do serviço (min) | Tempo de Espera na Fila (min) | Tempo total na loja (min) | Tempo ocioso atendente (min) |
|---|---|---|---|---|---|---|---|---|
| | | | Início | Fim | | | | |
| 1 | 0 | 0 | 0 | 4 | 4 | 0 | 4 | 0 |
| 2 | 5 | 5 | 5 | 8 | 3 | 0 | 3 | 1 |
| 3 | 3 | 8 | 8 | 13 | 5 | 0 | 5 | 0 |
| 4 | 1 | 9 | 13 | 15 | 2 | 4 | 6 | 0 |
| 5 | 1 | 10 | 15 | 16 | 1 | 5 | 6 | 0 |
| 6 | 4 | 14 | 16 | 19 | 3 | 2 | 5 | 0 |
| 7 | 1 | 15 | 19 | 21 | 2 | 4 | 6 | 0 |
| 8 | 3 | 18 | 21 | 24 | 3 | 3 | 6 | 0 |
| 9 | 6 | 24 | 24 | 26 | 2 | 0 | 2 | 0 |
| 10 | 7 | 31 | 31 | 37 | 6 | 0 | 6 | 5 |
| 11 | 7 | 38 | 38 | 42 | 4 | 0 | 4 | 1 |
| 12 | 2 | 40 | 42 | 46 | 4 | 2 | 6 | 0 |
| 13 | 4 | 44 | 46 | 51 | 5 | 2 | 7 | 0 |
| 14 | 8 | 52 | 52 | 56 | 4 | 0 | 4 | 1 |
| 15 | 8 | 60 | 60 | 63 | 3 | 0 | 3 | 4 |
| 16 | 6 | 66 | 66 | 70 | 4 | 0 | 4 | 3 |
| 17 | 8 | 74 | 74 | 78 | 4 | 0 | 4 | 4 |
| 18 | 5 | 79 | 79 | 81 | 2 | 0 | 2 | 1 |
| 19 | 1 | 80 | 81 | 85 | 4 | 1 | 5 | 0 |
| 20 | 4 | 84 | 85 | 87 | 2 | 1 | 3 | 0 |
| | | | | | *67* | *24* | *91* | *20* |

– Comentários:

- Podemos, a partir da simulação estimar alguns parâmetros do processo em questão:
- O tempo médio de espera de um cliente para ser atendido foi: 1,2 min., onde:

$$\text{Tempo de espera médio} = \frac{\text{Tempo total de espera na fila}}{\text{Quantidade total de clientes}}$$

A proporção de tempo que o atendente fica ocioso foi: 0,23

$$\text{Probabilidade ociosidade} = \frac{\text{Tempo ocioso total do atendente}}{\text{Duração da simulação}}$$

O tempo médio de serviço foi: 3,35

$$\text{Tempo médio de serviço} = \frac{\text{Duração total do serviço}}{\text{Quantidade total de clientes}}$$

A probabilidade de um cliente esperar na fila foi: 0,45

$$\text{Probabilidade de espera} = \frac{\text{Qte. total de clientes esperando na fila}}{\text{Quantidade total de clientes}}$$

Verifica-se assim que, embora a maioria dos clientes tenha de esperar o atendimento, a espera não se mostra demasiada e que o atendente não fica muito tempo ocioso. A decisão a ser tomada deverá considerar um balanço entre o custo de se colocar mais atendentes e o custo de espera na fila.

Um tomador de decisão estaria interessado nos resultados obtidos acima, salientando-se que quanto maior a duração da simulação, maior a proximidade com o tempo médio esperado.

## 2.11 CONCLUSÃO:

No menu ‘Análise de Dados’ encontra-se também a ferramenta ‘Amostragem’ que retira amostras de uma população de valores. Há ainda a possibilidade de se utilizar a função RANDÔMICO, que retornará um número aleatório sempre que a planilha for calculada, permitindo sua integração à planilhas e possibilitando sua atuação sempre que a planilha for calculada. Um exemplo de utilização seria a montagem, em uma planilha de cálculo do Excel, do método de geração de variáveis aleatórias exponenciais desenvolvido no capítulo anterior, permitindo sua atuação tantas vezes quanto a planilha for calculada.

O Excel oferece assim amplas possibilidades em modelagens probabilísticas mais simples ou em situações que não exigem um grande esforço computacional. Percebe-se, entretanto que sua flexibilidade e facilidade permitem sua utilização em modelagens mais simples, mas nem por isso menos importantes, podendo ser usado como um facilitador para a disseminação do uso de simulação em processos decisórios ou para análise preliminar.

## 2.12 EXERCÍCIOS PROPOSTOS:

**1.** As probabilidades de um corretor de imóveis vender 0, 1, 2, 3, 4, 5 ou 6 casas em uma semana são:

0,14 0,27 0,27 0,18 0,09 0,04 0,01

Simule suas vendas durante 25 semanas consecutivas.

**2.** O proprietário de uma padaria sabe que as probabilidades de procura por 0, 1, 2, 3, 4, ou 5 unidades de um bolo de queijo altamente perecíveis são:

0,05 0,15 0,25 0,25 0,20 0,10

**a)** Simule a procura pelo bolo de queijo em 30 dias úteis consecutivos;

**b)** Se o padeiro tem um lucro de $2,00 em cada bolo de queijo que vende, mas perde $1,00 em cada bolo perdido (isto é, que não é vendido no dia da fabricação), determine o lucro ou o prejuízo do padeiro em cada um dos 30 dias referidos no item anterior, supondo que produza três bolos de queijo diariamente. Determine o lucro médio do padeiro nos 30 dias.

c) Refaça o item anterior, supondo que o padeiro faça diariamente quatro bolos de queijo em lugar de três. Não repita o item a). Qual dos resultados parece ser mais lucrativo?

**3.** Um time de futebol irá disputar 10 partidas em um torneio de classificação.

**a)** Supondo que sua chance de vitória em cada jogo é de 60%, simule sua possível campanha.

**b)** Simule agora se é esperado o seguinte desempenho em cada jogo: 50% de vitória, 30% de empate e 20% de derrota;

**c)** Para a situação descrita em (b), simule 12 possíveis campanhas para o time, e estude a variável $X$= *número de pontos obtidos* (vitória = 3, empate = 1 e derrota = 0);

**d)** Proponha outros parâmetros para o time e repita a questão (c).

**4.** Utilizando um gerador de números aleatórios, obtenha uma amostra de 100 observações de uma distribuição Binomial, com parâmetros $n$= 10, 30, 50, 100 e $p$ = 0,5; 0,7; e 0,9. Para cada combinação de $n$ e $p$, construa um histograma. Analise o comportamento dos histogramas com relação aos valores de $n$ e $p$. Baseando-se na forma dos histogramas construídos, o que pode ser dito a respeito da aproximação para a distribuição normal?

**5.** Com o objetivo de simular a distribuição amostral de $\overline{X}$ realize as seguintes tarefas:

a) Gere 100 amostras de tamanho 30 de uma Normal com média 200 e desvio-padrão 5. Calcule então a média de cada amostra e faça o histograma correspondente à esse conjunto de médias amostrais. Qual conclusão pode ser tirada? Calcule medidas descritivas das 100 médias obtidas e comente.

**e)** Repita o item a) para as amostras geradas de uma binomial com parâmetros $n = 50$ e $p = 0{,}45$. Comente a respeito do histograma e das medidas descritivas obtidas, tendo em vista o Teorema Central do Limite.

**6.** Uma pesquisa domiciliar irá entrevistar todos os moradores do domicílio e a distribuição do número de moradores por domicílio encontra-se abaixo. Será usada uma amostra de 5 domicílios:

a) Simule 100 amostras de tamanho 5.
b) Considere **X**: número médio de pessoas pro amostra. Qual a distribuição de freqüência empírica de **X**?
c) Construa a distribuição de **X**: número médio de pessoas por amostra.
d) Encontre para a população o valor $\mu$ = número médio de pessoas, e construa a distribuição empírica de $\mathbf{X} - \mu$. Como pode ser interpretada esta distribuição?
e) Se o entrevistador recebe \$2 por pessoa entrevistada, usando o resultado b), qual a probabilidade de uma amostra custar mais de \$12?

**7.** A altura **X** das pessoas segue aproximadamente uma curva normal com média $\mu$ e variância $\sigma^2$.

a) Proponha dois valores realísticos para $\mu$ e $\sigma$, e gere 10 alturas de uma população de homens. Calcule a média e o desvio-padrão desta população.
b) Com os mesmos parâmetros, gere uma outra amostra de 10 alturas. Olhando e analisando as duas amostras elas parecem vir de populações distintas?
c) Gere uma amostra de 10 alturas de uma população feminina. Compare com a amostra obtida em a), e diga se é possível afirmar que as duas amostras vêm de populações distintas;
d) Como você acha que os parâmetros influenciam para diferenciar bem as amostras? Dê exemplos.

**8.** Suponha que o número de horas que uma pessoa leva para aprender a operar uma certa máquina é uma variável aleatória com uma distribuição normal com média 5,8 e desvio-padrão 1,2. Suponha que é necessário que duas pessoas aprendam a operar uma máquina. Simule o tempo que 4 pares de pessoas levam para aprender como operar a máquina, ou seja, para cada par, calcule o máximo de dois tempos de aprendizagem.

**9.** Repita a simulação do problema anterior 100 vezes e faça um histograma de:

a) 400 tempos de aprendizagem para os pares de operadores;
b) os 100 valores representando o tempo para treinar quatro pares de operadores.

**10.** Em uma população 20% das pessoas compram o produto **C**. Seleciona-se, com reposição indivíduos dessa população, até encontrar um comprador de C. A Variável **X** indica o número de indivíduos entrevistados. Qual a distribuição simulada de **X**?

## REFERÊNCIAS 2


1. BUSSAB, W.O. e MORETTIN, P.A. *Estatística básica*. São Paulo: Saraiva, 2002.

2. SANTOS, M.P. *Introdução à simulação discreta*. Rio de Janeiro: UERJ, 1999.

3. NEUFELD, J.L. *Estatística aplicada à administração usando Excel.* São Paulo: Prentice Hall, 2003.

4. MILLER, I. & FREUND, J.E. *Probability and statistics for engineers.* Upper Saddle River: Prentice Hall, 2000.