Anexo VI “Técnicas Básicas de Simulação” do livro

“Apoio à Decisão em Manutenção na Gestão de Activos Físicos”

LIDEL, 2010

Rui Assis
rassis@rassis.com
http://www.rassis.com

ANEXO VI – Técnicas Básicas de Simulação

Simular é imitar a realidade

A simulação é uma forma de imitar a realidade sem correr os riscos, os custos e o tempo que resultariam se tivéssemos de experimentar. Para o fazermos, criamos um modelo matemático que descreva o comportamento ao longo do tempo do sistema que vamos estudar.

Conceito de sistema

Um modelo considera os elementos existentes no mundo real inter-relacionados e funcionando com objectivos. O construtor do modelo deve focar a sua atenção apenas nos elementos mais importantes e na natureza das suas inter-relações, com o objectivo último de conseguir melhorar o desenho e funcionamento do sistema representado pelo modelo.

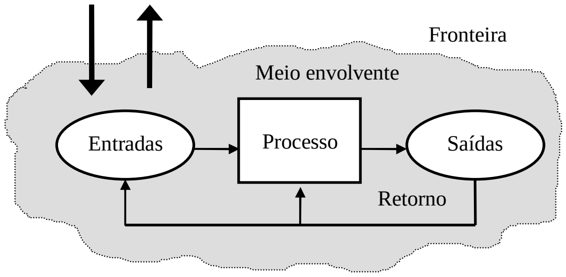


Figura VI.1 – Modelo de um sistema

Tipos de modelos

Existem muitos tipos de modelos. Uns são <u>físicos</u> outros são <u>simbólicos</u>. Os modelos físicos compreendem os modelos <u>icónicos</u> e os <u>analógicos</u>. Os modelos simbólicos compreendem os modelos <u>verbais</u> e <u>matemáticos</u>.

Modelos matemáticos

Os modelos matemáticos – os únicos que nos interessam neste curso – possuem várias características. Podem diferir nos <u>objectivos</u> (descrição *versus* optimização), <u>modo de análise</u> (analítico *versus* numérico) e <u>aleatoriedade</u> (determinístico *versus* probabilístico).

Relações entre variáveis

Os modelos de simulação contêm equações que expressam relações entre variáveis de interesse. Por exemplo: $CT = CF + c_v.Q$

Tipos de variáveis

As variáveis de um modelo de simulação podem classificar-se como de <u>entrada</u> (políticas, aleatórias ou determinísticas), <u>saída</u> e <u>saída intermédia</u>, (ver a Figura VI.2).

Modelos discretos e contínuos

Dependendo da natureza do sistema e das características de interesse pode usar-se um modelo de simulação <u>discreto</u>, <u>contínuo</u> ou <u>combinado</u>. Os modelos discretos são os mais populares.

Vantagens e desvantagens

A simulação é um método muito potente de análise que, devido ao tempo e custo necessários, só deve ser usado depois de esgotadas as alternativas:

- Modelos matemáticos;
- Experiências com o sistema real ou com um seu protótipo;
- Experiência e intuição pessoais.

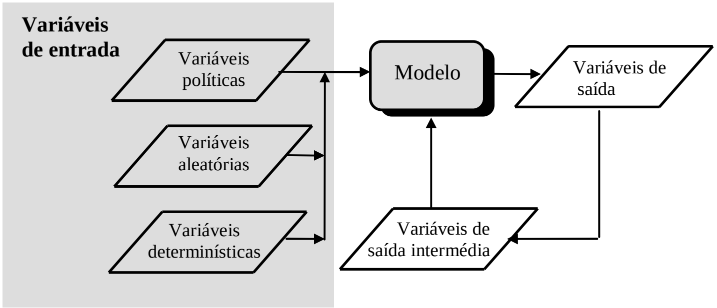


Figura VI.2 – Variáveis de um modelo

O ciclo de vida de um modelo de simulação compreende três fases.

Ciclo de vida de um modelo

- Na fase de <u>definição</u>, o problema é transmitido ao analista que o conceptualiza e constrói uma solução técnica;
- Na fase de <u>desenvolvimento</u>, constrói-se o modelo, valida-se e coloca-se em utilização;
- Na fase de <u>apoio à decisão</u>, o gestor manipula o modelo e decide colocando questões do tipo "O que é que acontece se ...?" (*what if ...?*).

Muitas vezes os modelos de simulação são componentes de sistemas de informação para gestão e de sistemas de apoio à decisão, compartilhando as mesmas bases de dados.

Simulação de *Monte-Carlo*

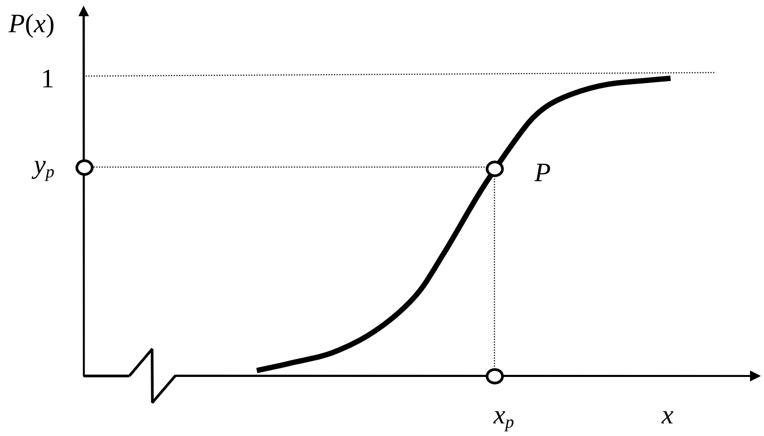
Um modelo de simulação pode proporcionar uma representação bastante fiel do mundo real. Esta representação não é possível, grande parte das vezes, com modelos analíticos. A simulação de processos probabilísticos pode ser realizada por vários métodos, sendo o de *Monte-Carlo* o mais popular. A sua adequabilidade, porém, depende dos casos concretos.

VI.1 – Método de Simulação de *Monte-Carlo*

O método de simulação de *Monte-Carlo* desenvolve-se ao longo das seguintes fases:

<u>Fase 1</u> **–** Definimos a função de probabilidade acumulada $P(x)$, da variável aleatória x, a qual pode ser uma distribuição teórica (Uniforme, Triangular, Normal, Beta, *Weibull*, etc.) ou uma distribuição empírica qualquer. A Figura VI.3 representa uma função de probabilidade acumulada $P(x)$ da variável aleatória contínua x;
<u>Fase 2</u> **–** Escolhemos um número aleatório equiprovável entre 0 e 1 numa tabela de números aleatórios (ou usando a função RAND() no EXCEL). Representamos este número y_p no eixo das ordenadas da função $P(x)$;
<u>Fase 3</u> **–** Projectamos y_p horizontalmente até à curva $P(x)$, definindo-se o ponto P. Projectamos este ponto, por sua vez, sobre o eixo das abcissas, definindo-se o valor x_p de uma amostra;
<u>Fase 4</u> **–** Repetimos o procedimento e obtemos uma amostra.

Figura VI.3 – Função de probabilidade acumulada P(x)



Vejamos um exemplo de geração de uma função de duas variáveis.

Exemplo VI.1

A variável dependente é função de duas variáveis independentes aleatórias x e y com a forma $Z = 5x + 2y$. As variáveis aleatórias x e y podem tomar os valores dentro dos intervalos de referência (probabilidades) descritas no Quadro VI.1. Gerar 8 resultados de Z.

Quadro VI.1 – Intervalos empíricos de probabilidade assumidas pelas variáveis x e y

Valores de *x*	Intervalos de referência	Valores de *y*	Intervalos de referência
0	00 - 09	2	00 - 24
1	10 - 29	3	25 - 54
2	30 - 54	4	55 - 79
3	55 - 79	5	80 - 99
4	80 - 94		
5	95 - 99		

Quadro VI.2 – Valores aleatórios assumidos pelas variáveis independentes x e y e dependente Z

Ensaios	Nº aleatório para *x*	*x*	Nº aleatório para *y*	*y*	Valores de *Z = 5x + 2y*
1	43	2	22	2	14
2	96	5	50	3	31
3	57	3	13	2	19
4	53	2	36	3	16
5	14	1	91	5	15
6	03	0	58	4	8
7	33	2	45	3	16
8	40	2	43	3	16

Caso em Gestão de *stocks*

Um modelo de simulação de *Monte-Carlo* pode prever o comportamento de um sistema de gestão do *stock* de um artigo[1], fornecendo como variáveis de saída: o nível médio e nível máximo de *stock* e os custos de posse, de encomendas e de roturas. Estes valores dependem do conjunto de variáveis de decisão: Ponto de Encomenda e quantidade por encomenda

[1] Ver o Anexo VII.

ou periodicidade de encomenda e nível de *stock* objectivo – conforme usarmos o modelo de revisão contínua ou de revisão periódica, respectivamente.

Caso de uma fila de espera

Um modelo de simulação de uma fila de espera permite obter como variáveis de saída: os tempos de espera dos clientes e o comprimento da fila, bem como, os tempos de ociosidade dos atendedores. Estes valores dependem do tempo entre chegadas e dos tempos de atendimento (variáveis independentes) e do número de atendedores (variável de decisão).

Optimização em simulação

A optimização por simulação consegue-se por aproximações sucessivas à zona óptima, fazendo variar o conjunto de todas as variáveis de decisão.

A figura VI.4 mostra o caso de sucessivas iterações para encontrar o conjunto de valores das variáveis de entrada *X* e *Y* que minimizam a variável de saída *Z*. O ponto *P* corresponde ao óptimo, ou seja, ao mínimo dos mínimos de *Z*.

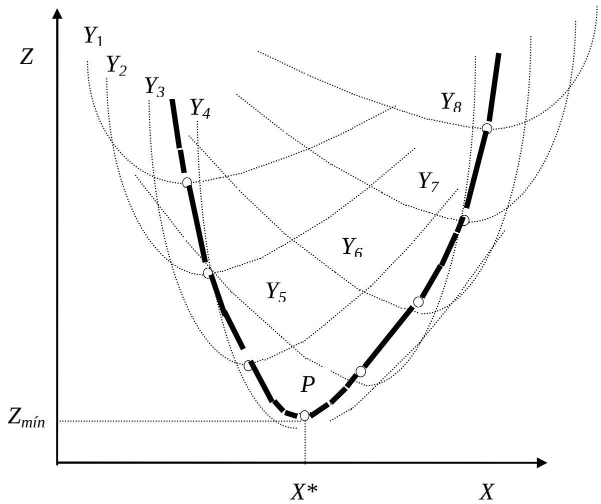


Figura VI.4 – Mínimo dos mínimos das várias funções Z(X,Y)

VI.2 Processos geradores de valores aleatórios

Distribuições de Probabilidade teóricas

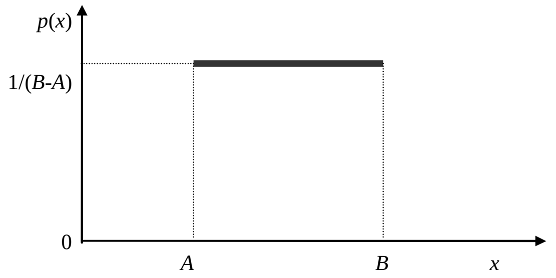
Uma componente muito importante de um modelo de simulação consiste na amostragem dos processos probabilísticos. O procedimento que permite executar uma amostragem designa-se processo gerador. O processo gerador de distribuições de probabilidade empíricas é mais difícil do que o das distribuições de probabilidade teóricas. O processo gerador das distribuições teóricas requer apenas que o analista introduza os parâmetros apropriados a cada tipo de distribuição.

VI.2.1 Variáveis contínuas

No caso das distribuições contínuas teóricas, tais como a Uniforme, a Exponencial negativa e a Triangular, o processo gerador pode ser obtido pelo método inverso. Com este método, a distribuição simples é integrada para obter a sua versão acumulada. A equação resultante é igualada a r (número aleatório uniforme) e resolvida de forma a obtermos a variável aleatória x em função de r.

Distribuição Uniforme

Figura VI.5 – Função densidade de probabilidade da distribuição Uniforme



O gerador da distribuição Uniforme é:

Expressão VI.1

$$x = A + r.(B - A)$$

Apoio do EXCEL

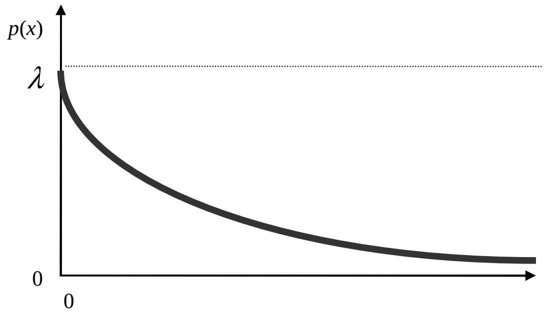
No EXCEL, podemos gerar qualquer valor contínuo x da distribuição Uniforme, entre B e A ($B > A$), fazendo:

Expressão VI.2

$$x = A + \text{RAND()}*(B - A)$$

Distribuição Exponencial negativa

Figura VI.6 – Função densidade de probabilidade da distribuição Exponencial



O gerador da distribuição Exponencial no EXCEL é:

Expressão VI.3

$$x = \text{-LN(RAND())} / \lambda$$

Distribuição Triangular

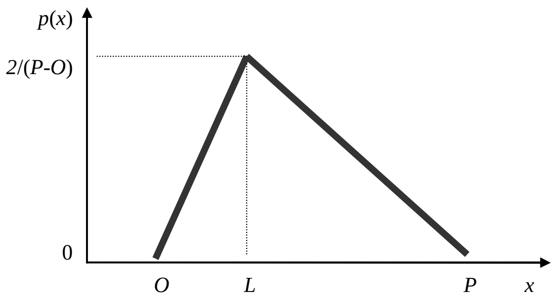


Figura VI.7 – Função densidade de probabilidade da distribuição Triangular

O gerador da distribuição Triangular, quando $r < (L – O) / (P – O)$, é:

$$x = O + \sqrt{r.(L-O).(P-O)}$$

Expressão VI.4

O gerador da distribuição Triangular, quando: $r \geq (L – O) / (P – O)$, é:

$$x = P - \sqrt{(1-r).(P-L).(P-O)}$$

Expressão VI.5

Distribuição Normal

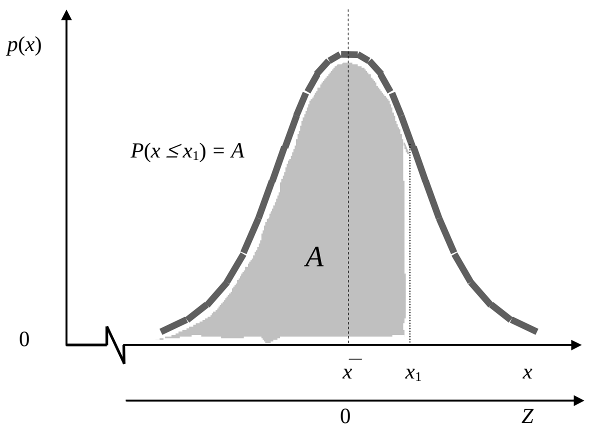


Figura VI.8 – Função densidade de probabilidade da distribuição Normal

O gerador da distribuição Normal no EXCEL é:

$$x = \text{NORMINV(RAND();}\mu\text{;}\sigma\text{)}$$

Expressão VI.6

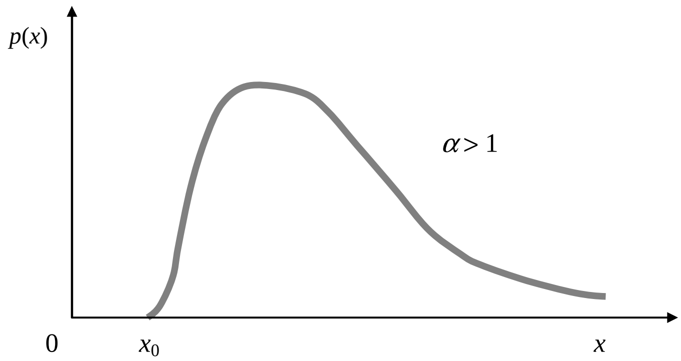
ou:

$$x = \mu + \text{NORMSINV(RAND())}*\sigma$$

Expressão VI.7

Distribuição de *Weibull*

Figura VI.9 – Função densidade de probabilidade da distribuição de Weibull



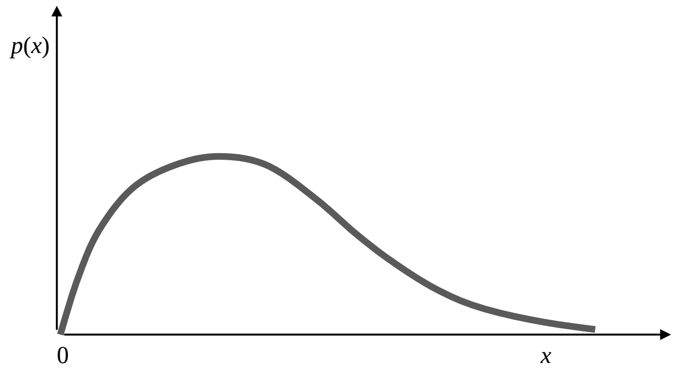
A função geradora da função de *Weibull* no EXCEL é a seguinte:

Expressão VI.8

$$x = x_0 + \beta.\left[-\ln\left(\text{RAND}()\right)\right]^{1/\alpha}$$

Distribuição LogNormal

Figura VI.10 – Função densidade de probabilidade da distribuição LogNormal



A função geradora da função LogNormal no EXCEL é a seguinte:

Expressão VI.9

$$x = \text{LOGINV}(\text{RAND}();\mu_y;\sigma_y)$$

VI.2.2 Variáveis discretas

A amostragem a partir de distribuições teóricas discretas, tais como a *Bernoulli*, Binomial (*n* processos de *Bernoulli*) e *Poisson*, é realizada através de um procedimento de contagem dos valores assumidos pela variável aleatória.

No caso de variáveis discretas, se existirem dados históricos cujo padrão possa repetir-se no futuro, analisam-se os dados em frequência. Se não existirem esses dados, estimam-se em probabilidade. Calculam-se, assim, as distribuições em probabilidade simples e acumulada de cada variável. Seguidamente, aloca-se a cada valor possível da variável, um intervalo de números aleatórios proporcional à probabilidade da sua ocorrência. Produz-se então uma amostra de valores da variável, gerando números aleatórios e seleccionando os valores da variável que lhe estão associados.

Nesta obra não se apresentam os processos geradores das distribuições teóricas discretas pois não é fácil programá-los no EXCEL. É, contudo, possível criar valores aleatórios das distribuições *Bernoulli*, Binomial, *Poisson* e *Discrete* (a Uniforme e a Normal são contínuas e a *Patterned* pode ser contínua ou discreta), usando a janela de diálogo mostrada na Figura VI.11, fazendo sucessivamente: *Tools*, *Data analysis*, *Random Number Generation*.

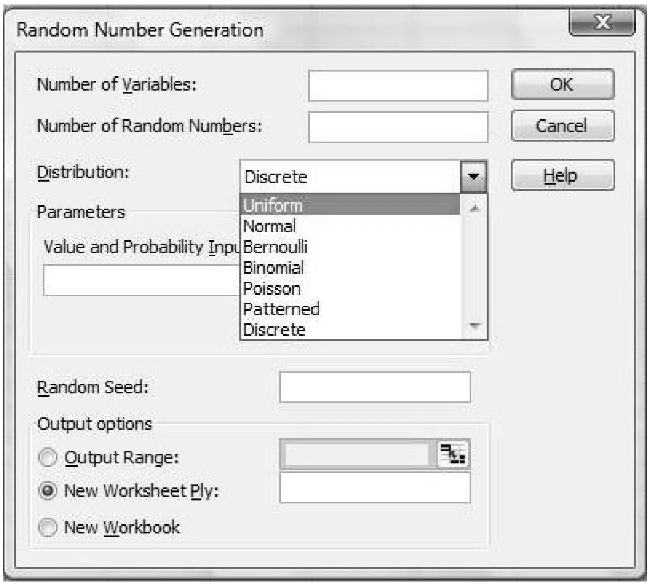


Figura VI.11 – Janela de diálogo no EXCEL para geração de números aleatórios

VI.3 Complementos das técnicas de simulação

A qualidade dos números aleatórios é fundamental para um estudo de simulação. Os números aleatórios podem ser obtidos numa tabela (das muitas publicadas) ou em *software* para o efeito. Por vezes, os números aleatórios não se encontram disponíveis e o analista tem que saber gerá-los e testar a sua qualidade.

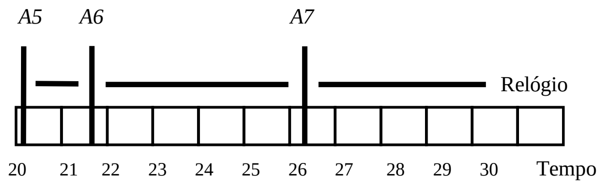
Um gerador de números aleatórios deve gerar números verdadeiramente aleatórios (equiprováveis), ser rápido, não requerer muito espaço de armazenagem, possuir um longo período de ciclo, não degenerar e ser capaz de repetir sequências predefinidas. Um gerador de números aleatórios deve passar qualquer teste estatístico de aleatoriedade.

O objectivo da fase de desenho experimental de um modelo de simulação, consiste em conseguir o desenho efectivo e eficiente de uma experiência, de forma a determinar o efeito de diferentes variáveis no comportamento do sistema que o modelo representa. Trata-se pois de uma forma de conhecer a resposta do sistema a diferentes factores. O número destes factores pode ser apenas um ou vários. Cada factor pode variar ao longo de todo o intervalo possível ou apenas ao longo de parte, dizendo-se que o desenho experimental é completo ou parcial, respectivamente.

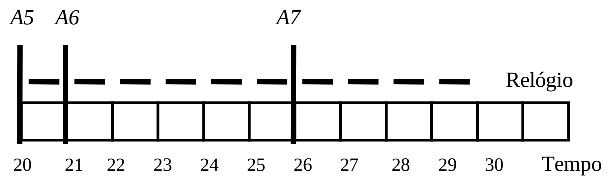
Os modelos de simulação empregam os seguintes métodos de avanço do tempo: próximo acontecimento (o relógio progride por saltos desiguais); incremento fixo (o relógio progride por saltos iguais). A selecção de um ou outro método depende do tipo de aplicação, da linguagem de programação usada, do tempo disponível para correr o modelo e da precisão da informação requerida.

<u>Próximo acontecimento:</u>

Figura VI.12 – Métodos de avanço do tempo: próximo acontecimento ou por incremento fixo

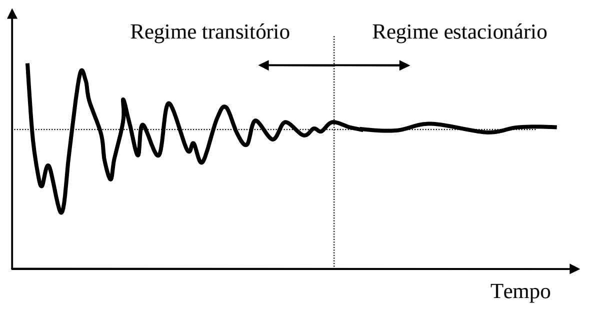


<u>Incremento temporal fixo:</u>



Quando se inicia uma corrida de um modelo de simulação, existe um período inicial <u>transitório</u> antes de se entrar no regime <u>estacionário</u>. O interesse da análise pode situar-se no período transitório ou no estacionário ou em ambos, dependendo das características de interesse do sistema que se pretendem captar. Este facto determina quais as condições do estado inicial que deverão ser usadas e o período de tempo a simular de forma a poder coligir e analisar dados suficientes.

Figura VI.13 – Regime transitório e regime estacionário



Quando as observações das variáveis de saída são estatisticamente independentes (ou quase), pode utilizar-se as fórmulas da amostragem simples para calcular o número necessário de iterações.

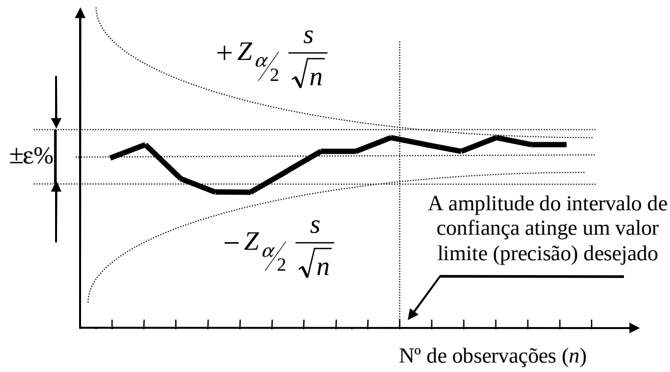


Figura VI.14 – A amplitude do intervalo de confiança decresce com o aumento do nº de observações n

A validação de um modelo ao longo das várias fases de desenvolvimento é muito importante. A validação compreende:

- O teste dos pressupostos;
- O teste empírico das relações usadas;
- A comparação das saídas do modelo com os dados colhidos do sistema real representado.

Devido à natureza probabilística das variáveis de saída, os seus valores médios devem ser expressos em intervalos de confiança conforme mostra a Figura VI.15 e é descrito nas Expressões VI.10, VI.11 e VI.12.

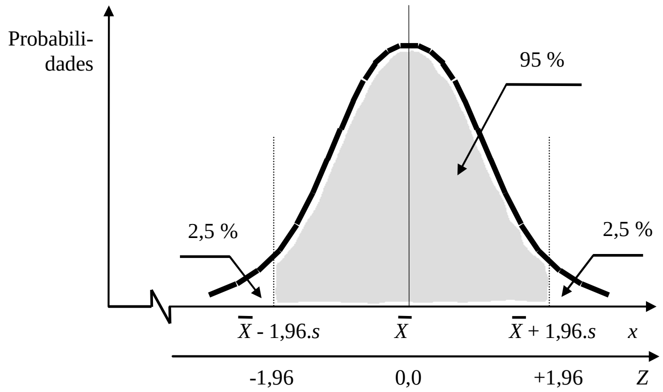


Figura VI.15 – Normal reduzida para um nível de confiança de 95%

Intervalo de confiança da média quando $n \leq 30$ (usamos a função t de *Student*):

$$\bar{x} \pm t_{\alpha/2;(n-1)} \cdot \frac{s}{\sqrt{n}}$$

Expressão VI.10

Intervalo de confiança da média quando $n > 30$ (usamos a função Normal reduzida Z):

Expressão VI.11

$$\bar{x} \pm Z_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

Intervalo de confiança no caso das proporções:

Expressão VI.12

$$\bar{p} \pm Z_{\alpha/2} \cdot \sqrt{\frac{\bar{p}.(1-\bar{p})}{n}}$$

Nestas três expressões, α representa o nível de significância e o segundo termo da equação é o erro amostral.

Quando se comparam duas alternativas entre si usando o mesmo modelo de simulação, e no qual cada alternativa se caracteriza por diferentes valores de uma ou mais das suas variáveis de entrada, torna-se necessário determinar se os resultados obtidos são significativamente diferentes ou não. Para tal, se μ_A e μ_B representarem os valores esperados de uma qualquer variável de saída resultantes da simulação das duas alternativas, e se considerarmos que a hipótese nula corresponde a $\mu_A = \mu_B$, então, podemos chegar a uma de três conclusões:

- $\mu_A = \mu_B$ (a hipótese nula é verdadeira);
- $\mu_A < \mu_B$ (a hipótese nula é falsa);
- $\mu_A > \mu_B$ (a hipótese nula é falsa).

A conclusão dependerá dos valores esperados $\overline{X}_A$ e $\overline{X}_B$ da variável em estudo, bem como dos desvios padrão s_A e s_B obtidos nas duas alternativas. Assumindo que as simulações são estatisticamente independentes, isto é, que os números aleatórios usados em cada alternativa foram diferentes, e assumindo normalidade na distribuição dos resultados de cada alternativa, podemos eleger Z para estatística de teste.

Expressão VI.13

$$Z = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

Admitindo um nível de significância α, a regra de decisão será então a seguinte:

- Se $Z < -Z_{\alpha/2}$ então: $\mu_A < \mu_B$ e rejeitamos a hipótese nula;
- Se $-Z_{\alpha/2} < Z < Z_{\alpha/2}$ então: $\mu_A = \mu_B$ e aceitamos a hipótese nula;
- Se $Z > Z_{\alpha/2}$ então: $\mu_A > \mu_B$ e rejeitamos a hipótese nula.

O desenvolvimento bem-sucedido de um modelo de simulação requer mais do que competência técnica. O analista deve seguir um processo lógico e sistemático de desenvolvimento do modelo e deve saber relacionar-se com os futuros utilizadores de forma a gerar interesse e aceitação.

As fases do processo de concepção, teste e implementação de um modelo incluem:

Concepção de modelos de simulação

- A identificação do problema e definição de objectivos;
- A ponderação de potenciais custos e benefícios;
- A colheita de dados e desenvolvimento do modelo;
- A validação do modelo;
- A implementação dos resultados.

Durante o desenvolvimento de um modelo é fundamental que haja envolvimento do utilizador final e haja apoio por parte dos gestores.

Deve haver muito cuidado na definição do problema e na selecção da informação necessária. O desenho do modelo deve ter em conta a verdade dos dados, a forma como os utilizadores tomam decisões e a forma como o modelo irá ser utilizado. Uma boa regra consiste em construir um modelo pequeno e expandi-lo mais tarde quando a situação o exigir.

Um modelo só deverá ser usado se o utilizador final perceber a sua validade e utilidade.