SIMULAÇÃO

I N T R O D U Ç Ã O   À   S I M U L A Ç Ã O

Uma rápida consulta ao "Dicionário da Língua Portuguesa" da Porto Editora (7ª ed.) indica-nos: "**simulação** - acto ou efeito de simular; fingimento; disfarce; (...) ; **representação de um sistema ou de um processo por um modelo estatístico com que se trabalha, como se tratasse desse sistema ou processo, para investigar os seus efeitos**." Como se vê, é muito útil ter um dicionário por perto...

O vocábulo "simulação" não é, obviamente, desconhecido da generalidade das pessoas.

Quem não ouviu falar dos **simuladores de voo** ? Para nossa tranquilidade (e porque os aviões são 'bichos muito caros'), os pilotos de aeronaves, para além de outro tipo de formação, recebem muito treino em simuladores de voo, onde se vêem confrontados com um grande número de 'ocorrências', que, se Deus quiser, nunca terão que viver na realidade.

Quem não se impressionou ao passar perto de uma **barragem** ? Uma estrutura (geralmente) de betão a reter um verdadeiro 'mar', permitindo a produção de energia eléctrica e a regularização de situações de cheia e de falta de água. Uma visita ao Laboratório Nacional de Engenharia Civil permite-nos ver uma série de modelos reduzidos de barragens com vista à **simulação de condições 'especiais'**, como por exemplo **sismos**. Uma série de pequenos sensores permite estudar as deformações e deslocamentos sofridos pelo modelo quando submetido a uma 'solicitação especial', permitindo antever o comportamento da estrutura real ... As **grandes pontes** são também objecto de estudos de simulação de comportamento a partir de **modelos reduzidos**.

A concepção de **aeronaves** passa sempre pela **simulação da sua resistência aerodinâmica** com modelos reduzidos em 'túneis de vento'.

A **gestão do tráfego rodoviário** numa grande cidade passa actualmente por sistemas que controlam automaticamente os semáforos, de modo a aumentar a fluidez do trânsito. Em Lisboa, o sistema Gertrudes, gere os semáforos das zonas mais movimentadas. Mas como foi 'programado / calibrado' esse sistema ? Certamente que não se fizeram experiências no **'sistema real'**, isto é, ninguém se atreveu a experimentar diferentes políticas de tempos de verde e vermelho em directo na Praça Marquês de Pombal... É que essa experiência poderia ser fatal ... para o experimentador ... Para evitar dramas 'reais', concebe-se um **modelo** que descreve 'o melhor possível' o **sistema**, implementa-se esse modelo em termos informáticos e, finalmente, testa-se diferentes **'políticas'** sobre o modelo na segurança de um gabinete ... **A resposta do modelo, que simula a realidade submetida às políticas testadas, permite antecipar as respostas que ocorreriam na realidade** (... e sem pôr em risco a vida do experimentador). Em função dos resultados das simulações, pode-se adoptar uma 'política' de gestão que se considere adequada.

Poderemos agora olhar novamente para a definição que o Dicionário apresentava de **Simulação**: **representação de um sistema ou de um processo por um modelo estatístico com que se trabalha, como se tratasse desse sistema ou processo, para investigar os seus efeitos**.

Podemos considerar que um **sistema** é, *grosso modo*, um conjunto de **entidades** que interagem entre si, com vista a um determinado fim comum - no exemplo acabado de referir (gestão do tráfego rodoviário), o sistema seria constituído por diferentes entidades (veículos (e seus condutores), peões, vias de comunicação e semáforos) que interagem entre si (peões e condutores pretendem atingir os seus destinos, devendo conformar-se com as vias de comunicação existentes e com as indicações dadas pelos semáforos).

Se estivéssemos interessados em determinar a temperatura, a pressão e a concentração química dos intervenientes numa determinada reacção química, poderíamos **simular o** correspondente **processo** químico.

Já referimos a existência de **modelos físicos** de um sistema, a uma escala normalmente reduzida (o exemplo das barragens) e **modelos matemáticos** (a representação do funcionamento dos semáforos numa dada zona pode ser feita com um modelo físico, pouco útil para testar diferentes políticas de gestão, ou com um modelo matemático, que pode ser implementado computacionalmente, permitindo testar diferentes políticas de gestão). Nos modelos matemáticos teremos sempre uma base constituída por **relações de lógica** (do tipo "se está o semáforo 'está verde', então passam os automóveis e param os peões"), sendo usual ter-se uma importante **componente de ordem probabilística**, dado que muitos dos intervenientes no funcionamento de um dado sistema não são determinísticos (por exemplo, os instantes de chegada de automóveis a um semáforo são obviamente aleatórios...).

A **Simulação** que apresentaremos terá por base modelos matemáticos (envolvendo as relações de lógica e as componentes probabilísticas) para descrição de sistemas ou processos.

Comecemos por considerar o seguinte **problema:**

A **Barragem de Pós-Boa** é relativamente pequena, com uma capacidade máxima de retenção, **VMAX**, de 100 u.vol. de água.

A equação simplificada de "balanço hidrológico mensal" é

$$\mathbf{V_f = V_0 + P - T - N - DS},$$

sendo

V_f — Volume no final do mês (u.vol.)
V_0 — Volume no início do mês (u.vol.)
P — Volume de precipitação mensal (u.vol.)
T — Volume turbinado durante o mês (u.vol.)
N — Volume para satisfação de necessidades de abastecimento mensal (u.vol.)
DS — Volume de descarga de superfície durante o mês (u.vol.)

Por razões de equilíbrio ecológico e ambiental nunca se pode permitir que o nível da barragem desça abaixo de dez por cento da capacidade máxima de retenção.

As necessidades de água a jusante, **Nec** (u.vol.), têm-se mantido relativamente inalteradas nos últimos dez anos, sendo dadas na tabela seguinte:

Mês	J	F	M	A	M	J	J	A	S	O	N	D
Nec (u.vol.)	8	8	8	9	10	13	15	15	15	13	9	8

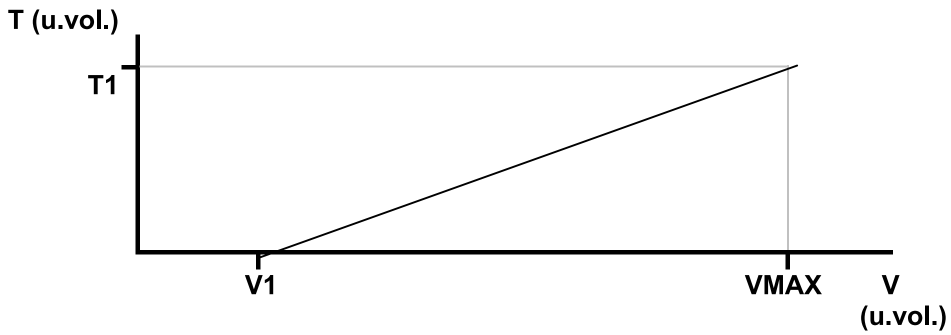
No <u>início do mês</u>, o gestor da barragem decide qual o volume de água que vai destinar à satisfação das necessidades previstas para esse mês, com base no volume de água então disponível na barragem. Sempre que possível (isto é, sempre que não seja

(continua)

(continuação)

violada a condição de equilíbrio ambiental) o volume de água destinado à satisfação das necessidades será igual ao valor previsto (indicado na tabela anterior).

Em função do volume, **V**, de água na barragem no início do mês após a afectação do volume de água destinado à satisfação das necessidades o gestor da barragem determina o volume de água a turbinar durante esse mês, de acordo com o gráfico seguinte:



O gestor da barragem pretende definir o valor dos parâmetros **V1** e **T1** da política de gestão de água a turbinar.

Por razões de segurança exige-se que V1 não seja inferior a vinte por cento da capacidade máxima de retenção.

O lucro mensal L (u.m.) associado ao volume turbinado T (u.vol.) é dado pela função $\mathbf{L = T^2}$.

Sabe-se que uma descarga total (T + DS) superior ou igual a 30 u.vol. origina situações de cheia a jusante.

Para decidir quais os valores a adoptar para os parâmetros **V1** e **T1** o gestor consultou os seus registos, podendo observar o valor das precipitações (em u.vol.) relativos a 1992, 1993 e 1994 indicados no quadro seguinte:

Precipitação Mensal (u.vol.)			
Mês	**1992**	**1993**	**1994**
J	30	43	25
F	40	48	23
M	40	39	18
A	30	28	10
M	25	29	5
J	20	22	0
J	12	15	0
A	0	0	0
S	12	13	1
O	18	20	5
N	29	28	13
D	30	29	18

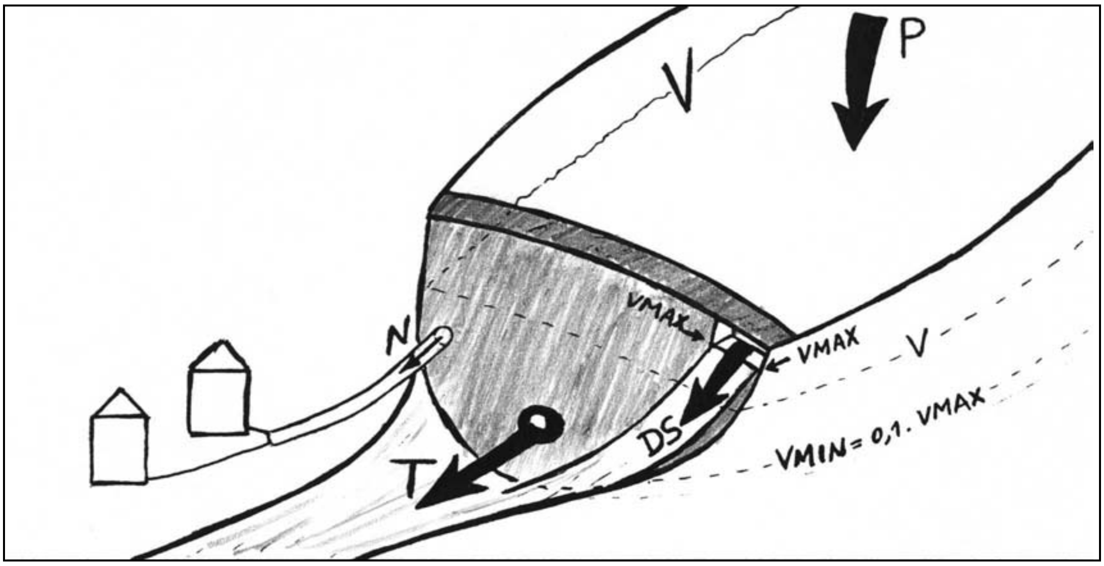
Que conselho poderemos dar ao gestor da barragem relativamente aos valores a adoptar para os parâmetros **V1** e **T1** da política de gestão de águas referida ?

Eis-nos perante um problema de **gestão de recursos hídricos** (ainda que em versão obviamente simplificada).

O **sistema** descrito no problema é constituído pela barragem, pelas precipitações (retidas pela barragem) e pelas necessidades de água a jusante. Quer as precipitações, quer as necessidades são **factores não controláveis** pelo gestor da barragem, mas que afectarão claramente a política de gestão a utilizar (o único **elemento controlável** pelo gestor).

De referir ainda que os factores não controláveis são, neste problema, **determinísticos**. É óbvio que, na realidade, tal não sucederá ! As precipitações e as necessidades de água a jusante serão certamente **aleatórias**.

Façamos uma primeira representação esquemática do sistema, para melhor compreendermos o seu funcionamento:



Fixada uma política de gestão de águas, isto é, fixados os valores dos parâmetros V1 ($V1 \geq 20$) e T1, no início de um mês genérico, o gestor turbinará um volume de água destinado à satisfação das necessidades de água a jusante. Estas necessidades serão, sempre que possível, totalmente satisfeitas - é necessário respeitar-se sempre a manutenção de um volume mínimo de água de 10 u.vol. . Assim, o volume de água a destinar à satisfação das necessidades, N (u.vol.) será função do volume de água na albufeira nessa altura (V0) e das necessidades desse mês (Nec):

$$N = \begin{cases} Nec & ; V0 - Nec \geq 10 \\ V0 - 10 & ; V0 - Nec < 10 \end{cases}$$

O volume de água na albufeira, V, passará a ser $V = V0 - N$.

De seguida, o gestor turbinará o volume T (u.vol.), que será função do nível de água na albufeira nesse momento, V (u.vol.), de acordo com a seguinte relação:

$$T = \begin{cases} 0 & ; V < V1 \\ \dfrac{T1 \cdot (V - V1)}{100 - V1} & ; V \in [V1 ; 100] \end{cases}$$

De notar que, dado que V1 não é inferior a 20 u.vol., nunca será violada a restrição ambiental pelo facto de se turbinar o volume T.

O volume na albufeira passa, então, para $V = V0 - N - T$.

De notar que o lucro mensal, L (em u.m.) é função de T (em u.vol.): $L = T^2$.

Em seguida 'processaremos' as precipitações desse mês, P. Se o volume das precipitações adicionado ao volume de água existente na albufeira ultrapassar VMAX (100 u.vol.), ter-se-á que proceder a uma descarga de superfície DS (u.vol.), dada por:

$$DS = \begin{cases} 0 & ; V0 - N - T + P \leq 100 \\ V0 - N - T + P - 100 & ; V0 - N - T + P > 100 \end{cases}$$

O volume da albufeira passa, então, para V = V0 - N - T + P - DS.

De notar que se T + DS exceder 30 u.vol. se originará uma situação de cheia a jusante.

Interessará estar atento à ocorrência quer de situações de abastecimento de água insuficiente face às necessidades, quer de situações de cheias.

Acabámos de reduzir o enunciado aos seus aspectos mais importantes.

E agora ? Como propor V1 e T1 ?

Se arbitrarmos uma **política de gestão** (isto é, valores para V1 e T1), poderemos aproveitar os dados relativos às precipitações mensais dos anos de 1992, 1993 e 1994 para **avaliar o desempenho do sistema**, isto é, avaliar o correspondente valor do Lucro (por exemplo, avaliando o valor do "Lucro Mensal Médio"), avaliar a ocorrência de situações de abastecimento de água insuficiente face às necessidades (por exemplo, avaliando o número de meses com 'restrições' e o valor da 'restrição média') e avaliar a ocorrência de situações de cheias (por exemplo, avaliando o número de meses com cheias e o valor da 'cheia média'). Poderemos designar o valor do "Lucro Mensal Médio", o número de meses com 'restrições' , o valor da 'restrição média', o número de meses com cheias e o valor da 'cheia média' como **medidas de desempenho do sistema** (ou, **medidas de 'performance'**).

Assim, **a cada política de gestão corresponderá um 'retrato' que se traduz nos valores das medidas de desempenho definidas**. Caberá ao **decisor** (neste caso ao gestor da barragem) optar pela política de gestão que considera mais *útil* (o que habitualmente se traduz num problema de 'Decisão Multicritério'...).

Para começarmos a **simular o funcionamento deste sistema** precisaremos de definir uma **condição inicial**: "Qual o volume de água da albufeira no início de Janeiro de 1992 ?" . Só a partir desse valor inicial poderemos prosseguir a simulação. Claro que se adoptarmos um valor muito baixo corremos o risco de obter um 'retrato muito negro', no que diz respeito às necessidades de água a jusante ... Se, pelo contrário, formos muito 'generosos' na indicação desse valor inicial, certamente teremos um quadro 'muito interessante' no que diz respeito às cheias ... Ora cá está uma ***questão técnica*** interessante que abordaremos posteriormente com mais detalhe: **a influência das condições iniciais na simulação**.

Relativamente ao nosso problema poderemos admitir que o volume da albufeira no início de Janeiro de 1992 era igual a 50 u.vol. (albufeira meia cheia).

Sistematizemos, em seguida, a simulação do nosso sistema:

1 - Inicializar as variáveis associadas às medidas de desempenho:

SL = somatório dos lucros mensais ;
NMCC = nº meses com cheias ; **SC** = Σ dos volumes de cheia
NMCR = nº meses com restrições ; **SR** = Σ dos vol.s corresp.s às restrições

Igualar a zero as variáveis indicadas.

2 - Indicar uma política de gestão.
Indicar o valor a atribuir à condição inicial (V0 = 50 u.vol.).

3 - Iniciar a simulação:

Repetir para os anos de 1992 a 1994:

Repetir para os meses de Janeiro a Dezembro:

$$N = \begin{cases} Nec & ; V0 - Nec \geq 10 \\ V0 - 10 & ; V0 - Nec < 10 \end{cases}$$

$$V = V0 - N$$

$$V0 - Nec < 10 \quad \Rightarrow \quad \begin{cases} NMCR = NMCR + 1 & [\text{ v. Nota 1 }] \\ SR = SR + (Nec - V0 + 10) \end{cases}$$

$$T = \begin{cases} 0 & ; V < V1 \\ \dfrac{T1 \cdot (V - V1)}{100 - V1} & ; V \in [V1 ; 100] \end{cases}$$

$$V = V0 - N - T$$

$$L = T^2 \quad ; \quad SL = SL + L$$

$$DS = \begin{cases} 0 & ; V0 - N - T + P \leq 100 \\ V0 - N - T + P - 100 & ; V0 - N - T + P > 100 \end{cases}$$

$$VF = V0 - N - T + P - DS$$

$$T + DS > 30 \quad \Rightarrow \quad \begin{cases} NMCC = NMCC + 1 \\ SC = SC + (T + DS - 30) \end{cases}$$

$$V0 = VF \quad [\text{ v. Nota 2 }]$$

4 - Apurar os valores das Medidas de Desempenho:

Lucro Médio Mensal = SL / 36

Nº Meses Com Cheias = NMCC ; Cheia Média = SC / NMCC

Nº Meses Com Restrições = NMCR ; Restrição Média = SR / NMCR

Notas: **1 - " X = X + ∆X " corresponde à instrução " Incremente-se a variável X de ∆X " ;**
2 - V0 = VF ⇒ V0 passa a ter o valor de VF (vol. no final do mês), que será o volume inicial do próximo mês e estará na base do cálculo de N.

Se se pretender fazer uma Simulação "manual", poderemos começar por representar um Quadro como o seguinte, onde destacamos os dados do problema (apresentados nas duas primeiras colunas):

Mês	P	Nec	V0	N	T	DS	VF	L	R	C
J 92	30	8	50							
F 92	40	8								
M 92	40	8								
A 92	30	9								
M 92	25	10								
J 92	20	13								
J 92	12	15								
A 92	0	15								
S 92	12	15								
O 92	18	13								
N 92	29	9								
D 92	30	8								
J 93	43	8								
F 93	48	8								
M 93	39	8								
...	...	...	...	...	...	...	...	...	...	...

Notas: V0 - volume no início do mês; N - volume destinado à satisfação de necessidades de água a jusante; T - volume turbinado; DS - volume correspondente a Descargas de Superfície; VF - Volume no final do mês; L - lucro; R - restrição ao consumo (volume em falta); C - cheia (volume em excesso além do "Limite de Cheia" 30 u.vol. para T + DS)

Indicou-se no Quadro o valor inicial V0 = 50 para Janeiro de 1992. Em função de uma dada política de gestão, isto é de valores para V1 e T1, poderíamos preencher o Quadro e obter os valores correspondentes das Medidas de Desempenho do Sistema.

Como é evidente, o preenchimento de um Quadro correspondente a uma dada Política de Gestão é uma tarefa muito pouco interessante e consideravelmente morosa para ser feita manualmente. A utilização de uma "Folha de Cálculo" (tipo EXCEL) revela-se uma opção particularmente adequada a este problema. Depois de programada a folha de cálculo, basta alterar os valores correspondentes à política de gestão, V1 e T1, para a folha ser 'recalculada' para essa nova política de gestão, com a indicação dos valores correspondentes às Medidas de Desempenho do Sistema.

Já sabemos que V1 é superior ou igual a 20 u.vol. e inferior a 100 u.vol. . Por outro lado, T1 representa o volume a turbinar quando o volume de água na albufeira é máximo (igual a 100 u.vol.), pelo que, poderemos concluir que T1 é limitado superiormente por 90 u.vol. (para não se violar a 'restrição ambiental'). Assim, numa primeira análise, poderíamos experimentar para V1 os valores 20, 30, 40, ..., 90 u.vol. e para T1 os valores 10, 20, 30, ..., 90 u.vol. .

Para cada par de valores (V1 , T1) poderemos obter os valores de lucro médio mensal, número de meses com cheia, volume médio de cheia, número de meses com restrições no abastecimento e volume médio de restrição no abastecimento.

O gestor deseja maximizar o lucro e, simultaneamente, minimizar as situações de cheias e de restrições no abastecimento, o que são objectivos obviamente 'contraditórios'. Com efeito, alguns dos valores mais elevados de lucro neste problema estão associados a situações de cheias e/ou restrições com razoável amplitude.

Apresentaremos de seguida a Folha de Cálculo correspondente à simulação do funcionamento da barragem com a política de gestão traduzida por V1 = 20,00 u.vol. e T1 = 30,00 u.vol. :

V1 =20,00 **T1=30,00** **VMAX = 100,00** **LCheia=30,00**

	P	Nec	V0	N	T	DS	VF	L	Rtr.Cons.	Cheia
J92	30	8	**50,00**	8,00	8,25	0,00	63,75	68,06	0,00	0,00
F92	40	8	63,75	8,00	13,41	0,00	82,34	179,73	0,00	0,00
M92	40	8	82,34	8,00	20,38	0,00	93,96	415,30	0,00	0,00
A92	30	9	93,96	9,00	24,36	0,00	90,60	593,50	0,00	0,00
M92	25	10	90,60	10,00	22,73	0,00	82,88	516,48	0,00	0,00
J92	20	13	82,88	13,00	18,70	0,00	71,17	349,83	0,00	0,00
J92	12	15	71,17	15,00	13,56	0,00	54,61	184,01	0,00	0,00
A92	0	15	54,61	15,00	7,35	0,00	32,26	54,07	0,00	0,00
S92	12	15	32,26	15,00	0,00	0,00	29,26	0,00	0,00	0,00
O92	18	13	29,26	13,00	0,00	0,00	34,26	0,00	0,00	0,00
N92	29	9	34,26	9,00	1,97	0,00	52,28	3,88	0,00	0,00
D92	30	8	52,28	8,00	9,11	0,00	65,18	82,93	0,00	0,00
J93	43	8	65,18	8,00	13,94	0,00	86,24	194,37	0,00	0,00
F93	48	8	86,24	8,00	21,84	4,40	100,00	476,92	0,00	0,00
M93	39	8	100,00	8,00	27,00	4,00	100,00	729,00	0,00	1,00
A93	28	9	100,00	9,00	26,63	0,00	92,38	708,89	0,00	0,00
M93	29	10	92,38	10,00	23,39	0,00	87,98	547,12	0,00	0,00
J93	22	13	87,98	13,00	20,62	0,00	76,37	425,15	0,00	0,00
J93	15	15	76,37	15,00	15,51	0,00	60,85	240,62	0,00	0,00
A93	0	15	60,85	15,00	9,69	0,00	36,16	93,99	0,00	0,00
S93	13	15	36,16	15,00	0,43	0,00	33,72	0,19	0,00	0,00
O93	20	13	33,72	13,00	0,27	0,00	40,45	0,07	0,00	0,00
N93	28	9	40,45	9,00	4,29	0,00	55,16	18,44	0,00	0,00
D93	29	8	55,16	8,00	10,18	0,00	65,97	103,72	0,00	0,00
J94	25	8	65,97	8,00	14,24	0,00	68,73	202,78	0,00	0,00
F94	23	8	68,73	8,00	15,28	0,00	68,46	233,33	0,00	0,00
M94	18	8	68,46	8,00	15,17	0,00	63,29	230,19	0,00	0,00
A94	10	9	63,29	9,00	12,86	0,00	51,43	165,31	0,00	0,00
M94	5	10	51,43	10,00	8,04	0,00	38,39	64,58	0,00	0,00
J94	0	13	38,39	13,00	2,02	0,00	23,37	4,09	0,00	0,00
J94	0	15	23,37	13,37	0,00	0,00	10,00	0,00	1,63	0,00
A94	0	15	10,00	0,00	0,00	0,00	10,00	0,00	15,00	0,00
S94	1	15	10,00	0,00	0,00	0,00	11,00	0,00	15,00	0,00
O94	5	13	11,00	1,00	0,00	0,00	15,00	0,00	12,00	0,00
N94	13	9	15,00	5,00	0,00	0,00	23,00	0,00	4,00	0,00
D94	18	8	23,00	8,00	0,00	0,00	33,00	0,00	0,00	0,00
								6886,55	47,63	1,00

Lucro Médio Mensal = 191,29

Nº Meses Com Cheias = 1 **Nº Meses Com Restrições = 5**

Cheia Média = 1,00 **Restr.Média = 9,53**

Pode-se observar um Lucro Médio Mensal de 191,29 u.m., associado à ocorrência de um único mês com cheias (de baixo valor: 1,00 u.vol. acima do "limite de cheia" 30,0 u.vol.) e à ocorrência de 5 meses (de entre os 36 meses simulados) com restrições no abastecimento de água, sendo o volume médio em falta relativamente elevado: 9,53 u.vol. Assim, embora no respeitante às cheias a situação seja francamente aceitável, as restrições no abastecimento de água parecem-nos claramente indesejáveis, pelo que a política testada (V1 = 20,00 ; T1 = 30,00) não nos parece muito recomendável.

No Quadro seguinte apresentamos os resultados obtidos nas simulações correspondentes às políticas traduzidas por V1 = 20, 30, ..., 90 / T1 = 10, 20, ..., 90:

V1→ T1↓	20	30	40	50	60	70	80	90
10	44,53 C2 R2 5,50 7,28	41,67 C3 R2 3,95 5,90	— **C4** R2	— **C4** R2	— **C4** R1	— **C4** R1	— **C4** R0	— **C4** R0
20	— C2 **R4**	119,50 C2 R3 5,50 10,2	114,78 C2 R3 5,50 8,33	108,62 C3 R3 3,80 6,31	— **C4** R2	— **C4** R2	— **C4** R0	— **C4** R0
30	— C1 **R5**	— C2 **R4**	— C2 **R4**	174,22 C2 R3 5,50 8,62	164,06 C3 R3 4,33 6,04	— **C4** R2	— **C4** R1	— **C4** R0
40	— C2 **R7**	— C2 **R4**	— C3 **R4**	— **C4** R3	— **C4** R3	— **C4** R2	— **C4** R1	— **C4** R0
50	— C2 **R8**	— **C4 R5**	— **C4 R4**	— **C4** R3	255,90 C3 R3 8,96 6,84	241,22 C3 R2 7,78 5,65	— **C4** R2	— **C4** R0
60	— **C4 R8**	— **C5 R5**	— **C4 R4**	— **C4** R3	268,23 C3 R3 11,63 7,01	269,78 C3 R2 11,33 7,00	— **C6** R0	— **C4** R0
70	— **C5 R8**	— **C4 R5**	— **C4 R4**	263,77 C3 R3 8,78 10,30	283,21 C3 R3 13,64 7,18	— **C8** R2	— **C7** R2	— **C4** R0
80	— **C4 R8**	— **C4 R5**	— **C4 R4**	— C3 **R4**	— **C4** R3	— **C5** R3	— **C7** R2	— **C4** R0
90	— **C4 R8**	— **C4 R5**	— **C5 R4**	— C3 **R4**	— **C7** R3	— **C6** R3	— **C5** R1	— **C4** R0

Legenda:

V1 = 20

T1 = 10

44,53	
C2	R2
5,50	7,28

← Lucro Médio Mensal (u.m.)
← R2: 2 meses com 'restrições'
← vol. médio de 'restrição' nesses 2 meses = 7,28 u.vol.
↖ C2: 2 meses com cheias ;
vol. médio de cheia nesses 2 meses = 5,50 u.vol.

Nota: Sempre que tenham ocorrido mais do que 3 meses com restrições, ou mais do que 3 meses com cheias, considerou-se a correspondente política como 'desadequada', pelo que não se indicou os correspondentes valores médios mensais de lucro, volume de 'restrição' ou volume de cheia.

O Quadro anterior sugere-nos alguns comentários imediatos:

- É importante saber **delinear as experiências** a levar a cabo ! Normalmente (embora, neste problema, não tenha sido o caso) a simulação correspondente a 'uma experiência' é relativamente morosa, pelo que é importante não fazer (pelo menos, de início) uma análise 'muito fina' (neste problema, por exemplo com incrementos unitários de V1 e T1...) que pouco acrescentaria, em termos de análise de resultados e que certamente se traduziria num grande acréscimo de tempo de simulação/computação. Neste problema, após uma primeira análise, poder-se-ia restringir o domínio de variação de V1 e T1 para levar a cabo uma análise 'mais fina'...

- É importante saber **escolher criteriosamente as medidas de desempenho do sistema** ! Por um lado, pretendemos ter uma 'fotografia' tão rica quanto possível mas, por outro lado, pretendemos ter uma 'fotografia' que seja facilmente analisável ! Assim, um número muito elevado de medidas de desempenho pode traduzir-se numa maior dificuldade de análise, acabando por não contribuir para uma melhor análise do funcionamento do sistema ... É, ainda, importante **saber apresentar os resultados 'inteligentemente'** ... A escolha de apresentações tabulares ou gráficas pode ser importante por permitir, com maior facilidade, a comparação do funcionamento do sistema quando submetido a diferentes 'políticas'.

- **A análise de resultados associados a uma série de 'experiências' levadas a cabo com um modelo de simulação, não é, em geral, uma tarefa *linear* ...** Se tomarmos como exemplo o problema que apresentamos (um problema de simulação muito, muito simples !), constataremos que não é óbvia a escolha da 'melhor' política (de entre as 72 que foram simuladas). Com efeito, o que é melhor para o empresário, não é certamente interessante para o consumidor (que mora a jusante da barragem) ... E o autarca que tivesse que seleccionar uma política de gestão de águas desta albufeira teria uma tarefa bem ingrata: para defender os interesses dos consumidores [Atenção às eleições autárquicas !], o autarca deveria estar disposto a renunciar a grande parte dos lucros que poderia obter !

Na nossa situação privilegiada de 'observadores independentes', poderíamos tentar fazer uma breve análise dos resultados obtidos:

Comecemos por recordar que o 'Limite de Cheia' é igual a 30 u.vol., pelo que não nos parece muito recomendável que se ultrapasse esse limite em mais de 20 %, isto é, em mais de 6 u.vol.. Por outro lado, as necessidades mensais a jusante variam, com o mês, de 8 u.vol. a 15 u.vol., apresentando um valor médio aproximadamente igual a 10 u.vol. - assim, nos meses com restrições (provavelmente, no Verão quando as necessidades atingem as 15 u.vol.) será importante tentar que o volume em falta seja minimizado (por exemplo, poderemos considerar desejável que esse volume não ultrapasse as 6 u.vol.).

Relativamente aos resultados obtidos, pode-se referir em primeiro lugar que dos 72 cenários estudados, 59 são de imediato classificados de 'desadequados' ! Dos 13 cenários 'sobreviventes' parece-nos que os cenários correspondentes a T1 superior a 40 u.vol. podem ser rejeitados, já que os volumes médios de restrição ou de cheia são inaceitavelmente elevados ! ... E já só restam 7 ! Olhando com atenção para esses sete cenários, parece--nos ser justo destacar dois:

V1 = 30 u.vol.; T1 = 10 u.vol. $\Rightarrow$ L = 41,67 u.m.; C3 (3,95 u.vol.); R2 (5,90 u.vol.)
V1 = 60 u.vol.; T1 = 30 u.vol. $\Rightarrow$ L = 164,06u.m.; C3 (4,33 u.vol.); R3 (6,04 u.vol.)

O primeiro dos dois cenários destacados é o preferível na óptica do consumidor: os dois cenários apresentam 3 meses com cheias (de entre os 36 meses simulados), embora o volume médio de cheia seja ligeiramente inferior no primeiro cenário; o volume médio correspondente às restrições é idêntico nos dois cenários mas, é importante realçar que o primeiro cenário apenas apresenta 2 meses com restrições, enquanto que o segundo cenário corresponde a 3 meses com restrições (nos 36 meses simulados).

Numa perspectiva de lucro, o primeiro cenário é 'quatro vezes pior' do que o segundo !

Assim, tal como Pilatos, lavamos as mãos e deixamos ao 'decisor político' a responsabilidade pela escolha final !

O problema apresentado merece-nos ainda um comentário adicional sobre os dados fornecidos no enunciado.

Relativamente às **necessidades de água a jusante**, o enunciado indica **valores determinísticos** (ainda que variáveis com o mês) independentes do ano. Tal corresponde, obviamente, a uma simplificação da realidade. Num ano mais quente e seco é natural que as necessidades de água sejam maiores ... Assim, as necessidades de água a jusante deveriam ser, numa representação mais *real*, traduzidos por dados de **natureza aleatória** com determinadas distribuições cujos parâmetros variariam não só com o mês, mas também com o facto de se estar perante um ano mais (ou menos) seco.

Relativamente à **precipitação**, o enunciado fornece-nos dados correspondentes aos anos de 1992, 1993 e 1994. Assim, a **simulação com** esses **'dados históricos'**, corresponderá sempre a uma avaliação do tipo "O que teria acontecido em 1992, 1993 ou 1994, se a barragem tivesse sido gerida com esta política ?". Ora, se essa abordagem pode ser importante, especialmente para avaliar o desempenho do sistema numa dada situação gravosa do passado, resta sempre a dúvida sobre o futuro e a sua imprevisibilidade/aleatoriedade. Assim, para o estabelecimento de uma 'política de gestão' com uma razoável 'confiança estatística', seria mais útil proceder <u>não</u> a uma **simulação com dados históricos correspondentes a 3 anos**, mas a uma **simulação com 'dados aleatórios'** (*com comportamento estatístico aceitável...*) **correspondentes a um período muito mais longo** (por exemplo, 100 ou 200 anos).

Tudo questões muito interessantes ...

A L G U N S 'A S P E C T O S T É C N I C O S' D A S I M U L A Ç Ã O

Nesta secção faremos uma abordagem de alguns *aspectos técnicos* mais importantes a ter em conta na abordagem de um problema de Simulação. Começaremos por apresentar inicialmente algumas noções básicas sobre fluxogramas, apresentaremos em seguida conceitos básicos de Estatística fundamentais em Simulação, discutiremos com brevidade a influência das condições iniciais num modelo de simulação e, finalmente, discutiremos a influência da duração da simulação na precisão dos resultados.

• Noções Básicas sobre Fluxogramas

Recordemo-nos do problema apresentado na secção anterior ("A Barragem de Pós-Boa") e das instruções que correspondiam ao 'núcleo' do modelo de simulação:

1 - Inicializar as variáveis associadas às medidas de desempenho, igualando-as a zero:
SL = somatório dos lucros mensais; **NMCC** = nº meses com cheia ; **SC** = Σ dos volumes de cheia;
NMCR = nº meses com restrições ; **SR** = Σ dos vol.s corresp.s às restrições

2 - Indicar uma política de gestão (V1 ; T1). Indicar o valor a atribuir à condição inicial (V0).

3 - Iniciar a simulação:

Repetir para os anos de 1992 a 1994:

Repetir para os meses de Janeiro a Dezembro:

$$N = \begin{cases} Nec & ;\ V0 - Nec \geq 10 \\ V0 - 10 & ;\ V0 - Nec < 10 \end{cases}$$

$V = V0 - N$

$$V0 - Nec < 10 \quad \Rightarrow \quad \begin{cases} NMCR = NMCR + 1 & \text{[v. Nota 1]} \\ SR = SR + (Nec - V0 + 10) \end{cases}$$

$$T = \begin{cases} 0 & ;\ V < V1 \\ \dfrac{T1 \cdot (V - V1)}{100 - V1} & ;\ V \in [V1 ; 100] \end{cases}$$

$V = V0 - N - T$

$L = T^2$; $SL = SL + L$

$$DS = \begin{cases} 0 & ;\ V0 - N - T + P \leq 100 \\ V0 - N - T + P - 100 & ;\ V0 - N - T + P > 100 \end{cases}$$

$VF = V0 - N - T + P - DS$

$$T + DS > 30 \quad \Rightarrow \quad \begin{cases} NMCC = NMCC + 1 \\ SC = SC + (T + DS - 30) \end{cases}$$

$V0 = VF$ [v. Nota 2]

4 - Apurar os valores das Medidas de Desempenho:

Lucro Médio Mensal = SL / 36
Nº Meses Com Cheias = NMCC ; Cheia Média = SC / NMCC
Nº Meses Com Restrições = NMCR ; Restrição Média = SR / NMCR

Será esta a melhor forma de apresentarmos o 'núcleo' do nosso modelo de simulação ? Cremos que não. Especialmente quando pretendemos proceder a uma posterior implementação informática do modelo, o **fluxograma** corresponde a uma apresentação 'mais natural'.

Um fluxograma é uma **representação gráfica** de uma **sequência** de **instruções**. A leitura de um fluxograma torna-se muito mais rápida, do que a leitura da correspondente série de instruções 'escritas por extenso'. É particularmente útil podermos dispor de um fluxograma, já que tal, para a generalidade das linguagens de programação, facilita a 'transcrição' das instruções no correspondente programa. Um fluxograma permite ainda uma fácil identificação de 'blocos de instruções' que se repitam num dado programa. Cada um desses 'blocos' pode ser programado como uma 'rotina', que será invocada repetidamente pelo 'programa principal' (ou, eventualmente, por outras 'rotinas').

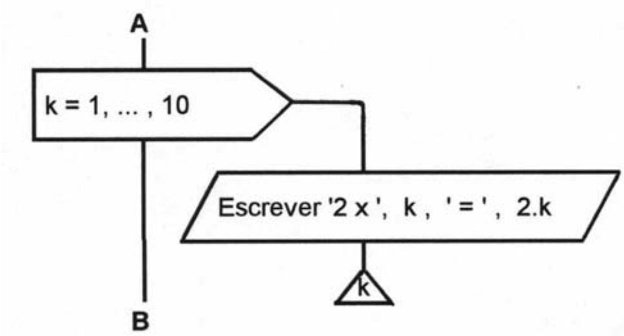
Os tipos de instruções mais importantes são:

• **instruções de atribuição** (por exemplo, "atribua-se a X o valor 2", isto é, "X = 2"), **instruções de cálculo** (por exemplo, "some-se 2 ao valor atribuído à variável X e atribua-se o resultado a Y", isto é, "Y = X + 2", ou ainda, "adicione-se 2 unidades à variável X", ou seja, "some-se 2 ao valor atribuído à variável X e atribua-se o resultado à variável X", isto é, "X = X + 2"). Num fluxograma, as instruções de atribuição e de cálculo são representadas dentro de **rectângulos**.

• **instruções de entrada ou saída** (em terminologia anglo-saxónica, "input" e "output"), ou seja, instruções de leitura ou de escrita que permitem a interacção do computador com o utilizador (por exemplo, "Pedir ao utilizador o valor de X", ou seja, "X = ?", ou ainda, "Imprimir o valor de X", isto é, "Imprimir 'X = ' seguido do valor atribuído à variável X", ou seja, "Imprimir 'X = ', X"). Num fluxograma, as instruções de entrada ou saída são representadas dentro de **paralelogramos**.

• **instruções de teste** (por exemplo, "O valor atribuído à variável X é igual a 3 ?", isto é, "X = 3 ?", ou ainda, "O valor atribuído à variável X é igual ao valor atribuído à variável Y ?", isto é, "X = Y ?"). Num fluxograma, as instruções de entrada ou saída são representadas dentro de **losangos**.

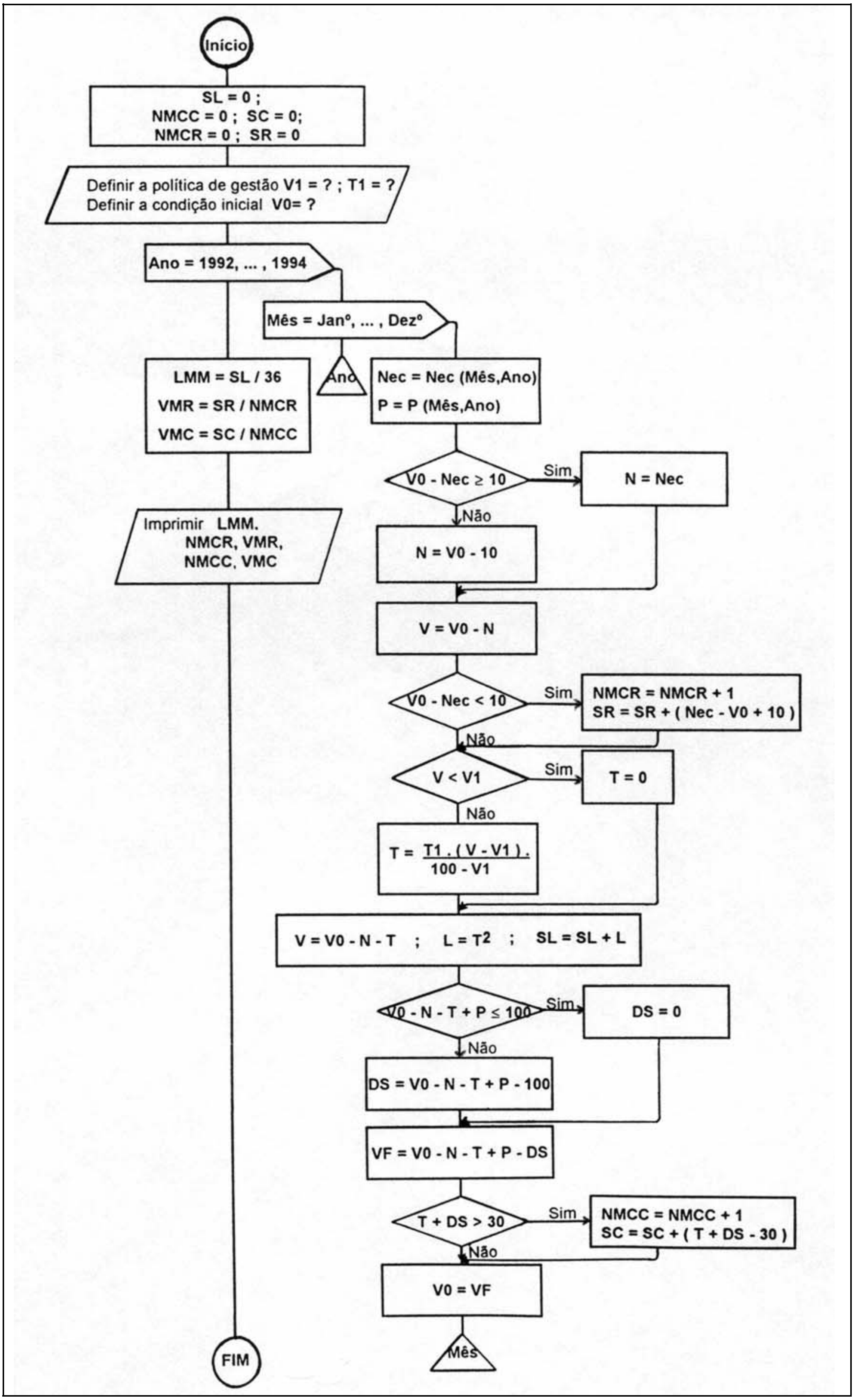
• **instruções de repetição**, ou de ciclo (por exemplo, "repetir para k de 1 até 10: "escrever '2 x ', k, ' = ', 2.k " ", isto é, escrever a tabuada dos 2). Representa-se, em seguida, o troço de fluxograma que contém a instrução que se apresentou como exemplo:



Depois de executadas a instrução **A**, inicia-se o ciclo controlado pela variável k, que é inicializada (k = 1).

Inicialmente (k=1), escreve-se '2 x 1 = 2' e k é incrementada de 1 unidade, passando a k=2. Como k ≤ 10, volta a executar-se a instrução de escrita, escrevendo-se, então, '2 x 2 = 4'. A variável k volta a ser incrementada de 1 unidade, passando a k=3. Como k ≤ 10, volta a executar-se a instrução de escrita, escrevendo-se, então, '2 x 3 = 6. A variável k volta a ser incrementada (...), passando a k=10. Como k ≤ 10, volta a executar-se a instrução de escrita, escrevendo-se, então, '2 x 10 = 20. A variável k volta a ser incrementada passando a k=11. Como k > 10, encerra-se o ciclo, passando-se à execução da instrução **B**.

Executemos agora o fluxograma correspondente ao 'núcleo do modelo desenvolvido para o problema da Barragem de Pós-Boa:


Início
SL = 0 ;
NMCC = 0 ; SC = 0;
NMCR = 0 ; SR = 0
Definir a política de gestão V1 = ? ; T1 = ?
Definir a condição inicial V0= ?
Ano = 1992, ... , 1994
Mês = Janº, ... , Dezº
Ano
Nec = Nec (Mês,Ano)
P = P (Mês,Ano)
LMM = SL / 36
VMR = SR / NMCR
VMC = SC / NMCC
V0 - Nec ≥ 10
Sim
N = Nec
Não
Imprimir LMM,
NMCR, VMR,
NMCC, VMC
N = V0 - 10
V = V0 - N
V0 - Nec < 10
Sim
NMCR = NMCR + 1
SR = SR + (Nec - V0 + 10)
Não
V < V1
Sim
T = 0
Não
T = T1 . (V - V1) . / 100 - V1
V = V0 - N - T ; L = T2 ; SL = SL + L
V0 - N - T + P ≤ 100
Sim
DS = 0
Não
DS = V0 - N - T + P - 100
VF = V0 - N - T + P - DS
T + DS > 30
Sim
NMCC = NMCC + 1
SC = SC + (T + DS - 30)
Não
V0 = VF
Mês
FIM

Relativamente ao fluxograma apresentado, poderemos referir que, tal como seria de esperar, a sequência 'gráfica' de instruções segue de modo evidente a sequência 'por extenso' apresentada anteriormente. É importante referir-se que, logo após o início do ciclo 'Mês', se afecta às variáveis **Nec** e **P**, respectivamente, os valores correspondentes das necessidades de água a jusante para aquele mês e ano (**Nec (Mês,Ano)**) e da precipitação também para esse mês e ano (**P (Mês,Ano)**). Esses valores poderiam estar 'guardados' em dois vectores, para onde previamente se deveria 'transferir' essa informação, ou, alternativamente, poder-se-ia fazer a leitura desses valores a partir de dois 'ficheiros'. Num cenário de simulação 'a sério', isto é, para uma simulação sem ser com dados históricos, esse rectângulo do fluxograma poderia corresponder a "**Gerar P e Nec, de acordo com as distribuições previamente indicadas**". Retomaremos posteriormente esta questão.

De referir ainda que, a maior parte das instruções peretencem ao ciclo 'Mês', isto é, são executadas 'todos os meses'. Com a instrução V0 = VF, isto é, atribuição do valor de VF a V0, termina-se o ciclo 'Mês'.

O ciclo 'Ano' apenas contém 'uma instrução': o ciclo 'Mês'. Terminado o ciclo 'Ano', são calculados os valores das **medidas de desempenho do sistema**: LMM (lucro médio mensal), VMR (volume médio de 'restrições') e VMC (volume médio de cheias). Após a impressão dos valores das referidas medidas de desempenho, bem como de NMCR (número de meses com 'restrições') e NMCC (número de meses com cheias) termina-se a execução.

☞ Se sabe programar computadores, esta é uma boa altura para converter o fluxograma apresentado num programa muito simples, que lhe permitirá testar diferentes políticas de gestão de águas da Barragem de Pós-Boa !

Se não sabe programar, não sabe o que está a perder ! ...

• Conceitos Básicos de Estatística Fundamentais em Simulação

Há alguns conceitos básicos de Estatística que são particularmente importantes para se elaborar um modelo de simulação com 'qualidade'.

No exemplo que apresentamos anteriormente ('A Barragem de Pós-Boa'), efectuamos uma simulação com dados históricos, correspondente aos anos de 1992, 1993 e 1994. Como se viu, para cada política de gestão de águas interessava-nos obter algumas Medidas de Desempenho do Sistema, para podermos ajuizar da adequabilidade dessa política de gestão. Embora tivéssemos simulado apenas três anos, não tem grande interesse saber o que aconteceu 'mês a mês'. Preferimos obter uma 'visão global' com menor volume de informação - daí que o **Valor Médio** seja uma noção que, obviamente, esteja sempre subjacente ao apuramento da generalidade das Medidas de Desempenho de um Sistema.

Claro que todos sabemos estimar o valor médio de uma amostra. Basta recorrer ao **estimador 'Média Amostral'** $\bar{X}$ e, obter uma **estimativa pontual** $\bar{x}$, a partir de uma amostra ($x_1, x_2, ..., x_n$):

$$\bar{X} = \left(\sum_{i=1}^{n} X_i \right) / n \qquad \rightarrow \qquad \bar{x} = \left(\sum_{i=1}^{n} x_i \right) / n$$

Ao utilizarmos o estimador Média Amostral reduzimos consideravelmente a informação: fazemos corresponder à amostra ($x_1, x_2, ..., x_n$) apenas um valor - a estimativa pontual do valor médio, $\bar{x}$. Esta estimativa é muito importante, já que, *grosso modo*, representa o 'centro de gravidade' da distribuição de X.

Poderemos discutir a *qualidade* dessa estimativa pontual, acabando por introduzir a noção de **estimativa por intervalo de confiança** que também será muito útil nos modelos de Simulação.

Antes, porém, parece-nos mais importante referir que, em muitas aplicações, não só é importante, para uma dada variável, ter noção da **localização** do 'centro de gravidade' da sua distribuição, como também é relevante ter-se noção da **dispersão** dessa mesma distribuição. A **Variância** é uma importante medida da dispersão de uma variável aleatória, podendo-se, a partir do estimador 'Variância Amostral' S'^2 e de uma amostra (x_1, x_2, ..., x_n) obter uma estimativa pontual s'^2:

$$S'^2 = \frac{n}{n-1} \cdot \left[\left(\sum_{i=1}^{n} X_i^2 \right) / n - \bar{X}^2 \right]$$

$$\downarrow$$

$$s'^2 = \frac{n}{n-1} \cdot \left[\left(\sum_{i=1}^{n} x_i^2 \right) / n - \bar{x}^2 \right]$$

O **Desvio Padrão** (igual à raiz quadrada da Variância), é expresso nas mesmas unidades que o Valor Médio.

O conhecimento do valor médio e do desvio padrão de uma variável permite-nos, de imediato, ter uma noção muito razoável da sua 'distribuição'. Com efeito, a **Desigualdade de Tchebycheff** indica que, para uma dada variável aleatória X (contínua ou discreta e independentemente da sua distribuição) com valor médio μ e desvio padrão σ e para qualquer número positivo k, se tem:

$$\mathbf{P (| X - \mu | < k . \sigma) \geq 1 - 1/k^2}$$

Ou seja, para k = 2, tem-se $P (\mu - 2 . \sigma \leq X \leq \mu + 2 . \sigma) \geq 0{,}750$ e, para k = 3, tem-se $P (\mu - 3 . \sigma \leq X \leq \mu + 3 . \sigma) \geq 0{,}889$.

O conhecimento do valor médio e do desvio padrão (ou da variância) da distribuição de uma variável são, como se referiu, muito importantes. No entanto, há situações em que é necessário conhecer algo mais, com vista à elaboração de propostas de ajustamento de distribuição. O que se passa com a **simetria da distribuição** ? E com o seu **'pico'** ? Para tal, poderemos determinar estimativas do **coeficiente de assimetria** e do **coeficiente de kurtosis**.

Sabe-se que uma distribuição simétrica tem coeficiente de assimetria nulo (mas, o inverso nem sempre é verdade...). Sabe-se que um coeficiente de assimetria positivo corresponde a uma distribuição assimétrica com 'cauda à direita', isto é, com maior probabilidade de ocorrência de valores extremos elevados (*máximos*), do que de valores extremos baixos (*mínimos*). A distribuição Exponencial tem coeficiente de assimetria igual a 2 (independentemente do valor médio da distribuição).

Por outro lado, a distribuição Normal tem um coeficiente de kurtosis igual a 3 (independentemente dos valores dos parâmetros) e um coeficiente superior a 3 indica um 'pico' mais acentuado que o da distribuição Normal.

Assim, antes de se pensar em ajustar uma distribuição Normal a um conjunto de observações é bom verificar-se se as estimativas do coeficiente de assimetria e do coeficiente de kurtosis se aproximam, respectivamente, de 0 e de 3.

Definamos, em seguida, o Coeficiente de Assimetria e o Coeficeiente de Kurtosis de uma distribuição (para tal, designemos por μ o valor médio, por σ o desvio padrão e por μ_i o i-ésimo momento central da distribuição):

Coeficiente de Assimetria	Coeficiente de Kurtosis
$\gamma_1 = \mu_3 / \sigma^3$ $(\mu_3 = E [(X - \mu)^3])$	$\gamma_2 = \mu_4 / \sigma^4$ $(\mu_4 = E [(X - \mu)^4])$

Designando por s' a estimativa do desvio padrão e por S1 o somatório das n observações x_1, x_2, ..., x_n , S2 o somatório dos quadrados das n observações e S3 o somatório dos cubos das n observações, pode-se determinar, do modo seguinte, uma **estimativa do valor do Coeficiente de Assimetria** :

$$g_1 = [\ S3 / n \ - \ 3 . S1 . S2 / n^2 \ + \ 2 . S1^3 / n^3 \ ] \ / \ s'^3$$

Designando, adicionalmente, por S4 o somatório das quartas potências das n observações x_1, x_2, ..., x_n , pode-se determinar, do modo seguinte, uma **estimativa do valor do Coeficiente de Kurtosis** :

$$g_2 = [\ S4 / n \ - \ 4 . S1 . S3 / n^2 \ + \ 6 . S1^2 . S2 / n^3 \ - \ 3 . S1^4 / n^4 \ ] \ / \ s'^4$$

Já tivemos a oportunidade de referir a grande importância da estimação pontual do valor médio da distribuição de uma variável aleatória. Desde logo se referiu que uma alternativa 'mais rica' à estimação pontual do valor médio é a **estimação por intervalos de confiança do valor médio**. Nesta estimação, envolvemos de algum modo a variabilidade da amostra, o que poderá ser muito útil em determinados modelos de simulação (posteriormente, aquando da abordagem da questão 'duração da simulação *versus* precisão dos resultados', aplicaremos este conceito).

Se X for proveniente de uma população Normal com média desconhecida μ e desvio padrão σ conhecido, poderemos indicar o **Intervalo de Confiança a 95 % para o Valor Médio μ:**

$$[\ \overline{X} - 1{,}96 . \sigma / \sqrt{n} \ ; \ \overline{X} + 1{,}96 . \sigma / \sqrt{n}\ ]$$

Para **n elevado**, o que acontece quando se está a simular 'a sério', poderemos, sem grande perda de rigor, concretizar uma **estimativa do Intervalo de Confiança a 95 % para o Valor Médio μ a partir de uma amostra (x_1, x_2, ..., x_n)**:

$$[\ \overline{x} - 1{,}96 . s' / \sqrt{n} \ ; \ \overline{x} + 1{,}96 . s' / \sqrt{n}\ ]$$

($\overline{x}$ e s' designam respectivamente as estimativas pontuais do valor médio e do desvio padrão obtidas a partir da amostra referida).

É importante notar que, para um dado nível de confiança, **a amplitude do intervalo de confiança para o Valor Médio**

- é independente da estimativa do valor médio
- aumenta com s', isto é, com a variabilidade da amostra
- diminui com o tamanho da amostra (embora não linearmente, mas com a raiz quadrada de n !).

A amplitude do intervalo de confiança para o valor médio aumenta com o nível de confiança... No entanto, tal não nos interessará particularmente, já que, nos modelos de simulação, só consideraremos o nível de confiança 95 %.

Um último conceito básico de Estatística que se revela muito útil em algumas análises de resultados de experiências com modelos de simulação diz respeito ao estudo da eventual **relação linear existente entre duas variáveis**.

Sejam (x_1 , y_1), (x_2 , y_2), ... , (x_n , y_n) pares de observações correspondentes às variáveis X e Y. Se se admitir a existência de uma relação linear entre as variáveis X e Y dada por **Y = A + B . X**, poderemos determinar, pelo Método dos Mínimos Quadrados as estimativas **a** e **b** seguintes, respectivamente, dos coeficientes **A** e **B**:

$$b = \frac{\Sigma x_i \cdot y_i - (\Sigma x_i \cdot \Sigma y_i)/n}{\Sigma x_i^2 - (\Sigma x_i)^2/n}$$

Nota: Todos os somatórios correspondem a i = 1, ... , n.

$$a = \Sigma y_i / n - b \cdot (\Sigma x_i / n)$$

Para avaliar a adequabilidade do ajustamento linear, poderemos estimar $\rho_{X,Y}$ com

$$r_{X,Y} = \frac{\Sigma x_i \cdot y_i - (\Sigma x_i \cdot \Sigma y_i)/n}{\sqrt{[\Sigma x_i^2 - (\Sigma x_i)^2/n] \cdot [\Sigma y_i^2 - (\Sigma y_i)^2/n]}}$$

Nota: Todos os somatórios correspondem a i = 1, ... , n.

Como se sabe, $\rho_{X,Y}$ toma valores no intervalo [-1 ; 1]. Valores muito próximos de + 1 ou de -1 indicam um muito bom ajustamento linear correspondente a uma recta com declive, respectivamente, positivo ou negativo . Valores entre -0,5 e 0,5 indicam um fraco ajustamento linear. Sabe-se que se X e Y são variáveis aleatórias independentes, então $\rho_{X,Y}$ = 0. Contudo não é legítimo afirmar que se $\rho_{X,Y}$ = 0, então X e Y são independentes ...

• Influência das Condições Iniciais num Modelo de Simulação

Recordemo-nos do problema "A Barragem de Pós-Boa". O modelo desenvolvido necessita de um **valor inicial** de volume de água na albufeira (V0) para se dar início à simulação. Nas simulações levadas a cabo e cujos resultados foram apresentados anteriormente adoptou-se V0 = 50 (isto é, a albufeira estava "meia cheia").

É evidente a influência de V0 nos resultados da simulação do funcionamento da Barragem de Pós-Boa relativa aos anos de 1992, 1993 e 1994. Quanto maior for o volume inicial, menor será a probabilidade de ocorrência de situações de seca e maior será a probabilidade de ocorrência de situações de cheia.

Nas simulações levadas a cabo com V0 = 50 u.vol. destacaram-se duas políticas de gestão de água: **V1 = 30 u.vol.; T1 = 10 u.vol.** e **V1 = 60 u.vol.; T1 = 30 u.vol.** . No Quadro seguinte observaremos a variação de desempenho dessas duas políticas em função de V0 (V0 = 10, 20 , ... , 100) :

V0→ (V1;T1)	**10**		**20**		**30**		**40**		**50**		**60**		**70**		**80**		**90**		**100**	
(30;10)	39,55		39,70		40,19		40,85		41,67		—		—		—		—		—	
	C3 3,95	R3 5,90	C3 3,95	R2 5,90	C3 3,95	R2 5,90	C3 3,95	R2 5,90	C3 3,95	R2 5,90	**C4**	R2	**C4**	R2	**C4**	R2	**C5**	R2	**C5**	R2
(60;30)	—		148,38		151,81		158,38		164,06		166,88		171,79		175,83		183,19		193,88	
	C2	**R4**	C2 5,50	R3 6,04	C2 5,50	R3 6,04	C2 5,50	R3 6,04	C3 4,33	R3 6,04	C3 4,33	R3 6,04	C3 4,33	R3 6,04	C3 4,33	R3 6,04	C3 4,33	R3 6,04	C3 4,33	R3 6,04

Legenda: **V0 = 10**

(V1;T1) = (30;10)

39,55	
C3 3,95	R3 5,90

← Lucro Médio Mensal (u.m.)
← R3: 3 meses com 'restrições'
← vol. med. 'restrição' nesses 2 meses = 5,90 u.vol.
↖ C3: 3 meses com cheias ; vol. médio de cheia nesses 3 meses = 3,95 u.vol.

Nota: Sempre que tenham ocorrido mais do que 3 meses com restrições, ou mais do que 3 meses com cheias, considerou-se a correspondente política como 'desadequada'.

Como se pode observar, o valor inicial V0 correspondente ao volume de água inicialmente na albufeira exerce uma razoável influência na simulação levada a cabo. Para cada uma das duas políticas testadas, podemos observar um aumento de lucro com V0 (o que é natural pois, havendo mais água na albufeira, mais água se pode turbinar e maior será o lucro). Observa-se, ainda, que um aumento de V0 está geralmente associado a uma aumento de situações de cheia e (claro está !) a uma diminuição das situações de 'restrições'. É importante recordarmo-nos que, quer o lucro, quer os volumes de cheia e de 'restrições' são variáveis médias mensais ... E, ainda assim, foi possível detectar a influência de V0.

Se tivéssemos simulado o funcionamento da barragem durante 10 anos, certamente que a influência de V0 se teria atenuado ligeiramente. E, é claro que se estivessemos a simular um número de anos muito elevado (por exemplo, 100), não seria muito credível que essas **medidas de desempenho *médio*** permitissem observar uma grande influência de V0. Com efeito, **em geral, quanto maior for a 'duração' da simulação, menor será a influência das condições iniciais**.

Assim, quando se recear que as condições iniciais possam influenciar os resultados da simulação, poderemos **reservar um período inicial da duração da simulação para atenuar a influência das condições iniciais**. Nesse período, é feita a simulação do sistema, mas não se 'actualiza' as variáveis que se destinam a apurar as medidas de desempenho do sistema. Por exemplo, no problema "A Barragem de Pós-Boa", não se actualizaria as variáveis SL, NMCC, SC, NMCR e SR durante esse "período inicial"; nesse período apenas se registaria a evolução do volume de água na albufeira. Na secção **IV** destes apontamentos retomaremos este problema e teremos oportunidade para aprofundar a 'questão da influência das condições iniciais'.

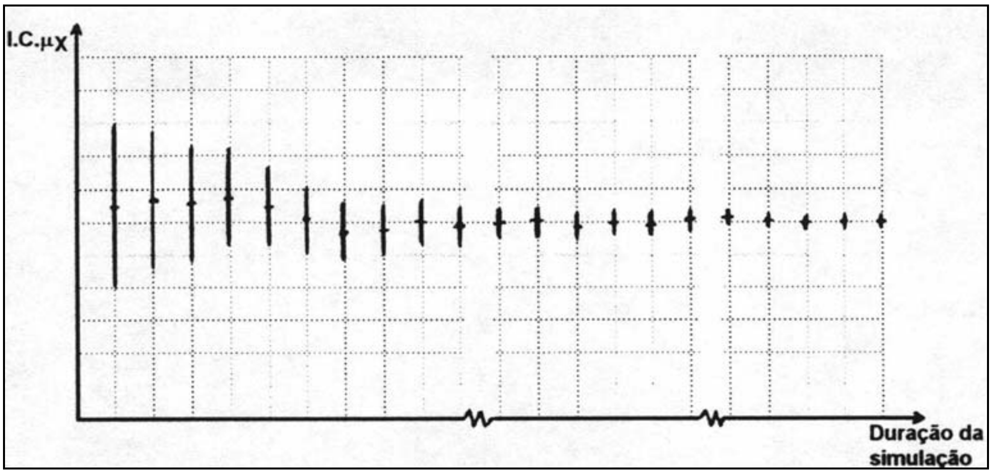
Um alerta deve ser feito relativamente à influência das condições iniciais: há sistemas que são muito sensíveis aos valores atribuídos às variáveis que traduzem as condições iniciais. Nesses casos, é preciso ter-se muito cuidado na atribuição desses valores, já que uma escolha irreflectida de valores para as variáveis que traduzem as condições iniciais pode 'desiquilibrar' irremediavelmente o sistema, não permitindo uma análise adequada .... Imaginemos que se está a simular o funcionamento de uma instituição bancária. A escolha de uma 'situação inicial muito má' pode, em determinados cenários, acarretar a falência (simulada) da instituição; pelo contrário, a escolha de uma 'situação inicial muito confortável' pode originar uma elevada rentabilidade (simulada) da instituição completamente desajustada da realidade.

• Duração da Simulação *versus* Precisão dos Resultados

Empiricamente podemos esperar que a um **aumento da 'duração' de uma simulação** esteja associado um **aumento da precisão dos resultados**. Com efeito, e tomando o problema "A Barragem de Pós-Boa" como exemplo, quando se aumenta a 'duração' da simulação (isto é, quando se aumenta o número de anos simulados), as **medidas de *desempenho médio* do sistema** passam a ser estimadas a partir de amostras maiores e, consequentemente, diminui a amplitude dos correspondentes intervalos de confiança para os valores médios dessas medidas de desempenho. Simular o funcionamento da Barragem de Pós-Boa durante 3 anos significa estimar um lucro médio mensal a partir de 36 observações. Aumentar o período simulado para 100 anos, corresponde a estimar o lucro médio mensal a partir de uma amostra com 1200 observações !

É importante recordarmo-nos que a amplitude de um intervalo de confiança para o valor médio de uma dada variável diminui com o aumento do tamanho da amostra, que serve de base ao cálculo desse intervalo. Contudo, essa **diminuição da amplitude** não é **proporcional** ao tamanho da amostra mas, **à raiz quadrada do tamanho da amostra**.

Daqui decorre uma importante constatação: **à medida que se aumenta a duração da simulação, a diminuição da amplitude dos intervalos de confiança para os valores médios das medidas de desempenho do sistema é cada vez menor** (embora continue a ocorrer), isto é, **verifica-se um acréscimo de precisão dos resultados cada vez menor**. Na figura seguinte esboça-se a variação da amplitude de um intervalo de confiança para o valor médio de uma medida de desempenho, X, de um sistema que está a ser simulado com a duração da simulação:



Como se pode observar, a partir de certa altura só com um grande acréscimo na duração da simulação / tempo de computação se consegue um pequeno acréscimo de precisão (isto é, uma diminuição ligeira da amplitude do intervalo de confiança).

Assim, fixar a duração de uma simulação é uma decisão muito importante, já que terá reflexos na precisão dos resultados.

Como, em geral, se desconhece a variabilidade das variáveis que são utilizadas para avaliar as medidas de desempenho do sistema, não é habitual indicar, *a priori*, a duração da simulação, tendo em vista a obtenção de determinado nível de precisão dos resultados. Alternativamente, pode-se **impor um determinado nível de precisão dos resultados** (explicitando as amplitudes máximas dos intervalos de confiança respectivos) e **indicar uma duração máxima para a simulação** (não vá dar-se o caso de se ter exigido uma precisão inatingível num tempo de simulação/computação aceitável). O próprio modelo de simulação poderá, **periodicamente, verificar se os níveis de precisão impostos para as diferentes variáveis são cumpridos** e, em caso afirmativo, terminar a simulação; caso contrário a simulação prossegue até nova avaliação, respeitando-se adicionalmente a duração máxima imposta inicialmente - trata-se de um **'*mecanismo auto-stop*'**. Também é habitual exigir-se uma **duração mínima da simulação**, independentemente da precisão exigida aos resultados, visando, assim, garantir-se uma *solidez estatística* mínima dos resultados. Esquematicamente, pode-se representar as diferentes *fases* de uma simulação, do modo seguinte: