Considere $X_1,\ X_2,\ \ldots,\ X_N$ números naturais e definam-se, para $i = 1, 2, \ldots, N$, $p_i = \dfrac{X_i}{\sum_{j=1}^{N} X_j}$.
Um algoritmo para a construção de uma variável aleatória Y cujos valores possíveis são 1, 2, …, N, consiste em repetir os seguintes passos o número necessário de vezes:

- fixa-se um inteiro qualquer $X^* \geq \max_{j=1,\ldots,N} X_j$;
- escolhem-se, ao acaso e de forma independente, um inteiro $r \leq N$ e um real $s \in (0, X^*)$;
- se $X_r \geq s$ fazemos $Y = r$, caso contrário ignora-se o par (r, s) e repete-se o passo anterior até aceitarmos um par.

Este algoritmo produz um valor para a variável Y, mas pode exigir várias repetições dos três passos descritos até que se consiga construir esse valor.

a) Mostre que $\mathrm{P}(Y = i, \text{uma repetição}) = \dfrac{1}{N}\dfrac{X_i}{X^*}$ e $\mathrm{P}(\text{uma repetição}) = \dfrac{1}{NX^*}\displaystyle\sum_{j=1}^{N} X_j =: p$.

b) Mostre que $\mathrm{P}(Y = i, \text{duas repetições}) = (1-p)\dfrac{X_i}{NX^*}$ e $\mathrm{P}(\text{duas repetições}) = (1-p)p$.

c) Proceda por indução para concluir que

$$\mathrm{P}(Y = i, m \text{ repetições}) = (1-p)^{m-1}\frac{X_i}{NX^*} \qquad \text{e} \qquad \mathrm{P}(m \text{ repetições}) = (1-p)^{m-1}p\,.$$

d) Mostre que $\mathrm{P}(Y = i) = \dfrac{X_i}{NX^*}$.

e) Mostre que o número médio de repetições até obter um valor para Y é igual a $\frac{1}{p}$, havendo portanto conveniência, do ponto de vista computacional, em escolher X^* o mais pequeno possível.

Exemplo: Explica-se de seguida o que se entendo por repetições do algoritmo descrito. Sejam $X_1 = 3$, $X_2 = 5$, $X_3 = 4$, $X_4 = 6$ e $X_5 = 2$. Logo $X_1 + X_2 + X_3 + X_4 + X_5 = 20$, $p_1 = \frac{3}{20} = 0.15$, $p_2 = \frac{5}{20} = 0.25$, $p_3 = \frac{4}{20} = 0.20$, $p_4 = \frac{6}{20} = 0.30$ e $p_5 = \frac{2}{20} = 0.10$.

1. Fixe-se $X^* = 10$. Escolha-se um inteiro $r \leq 5$ e um real $s \in (0, 10)$ ao acaso. Suponha-se $r = 3$ e $s = 2.342$. Como $X_r = X_3 = 4 \geq s = 2.342$ tomamos $Y = 3$ e obtivemos a simulação executando apenas uma repetição.

2. Fixe-se $X^* = 15$ Escolhe-se agora um inteiro $r \leq 5$ e um real $s \in (0, 15)$ ao acaso. Suponha-se que $r = 5$ e $s = 10.345$. Neste caso $X_r = X_5 = 2 < s = 10.345$ pelo que se rejeita o par $(5, 10.345)$. Procede-se a uma nova escolha de $r = 1$ e $s = 6.925$. Neste caso $X_r = X_1 = 3 < s = 6.925$, pelo que voltamos a rejeitar este par. Mais uma escolha de $r = 4$ e $s = 5.728$, que dá $X_r = X_4 = 6 \geq s = 5.728$ pelo que tomamos $Y = 4$ e necessitámos de 3 repetições dos passos do algoritmo para construir a simulação para Y.

Prazo de entrega: 13 de Março de 2009
paulo@mat.uc.pt