# Método de la urna
# Método de tablas
# Método del alias

**Patricia Kisbye**

FaMAF

8 de abril, 2010

# Método de la transformada inversa

$X : \{x_1, x_2, \dots\}$, con $P(X = x_i) = p_i$.

**Algoritmo: Transformada Inversa**

$F \leftarrow p_1$, $i \leftarrow 1$;
Generar $U \sim \mathcal{U}(0, 1)$;
**while** $U > F$ **do**
    $i \leftarrow i + 1$;
    $F \leftarrow F + p_i$;
**end**
$X \leftarrow x_i$

- El algoritmo recorre desde un valor en adelante.
- Para una v.a. discreta en general, realiza muchas búsquedas.

Veamos algunas alternativas.

# Método de la Urna

- $X$ toma un número finito de valores: $\{x_1, x_2, \ldots, x_n\}$
- Cada $p_i = k_i \, 10^{-q}$, con $k_i$ entero y $q$ fijo. Esto es:

$$\sum_{i=1}^{n} k_i = 10^q$$

  $q$: es el número máximo de dígitos decimales.
- $v$: un vector de tamaño $10^q$ que contiene $k_i$ veces cada elemento $x_i$.

$$v = \underbrace{x_1 \ldots x_1}_{k_1} \underbrace{x_2 \ldots x_2}_{k_2} \ldots \underbrace{x_n \ldots x_n}_{k_n}$$

---

Algoritmo: Variante A

---

Generar $U \sim \mathcal{U}(0,1)$;
$J \leftarrow \lfloor 10^q U \rfloor + 1$;
$X \leftarrow v_J$

---

# Método de la Urna

## Ejemplo

$X$ toma valores $x_1 = 8$, $x_2 = 9$, $x_3 = 10$ y $x_4 = 11$, con probabilidades $p_1 = 0.2$, $p_2 = 0.1$, $p_3 = 0.4$ y $p_4 = 0.3$.

$$q = 1$$

$$p_1 = 2 \cdot 10^{-1},\ p_2 = 1 \cdot 10^{-1},\ p_3 = 4 \cdot 10^{-1},\ p_4 = 3 \cdot 10^{-1}.$$

$$v = \boxed{8 \mid 8 \mid 9 \mid 10 \mid 10 \mid 10 \mid 10 \mid 11 \mid 11 \mid 11}$$

$$U = 0.378 \quad \Rightarrow \quad J = 4 \quad \Rightarrow \quad X \leftarrow 10$$

# Método de la Urna

¿Por qué funciona?

$$
\begin{aligned}
P(X = j) &= P(k_1 + \dots k_{j-1} < \lfloor 10^q U \rfloor \leq k_1 + \dots k_{j-1} + k_j) \\
&= \frac{k_j}{10^q} = p_j
\end{aligned}
$$

- Desventaja: Necesita $10^q$ lugares de almacenamiento.
- Advertencia: Si se redondean los datos para utilizar un $q$ más chico, recordar que $\sum k_i$ debe resultar 1.

# Método de la Tabla

Marsaglia (1963) propone la siguiente mejora el método de la urna.

- Se utiliza un vector $v_k$, para cada posición decimal: $k = 1, 2, \ldots, q$.
- En $v_k$ se almacena $x_i$ tantas veces como indique su posición decimal $k$-ésima.
- Se calcula $d_k$ la probabilidad de ocurrencia del dígito de orden $k$:

$$d_k = \frac{\text{suma de los dígitos de orden } k \text{ en las probabilidades}}{10^k}$$

# Método de la Tabla

## Ejemplo

Sea $X$ con distribución $p(1) = 0.15$, $p(2) = 0.20$, $p(3) = 0.43$, $p(4) = 0.22$,

$$v = \boxed{\begin{array}{|c|c|c|c|c|c|c|c|c|} 1 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 \end{array}} \quad m = 9$$

$$w = \boxed{\begin{array}{|c|c|c|c|c|c|c|c|c|c|} 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 4 & 4 \end{array}} \quad n = 10$$

$$d_1 = \frac{1+2+4+2}{10} = 0.9 \quad d_2 = \frac{5+0+3+2}{100} = 0.1$$

# Método de la tabla

$m = 9$
$n = 10$
$d_1 = 0.9$
$d_2 = 0.1$

---

Método de la tabla (Ejemplo)

---

**Input**: $v$, $w$, $m$, $n$
Generar $U \sim \mathcal{U}(0,1)$;
**if** $U < 0.9$ **then**
    Generar $V \sim \mathcal{U}(0,1)$;
    $J \leftarrow \lfloor mV \rfloor$;
    $X \leftarrow v_J$
**else**
    Generar $V \sim \mathcal{U}(0,1)$;
    $J \leftarrow \lfloor nV \rfloor$;
    $X \leftarrow w_J$
**end**

---

# Método de la tabla

¿Por qué funciona el algoritmo?

$$
\begin{aligned}
P(X=3) &= P(X=3 \mid \text{elegir decenas})P(\text{elegir decenas}) \\
&\quad + P(X=3 \mid \text{elegir centenas})P(\text{elegir centenas}) \\
&= \frac{4}{9}\,0.9 + \frac{3}{10}\,0.1 \\
&= 0.43
\end{aligned}
$$

- Ventaja sobre el anterior: mucho menor costo de almacenamiento. (19 lugares en lugar de 100).
- Desventaja: "Mayor" tiempo de ejecución.

# Método del Alias

- Este método fue presentado por A. J. Walker (1977).
- Asumimos $X : \{x_1, x_2, \ldots, x_n\}$.
- Se requiere calcular dos vectores de longitud $n$.
  - Un vector con valores de corte: $F_i \in (0, 1)$, $i = 1, 2, \ldots, n$.
  - Un vector de *alias*: $A_i \in$ Rango($X$), $i = 1, 2, \ldots, n$.
- Los valores de corte y los alias no son únicos.

Método de Alias

Generar $U, V \sim \mathcal{U}(0, 1)$;
$I \leftarrow \lfloor nU \rfloor + 1$;
**if** $V \leq F_I$ **then**
  $X \leftarrow x_I$
**else**
  $X \leftarrow A_I$
**end**

# Método del alias

---

Método para construir los vectores de alias

---

**for** $i = 1$ **to** $n$ **do**

  $F_i \leftarrow n\, p_i$

**end**

$G = \{i \mid F_i > 1\}$;

$S = \{i \mid F_i < 1\}$;

**while** *$S$ no esté vacío* **do**

  Quitar un elemento $k$ de $G$ y $j$ de $S$;

  $A_j \leftarrow X_k$;

  $F_k \leftarrow F_k - 1 + F_j$;

  **if** $F_k > 1$ **then**

    volver $k$ a $G$

  **else**

    si $F_k < 1$ agregar $k$ a $S$

  **end**

**end**

---

# Método del Alias

## Ejemplo

$X$ toma valores $x_1$, $x_2$, $x_3$, $x_4$, con probabilidades $p_1 = 0.3125$, $p_2 = 0.125$, $p_3 = 0.375$, $p_4 = 0.1875$.

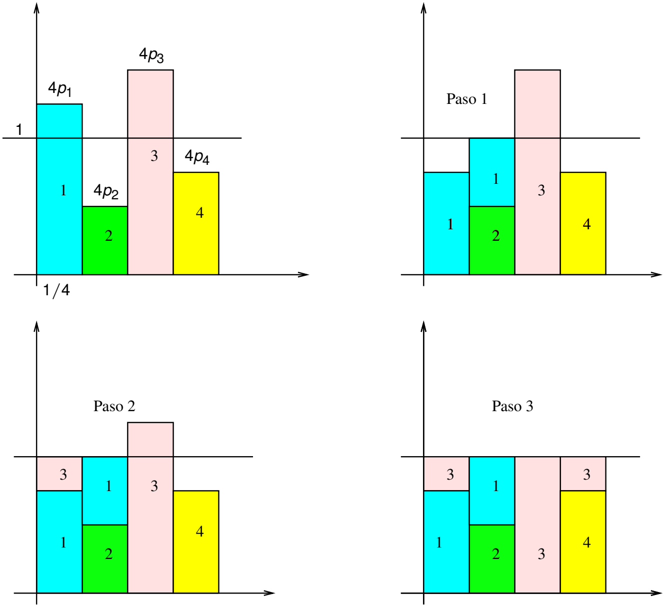
Paso 1 : $G = \{0, 2\}$,
$S = \{1, 3\}$

Paso 2 : $G = \{2\}$,
$S = \{0, 3\}$

Paso 3 : $G = \{2\}$, $S = \{3\}$

Paso 4 : $G = \{\ \}$, $S = \{\ \}$

| $i$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $F_i$ | 1.25 | 0.5 | 1.5 | 0.75 |
| $A_i$ | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
| $F_i$ | 0.75 | 0.5 | 1.5 | 0.75 |
| $A_i$ | $x_1$ | $x_1$ | $x_3$ | $x_4$ |
| $F_i$ | 0.75 | 0.5 | 1.25 | 0.75 |
| $A_i$ | $x_3$ | $x_1$ | $x_3$ | $x_4$ |
| $F_i$ | 0.75 | 0.5 | 1 | 0.75 |
| $A_i$ | $x_3$ | $x_1$ | $x_3$ | $x_3$ |



4p1
4p3
1
3
4p4
1
4p2
4
2
1/4
Paso 1
1
3
1
2
4
Paso 2
3
1
3
1
2
4
Paso 3
3
1
3
1
2
3
4