Generating random permutations with a prescribed set of descents

Philippe Marchal

Generating and counting

σ permutation of $\{1, 2, \dots n\}$.
Descent set : $D(\sigma) = \{i \in [1, n-1],\ \sigma(i) > \sigma(i+1)\}$.
Goal : generating at random, with uniform probability, and counting permutations with a given descent set A.

Generating and counting

σ permutation of $\{1, 2, \dots n\}$.
Descent set : $D(\sigma) = \{i \in [1, n-1],\ \sigma(i) > \sigma(i+1)\}$.
Goal : generating at random, with uniform probability, and counting permutations with a given descent set A.

Counting

Inclusion-exclusion formula : number of permutations with descent set A is

$$\sum_{B=\{i_1<i_2\dots<i_k\}\subset A^c} (-1)^{|A^c|-|B|} \binom{n}{n-i_k, i_k-i_{k-1}, \dots, i_2-i_1, i_1}$$

Generating and counting

σ permutation of $\{1, 2, \ldots n\}$.
Descent set : $D(\sigma) = \{i \in [1, n-1],\ \sigma(i) > \sigma(i+1)\}$.
Goal : generating at random, with uniform probability, and counting permutations with a given descent set A.

Counting

Inclusion-exclusion formula : number of permutations with descent set A is

$$\sum_{B=\{i_1<i_2\ldots<i_k\}\subset A^c} (-1)^{|A^c|-|B|} \binom{n}{n-i_k, i_k-i_{k-1}, \ldots, i_2-i_1, i_1}$$

Practical problem : number of computations to perform. If $n = 100$, $|A| = 50 \to 2^{50}$ terms in the summation.

Generating and counting

σ permutation of $\{1, 2, \dots n\}$.
Descent set : $D(\sigma) = \{i \in [1, n-1],\ \sigma(i) > \sigma(i+1)\}$.
Goal : generating at random, with uniform probability, and counting permutations with a given descent set A.

Counting

Inclusion-exclusion formula : number of permutations with descent set A is

$$\sum_{B=\{i_1<i_2\dots<i_k\}\subset A^c} (-1)^{|A^c|-|B|} \binom{n}{n-i_k, i_k-i_{k-1}, \dots, i_2-i_1, i_1}$$

Practical problem : number of computations to perform. If $n = 100$, $|A| = 50 \to 2^{50}$ terms in the summation.
More theoretical point of view : comparing the number of permutations with given descent sets A and A' (De Bruijn, Ehrenborg et al).

Generating and counting

Alternating permutations (André). Exponential generating function

$$\sum_n \frac{a_n}{n!} t^n = \sec t + \tan t$$

Other formulae for periodic patterns when there is exactly one descent in each period (Carlitz, Remmel).

Generating and counting

Alternating permutations (André). Exponential generating function

$$\sum_n \frac{a_n}{n!} t^n = \sec t + \tan t$$

Other formulae for periodic patterns when there is exactly one descent in each period (Carlitz, Remmel).
Other periodic patterns ? Asymptotics computed by Leeming-MacLeod, Bender et al.

Generating and counting

Alternating permutations (André). Exponential generating function

$$\sum_n \frac{a_n}{n!} t^n = \sec t + \tan t$$

Other formulae for periodic patterns when there is exactly one descent in each period (Carlitz, Remmel).
Other periodic patterns ? Asymptotics computed by Leeming-MacLeod, Bender et al.

Sampling

Number of alternating permutations of length n

$$\sim \frac{4}{\pi} \left(\frac{2}{\pi}\right)^{n+1} n!$$

The rejection algorithm has exponential complexity $\sim \frac{\pi}{4} \left(\frac{\pi}{2}\right)^{n+1}$

Continous space

Idea : construct a random sequence of reals $Y_i, 1 \leq i \leq n$, with the given descent profile. Recover the permutation from the (Y_i) by a sorting algorithm.

Continous space

Idea : construct a random sequence of reals $Y_i, 1 \leq i \leq n$, with the given descent profile. Recover the permutation from the (Y_i) by a sorting algorithm.

The sequence is defined by induction. We use the notion of *density* of a random variable.

$f : \mathbb{R} \to \mathbb{R}_+$ is the density of $X \in \mathbb{R}$ if, for all reals $a \leq b$,

$$\mathbb{P}(X \in [a, b]) = \int_a^b f(x)dx$$

Continous space

Idea : construct a random sequence of reals $Y_i, 1 \leq i \leq n$, with the given descent profile. Recover the permutation from the (Y_i) by a sorting algorithm.

The sequence is defined by induction. We use the notion of *density* of a random variable.

$f : \mathbb{R} \to \mathbb{R}_+$ is the density of $X \in \mathbb{R}$ if, for all reals $a \leq b$,

$$\mathbb{P}(X \in [a, b]) = \int_a^b f(x)dx$$

More generally, $f : \mathbb{R}^d \to \mathbb{R}_+$ is the density of $X \in \mathbb{R}^d$ if, for all Borel sets B,

$$\mathbb{P}(X \in B) = \int_B f(x)dx$$

A sequence of polynomials

- $f_n(x) = 1$
- If $i \in A$, $f_i(x) = \int_0^x f_{i+1}(y)dy$
- If $i \notin A$, $f_i(x) = \int_x^1 f_{i+1}(y)dy$

Number of computations : $O(n^2)$

A sequence of polynomials

- $f_n(x) = 1$
- If $i \in A$, $f_i(x) = \int_0^x f_{i+1}(y)dy$
- If $i \notin A$, $f_i(x) = \int_x^1 f_{i+1}(y)dy$

Number of computations : $O(n^2)$

Theorem

(i) The number of permutations with descent set A is

$$n! \int_0^1 f_1(x)dx$$

(ii) Let B be the set of local extrema of the permutation. Let $F(B)$ be the number of permutations with set of local extrema B. Then $F : \mathcal{P}(\{2, 3, \dots n-1\} \to \mathbb{N}$ is an increasing function.

Random sampling

Let $U_i, 0 \leq i \leq n$ be iid, uniform on $[0, 1]$.

- Y_1 has density proportional to f_1 :

$$U_1 \int_0^1 f_1(y)dy = \int_0^{Y_1} f_1(y)dy$$

If $U_1 = 0$, then $Y_1 = 0$. If $U_1 = 1$, then $Y_1 = 1$.
In the general case, $Y_1 \in [0, 1]$.

Random sampling

Let $U_i, 0 \leq i \leq n$ be iid, uniform on $[0,1]$.

- Y_1 has density proportional to f_1 :

$$U_1 \int_0^1 f_1(y)dy = \int_0^{Y_1} f_1(y)dy$$

  If $U_1 = 0$, then $Y_1 = 0$. If $U_1 = 1$, then $Y_1 = 1$.
  In the general case, $Y_1 \in [0,1]$.

- Y_{i+1} has conditional density proportional to f_{i+1} :
- If $i \in A$, recall that $f_i(x) = \int_0^x f_{i+1}(y)dy$. Then

$$U_{i+1} f_i(Y_i) = \int_0^{Y_{i+1}} f_{i+1}(y)dy$$

  If $U_{i+1} = 0$, then $Y_{i+1} = 0$. If $U_{i+1} = 1$, then $Y_{i+1} = Y_i$.
  In the general case, $Y_{i+1} \in [0, Y_i]$.

Random sampling

Let $U_i, 0 \leq i \leq n$ be iid, uniform on $[0,1]$.

- Y_1 has density proportional to f_1 :

$$U_1 \int_0^1 f_1(y)dy = \int_0^{Y_1} f_1(y)dy$$

  If $U_1 = 0$, then $Y_1 = 0$. If $U_1 = 1$, then $Y_1 = 1$.
  In the general case, $Y_1 \in [0,1]$.

- Y_{i+1} has conditional density proportional to f_{i+1} :
- If $i \in A$, recall that $f_i(x) = \int_0^x f_{i+1}(y)dy$. Then

$$U_{i+1} f_i(Y_i) = \int_0^{Y_{i+1}} f_{i+1}(y)dy$$

  If $U_{i+1} = 0$, then $Y_{i+1} = 0$. If $U_{i+1} = 1$, then $Y_{i+1} = Y_i$.
  In the general case, $Y_{i+1} \in [0, Y_i]$.

- If $i \notin A$, $U_{i+1} f_i(Y_i) = \int_{Y_{i+1}}^1 f_{i+1}(y)dy$.

The algorithm

Sektch of the proof

The density of Y_1 is $c_1 f_1$, where

$$c_1 = \frac{1}{\int_0^1 f_1(x)dx}$$

Sektch of the proof

The density of Y_1 is $c_1 f_1$, where

$$c_1 = \frac{1}{\int_0^1 f_1(x)dx}$$

Suppose for instance that 1 is a descent.
The density of Y_2, conditional on Y_1, is $c_2 f_2 \mathbf{1}_{\{Y_2 \leq Y_1\}}$.
Since $f_1(x) = \int_0^x f_2(y)dy$,

$$c_2 = \frac{1}{f_1(Y_1)}$$

Sektch of the proof

The density of Y_1 is $c_1 f_1$, where

$$c_1 = \frac{1}{\int_0^1 f_1(x)dx}$$

Suppose for instance that 1 is a descent.
The density of Y_2, conditional on Y_1, is $c_2 f_2 \mathbf{1}_{\{Y_2 \leq Y_1\}}$.
Since $f_1(x) = \int_0^x f_2(y)dy$,

$$c_2 = \frac{1}{f_1(Y_1)}$$

By independence, the density of the sequence $(Y_1, \ldots Y_n)$ is the product

$$c_1 f_1(Y_1) \frac{f_2(Y_2)}{f_1(Y_1)} \cdots \frac{f_n(Y_n)}{f_{n-1}(Y_{n-1})} \times \mathbf{1}_{\{D(Y)=A\}} = c_1 \mathbf{1}_{\{D(Y)=A\}}$$

Periodic descent set

Define f_n by increasing induction. Let $F(x, y) = \sum_n y^n f_n(x)$. Then

$$\sum_{n \geq 0} \frac{A_n}{n!} y^n = 1 + \int_0^1 yF(x, y)dx$$

A_n : number of permutations of length n with the desired pattern.

Periodic descent set

Define f_n by increasing induction. Let $F(x,y) = \sum_n y^n f_n(x)$. Then

$$\sum_{n\geq 0} \frac{A_n}{n!} y^n = 1 + \int_0^1 yF(x,y)dx$$

A_n : number of permutations of length n with the desired pattern.
If n is a descent (resp. an ascent), $f'_{n+1} = f_n$ (resp $f'_{n+1} = -f_n$).

$$\frac{\partial^p F(x,y)}{\partial x^p} = (-1)^k y^p F(x,y)$$

where p is the period and k the number of ascents in a period.

Periodic descent set

Define f_n by increasing induction. Let $F(x,y) = \sum_n y^n f_n(x)$. Then

$$\sum_{n \geq 0} \frac{A_n}{n!} y^n = 1 + \int_0^1 yF(x,y)dx$$

A_n : number of permutations of length n with the desired pattern.
If n is a descent (resp. an ascent), $f'_{n+1} = f_n$ (resp $f'_{n+1} = -f_n$).

$$\frac{\partial^p F(x,y)}{\partial x^p} = (-1)^k y^p F(x,y)$$

where p is the period and k the number of ascents in a period.

$$F(x,y) = \sum_{k=1}^{p} a_k(y) e^{\omega_k yx}$$

where the ω_k are the p-th roots of ± 1 and the a_k are power series : $a_k(y) = \sum_{n \geq 0} a_{k,n} y^n$.

Periodic descent set

- $A_n = (a_{1,n}, \ldots, a_{p,n})$. If $n-1$ is a descent, $A_n = 0$.
- V : Vandermonde matrix with i-th row $(\omega_1^{i-1}, \ldots, \omega_p^{i-1})$.
- U_n matrix with first column $(u_1^{(n)}, \ldots, u_p^{(n)})$ where $u_1^{(n)} = 1$ and $u_{j+1}^{(n)} = u_j^{(n)}$ (resp. $u_{j+1}^{(n)} = -u_j^{(n)}$) if $n+j-1$ is a descent (resp. an ascent). The other columns of U_n are zero.
- $R_n = (\rho_n, 0, \ldots, 0)$ with $\rho_0 = 1$ and if $n-1$ is an ascent,

$$\rho_n = -\sum_{j=1}^{n} \sum_{k=1}^{p} \frac{\omega_k^j a_{k,n-j}}{j!}$$

$$A_n = V^{-1} U_n R_n$$