Fisher-Yates

Every possible ordering of $\mathcal{P}$ is equally possible.

In place (no extra memory), $O(n)$, requires indexed (random) access.

```
# An unbiased shuffle of a population |𝒫| = n in place,
# given 𝒫 permits random access.

i ← 0
while ( i < n−1 ) do (
    j ← Equilikely(i,n-1)
    swap 𝒫[i] and 𝒫[j]
    increment i
)
# 𝒫 is now shuffled
```

Fisher-Yates Proof

Given population $\mathcal{P} = \{a_1, a_2, a_3, \ldots, a_n\}$, what is the probability that the original element a_i ends up after the shuffle at location $1 \leq k \leq n$?

$$
\begin{aligned}
\Pr(a_i \ @ \ k) &= \Pr(a_i \text{ not chosen for locations} < k) \cdot \Pr(a_i \text{ chosen for location } k) \\
&= \left(1 - \frac{1}{n}\right)^{\triangle} \left(1 - \frac{1}{n-1}\right)^{\odot} \left(1 - \frac{1}{n-2}\right) \cdots \left(1 - \frac{1}{n-(k-2)}\right) \left(\frac{1}{n-(k-1)}\right)^{*}
\end{aligned}
$$

$\triangle$ Choose one out of n (with probability $\frac{1}{n}$) for the first location,

$\odot$ Choose one out of $n-1$ (with probability $\frac{1}{n-1}$) for the second location,

$\vdots$

$*$ If $k = n$, no choice is left for the last element (probability $\left.\frac{1}{n-(k-1)}\right|_{k=n} = 1$).
Which is why the algorithm rearranges just the elements a_1 through a_{n-1}.

Fisher-Yates Proof

Given population $\mathcal{P} = \{a_1, a_2, a_3, \ldots, a_n\}$, what is the probability that the original element a_i ends up after the shuffle at location $1 \leq k \leq n$?

$$
\begin{aligned}
\Pr(a_i \ @ \ k) &= \Pr(a_i \text{ not chosen for locations} < k) \cdot \Pr(a_i \text{ chosen for location } k) \\
&= \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{n-2}\right)\cdots\left(1 - \frac{1}{n-(k-2)}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{n-1}{n}\right)\left(\frac{n-2}{n-1}\right)\left(\frac{n-3}{n-2}\right)\cdots\left(\frac{n-k+2-1}{n-k+2}\right)\left(\frac{1}{n-(k-1)}\right)
\end{aligned}
$$

Fisher-Yates Proof

Given population $\mathcal{P} = \{a_1, a_2, a_3, \ldots, a_n\}$, what is the probability that the original element a_i ends up after the shuffle at location $1 \leq k \leq n$?

$$
\begin{aligned}
\Pr(a_i \text{ @ } k) &= \Pr(a_i \text{ not chosen for locations} < k) \cdot \Pr(a_i \text{ chosen for location } k) \\
&= \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{n-2}\right)\cdots\left(1 - \frac{1}{n-(k-2)}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{n-1}{n}\right)\left(\frac{n-2}{n-1}\right)\left(\frac{n-3}{n-2}\right)\cdots\left(\frac{n-k+2-1}{n-k+2}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{\cancel{n-1}}{n}\right)\left(\frac{n-2}{\cancel{n-1}}\right)\left(\frac{n-3}{n-2}\right)\cdots\left(\frac{n-k+1}{n-k+2}\right)\left(\frac{1}{n-k+1}\right)
\end{aligned}
$$

Fisher-Yates Proof

Given population $\mathcal{P} = \{a_1, a_2, a_3, \ldots, a_n\}$, what is the probability that the original element a_i ends up after the shuffle at location $1 \leq k \leq n$?

$$
\begin{aligned}
\Pr(a_i \;@\; k) &= \Pr(a_i \text{ not chosen for locations} < k) \cdot \Pr(a_i \text{ chosen for location } k) \\
&= \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{n-2}\right)\cdots\left(1 - \frac{1}{n-(k-2)}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{n-1}{n}\right)\left(\frac{n-2}{n-1}\right)\left(\frac{n-3}{n-2}\right)\cdots\left(\frac{n-k+2-1}{n-k+2}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{\cancel{n-1}}{n}\right)\left(\frac{n-2}{\cancel{n-1}}\right)\left(\frac{n-3}{n-2}\right)\cdots\left(\frac{n-k+1}{n-k+2}\right)\left(\frac{1}{n-k+1}\right) \\
&= \left(\frac{1}{n}\right)\left(\frac{\cancel{n-2}}{1}\right)\left(\frac{n-3}{\cancel{n-2}}\right)\cdots\left(\frac{n-k+1}{n-k+2}\right)\left(\frac{1}{n-k+1}\right)
\end{aligned}
$$

Fisher-Yates Proof

Given population $\mathcal{P} = \{a_1, a_2, a_3, \ldots, a_n\}$, what is the probability that the original element a_i ends up after the shuffle at location $1 \leq k \leq n$?

$$
\begin{aligned}
\Pr(a_i \mathbin{@} k) &= \Pr(a_i \text{ not chosen for locations} < k) \cdot \Pr(a_i \text{ chosen for location } k) \\
&= \left(1 - \frac{1}{n}\right)\left(1 - \frac{1}{n-1}\right)\left(1 - \frac{1}{n-2}\right) \cdots \left(1 - \frac{1}{n-(k-2)}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{n-1}{n}\right)\left(\frac{n-2}{n-1}\right)\left(\frac{n-3}{n-2}\right) \cdots \left(\frac{n-k+2-1}{n-k+2}\right)\left(\frac{1}{n-(k-1)}\right) \\
&= \left(\frac{\cancel{n-1}}{n}\right)\left(\frac{n-2}{\cancel{n-1}}\right)\left(\frac{n-3}{n-2}\right) \cdots \left(\frac{n-k+1}{n-k+2}\right)\left(\frac{1}{n-k+1}\right) \\
&= \left(\frac{1}{n}\right)\left(\frac{\cancel{n-2}}{1}\right)\left(\frac{n-3}{\cancel{n-2}}\right) \cdots \left(\frac{n-k+1}{n-k+2}\right)\left(\frac{1}{n-k+1}\right) \\
&= \left(\frac{1}{n}\right)\left(\frac{1}{1}\right)\left(\frac{\cancel{n-3}}{1}\right) \cdots \left(\frac{\cancel{n-k+1}}{\cancel{n-k+2}}\right)\left(\frac{1}{\cancel{n-k+1}}\right) = \frac{1}{n} \qquad \square
\end{aligned}
$$

IOW: Every element has an equal probability of landing at each location after the shuffle.