

Algoritmos Genéticos Y Optimización Heurística

Dr. Adrian Will
Grupo de Aplicaciones de Inteligencia Artificial
Universidad Nacional de Tucumán
awill@herrera.unt.edu.ar

Algoritmos Genéticos - Operadores de Selección

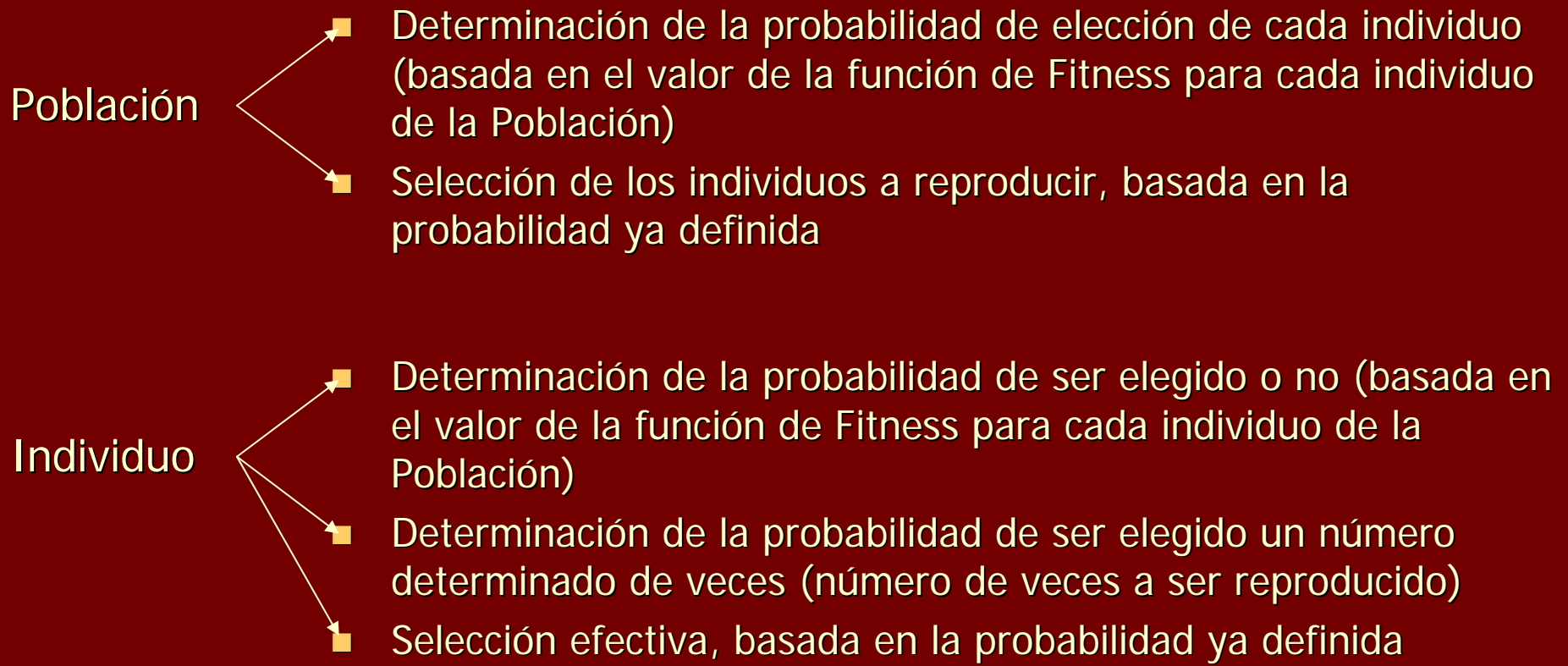
Selección: Proceso por el cual se determina la cantidad de veces que cada individuo será elegido para reproducción, y/o para pasar a la generación siguiente (dependiendo de la Arquitectura elegida).
Similarmente, operador que selecciona los individuos a reproducir.

Este proceso se realiza Con Repetición, es decir, cada individuo de la población puede ser elegido más de una vez para reproducirse.

Este proceso debe basarse en la Función de Evaluación o Fitness del problema, de manera tal de elegir mayor cantidad de veces a los individuos mejor adaptados (con mayor valor de Fitness), pero dar probabilidad de ser elegidos a todos los individuos

Algoritmos Genéticos - Operadores de Selección

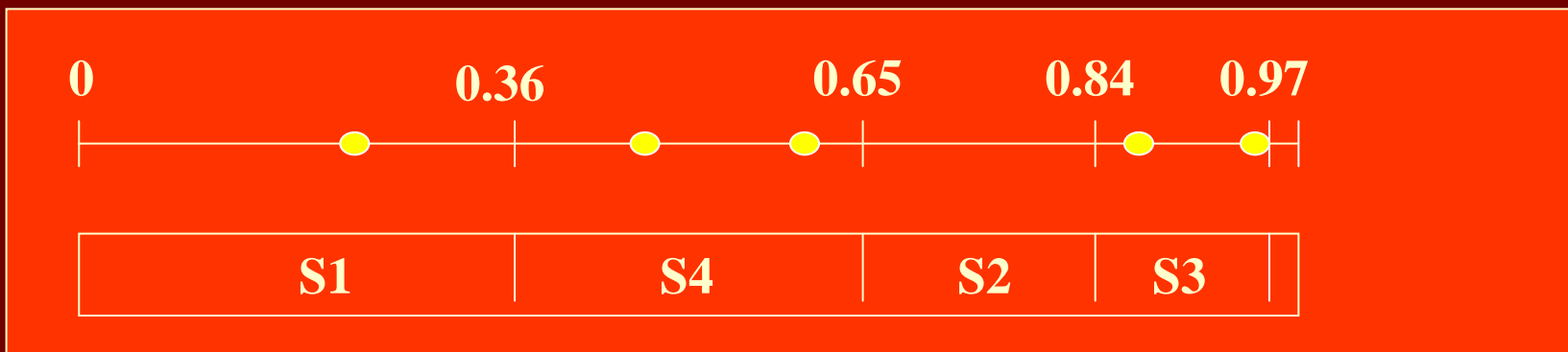
El proceso de Selección puede verse desde 2 puntos de vista, y en cada caso, puede separarse en 2 procesos distintos:



Algoritmos Genéticos – Operadores de Selección

En general, la probabilidad un individuo en particular de ser elegido, no coincide con el número de veces que el individuo es efectivamente elegido. O sea, individuos con alto valor de Fitness son elegidos pocas veces, e individuos con valores menores de Fitness son elegidos una mayor cantidad de veces.

Esto es debido a que los generadores aleatorios no responden exactamente a la distribución propuesta, especialmente en números bajos (cuando el tamaño de la muestra aumenta, esta diferencia disminuye).



Algoritmos Genéticos – Operadores de Selección

Se han desarrollado 3 medidas de la eficiencia de un operador de Selección, que deben ser tenidas en cuenta a la hora de utilizarlo:

BIAS o Desviación

SPREAD o Dispersión

Efficiency (Eficiencia Computacional)

SELECTIVE PRESSURE

GENETIC DIVERSITY

Operadores de Selección – Bias- Definición

BIAS o Desviación:

- Diferencia en valor absoluto, entre la probabilidad de un individuo de ser seleccionado, y el número de veces que es efectivamente seleccionado (*trials*).
- Diferencia en valor absoluto, entre la probabilidad de un individuo de ser seleccionado, y el valor de la función de Fitness Normalizada para el individuo. O sea,

$$F_{Norm}(x) = \frac{F(x)}{1 + F(x)} \quad (\text{si } F > 0), \text{ ó}$$
$$F_{Norm} = \frac{F(x) - m}{M - m} \quad (\text{si } m \leq F \leq M)$$

De este modo, F_{Norm} proporciona un valor entre 0 y 1 proporcional al valor de la función de Fitness F para el individuo, valor que debería a priori coincidir con la probabilidad del individuo de ser elegido.

Operadores de Selección – Bias - Definición

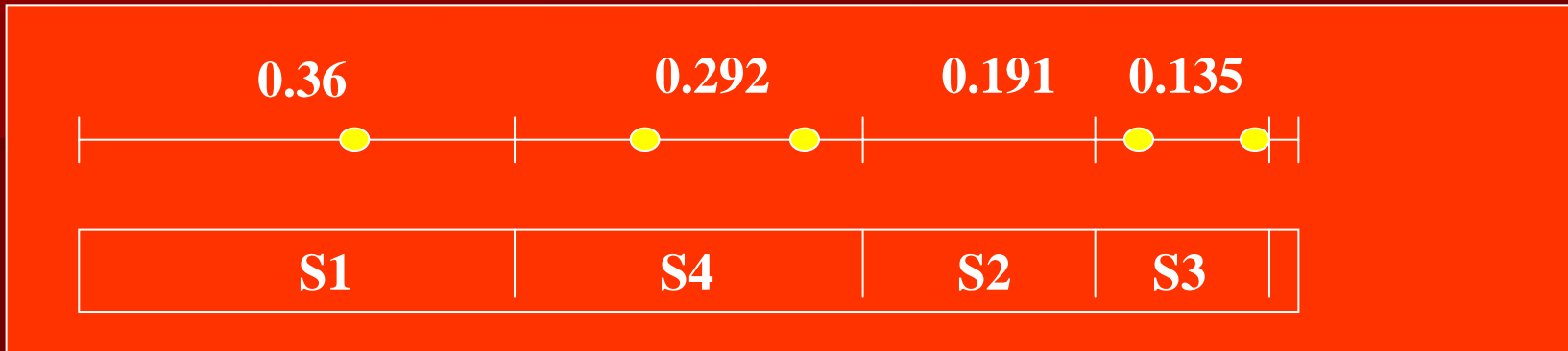
BIAS o Desviación:

- Diferencia en valor absoluto, entre la probabilidad de un individuo de ser seleccionado, y el valor de la función de Fitness Normalizada para el individuo F_{Norm} , o sea,

$$F_{Norm} : \text{Espacio de Soluciones} \xrightarrow{F} R \xrightarrow{G} [0,1]$$

donde F es la función de Fitness, G es una función de normalización, normalmente lineal o racional, de tal modo que F_{Norm} asigna a cada individuo un valor entre 0 y 1, proporcional al valor de la función original. Este proceso se denomina *Fitness Assignment*

Operadores de Selección – Bias - Ejemplo



$$S_1 \rightarrow 0.36 \rightarrow \frac{1}{5} = 0.2$$

$$S_4 \rightarrow 0.292 \rightarrow \frac{2}{5} = 0.4$$

$$S_2 \rightarrow 0.191 \rightarrow \frac{0}{5} = 0$$

$$S_3 \rightarrow 0.135 \rightarrow \frac{2}{5} = 0.4$$

$$S_5 \rightarrow 0.022 \rightarrow \frac{0}{5} = 0$$

Operadores de Selección – Bias - Características

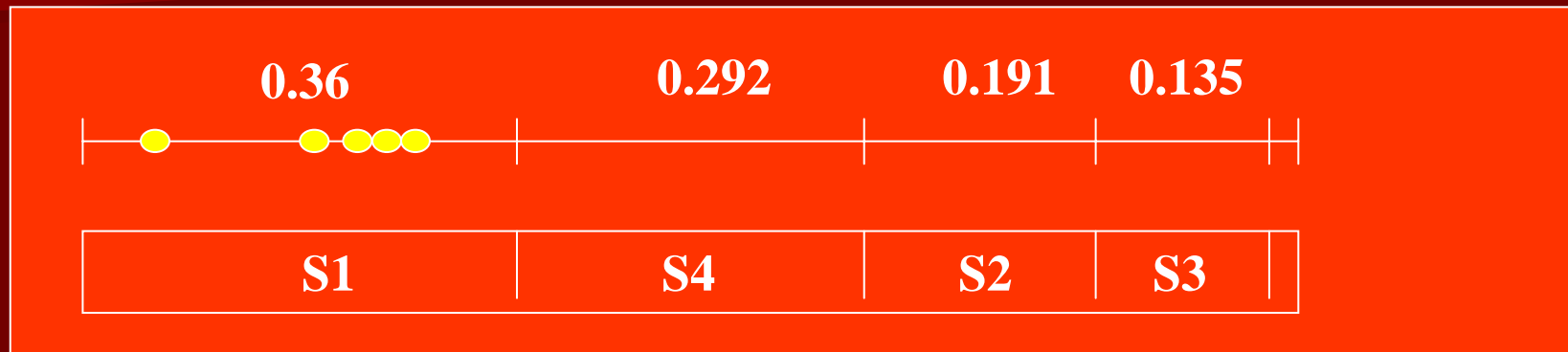
- El valor óptimo de Bias es $Bias = 0$, que se alcanza cuando la probabilidad de selección de los individuos coincide con las veces que el individuo es efectivamente elegido.
- Visto de otro modo, $Bias = 0$ cuando la probabilidad de selección de los individuos coincide con el número esperado de *trials* (expected number of trials) para el individuo y el operador seleccionados. O sea, la probabilidad calculada originalmente con la función de Fitness Normalizada, coincide con la probabilidad de ocurrencia del evento calculada para el operador en consideración.
- En general, el Bias cuantifica la tendencia del operador de no cumplir adecuadamente el procedimiento de la manera esperada.

Operadores de Selección – Spread - Definición

SPREAD o Dispersión:

- Rango en el espacio de Soluciones, que un individuo es capaz de conseguir, en una sola aplicación del operador de selección utilizado. En particular, dada la aleatoriedad involucrada en este tipo de procesos, representa la medida del “peor caso posible” (*Worst Case Scenario*)
- $\text{Spread} = [\min(V_i), \max(V_i)]$, donde V_i es el número de veces que el individuo i puede ser efectivamente elegido utilizando el operador seleccionado
- El Spread representa entonces el mayor número posible de veces que un individuo puede llegar a reproducirse. En particular, mide la mayor cantidad de veces que un *Buen* individuo puede conseguir, y por lo tanto, la máxima velocidad de *Dispersión* de una característica genética a través de la población

Operadores de Selección – Spread - Ejemplo



- Por ejemplo en el caso de la ruleta, el spread es el máximo posible $[0, n]$ (n = número de ensayos o trials), ya que es perfectamente posible (aunque poco probable) que un individuo no sea elegido nunca, o sea elegido todas las veces.

Operadores de Selección – Spread

- Spread → Velocidad de Dispersión de características genéticas
- En este caso, Velocidad de Dispersión de una buena solución a toda la población. En particular, tiempo que ésta tarda en llenar toda la población con copias de sí misma. Este es un problema común en Algoritmos Genéticos llamado *Crowding o Convergencia Prematura*
- En general, muchas de las estrategias y operadores que se utilizan, están diseñados específicamente para intentar evitar este problema, ya que la convergencia prematura en general ocurre cerca de óptimos locales, y evita la convergencia al óptimo global del problema. Se intenta entonces retrasar la convergencia, hasta asegurarse que el algoritmo ha tenido tiempo de explorar todo el espacio
- El Spread mínimo es entonces básico para que un Algoritmo Genético opere de manera eficiente, ya que en este caso una buena solución va a necesitar mayor número de generaciones para copiar la población.

Operadores de Selección – Eficiencia

- **EFICIENCIA:** El proceso de Selección es una de las partes computacionalmente más costosas de un Algoritmo Genético. En general, el resto de los procesos es de orden $O(nL)$, donde n es la dimensión del espacio de búsqueda (número de variables), y L es la longitud del vector solución. Mantener la complejidad de esta parte del proceso al mínimo es entonces claramente básico.

En general, **la parte computacionalmente más costosa de un Algoritmo Genético, es el cálculo de la función de evaluación**, que se realiza por lo menos n veces por generación (en algunas arquitecturas puede ser más).

Es importante entonces **dedicar tiempo a producir un código Computacionalmente Eficiente** para estos operadores, más que para los de Crossover o Mutación, que son relativamente más simples y en general aplicados a porciones menores de la población.

Operadores de Selección – Otras Características

- **SELECTIVE PRESSURE:** (presión de selección) Probabilidad de elección del mejor individuo de ser seleccionado, en comparación con el promedio de probabilidad de elección de toda la población
- Número efectivo de veces que el mejor individuo de la población es efectivamente elegido, en comparación al promedio de la población.

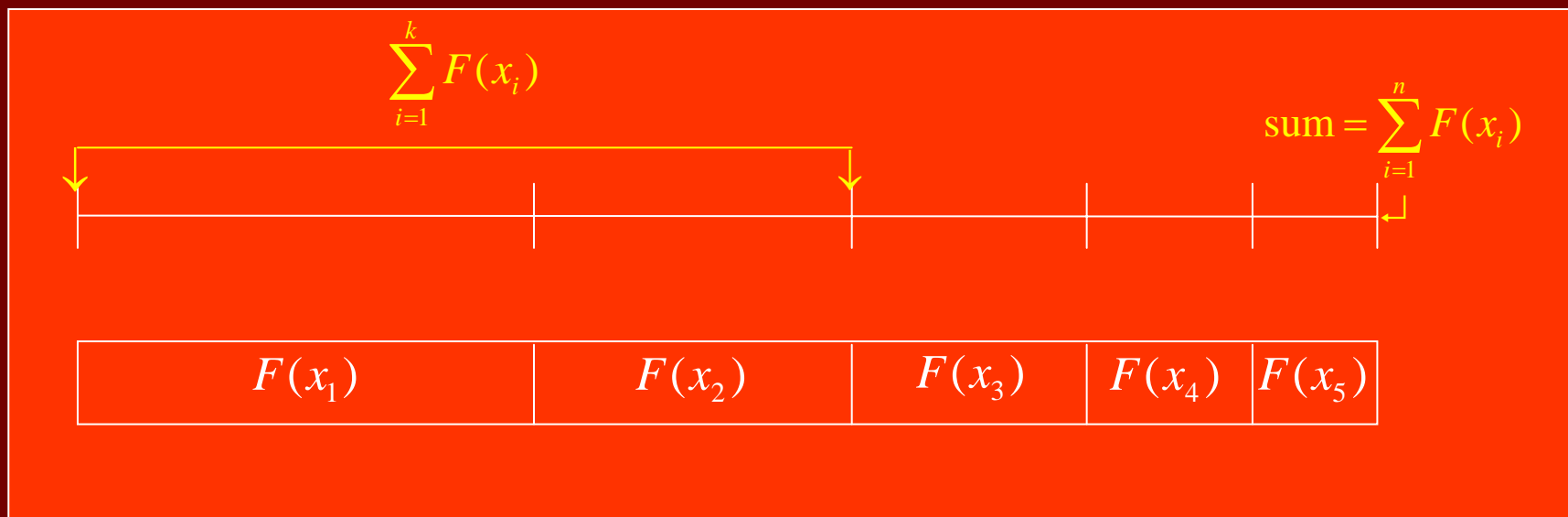
En general, se refiere a la preferencia que tiene el método de selección de elegir el mejor con preferencia a individuos promedio: cuando esta tendencia es más marcada, el mejor individuo será elegido una gran cantidad de veces en cada generación, copando rápidamente la población y produciendo *Crowding*. De ser la presión de selección excesivamente baja, individuos de baja calidad pueden reproducirse en exceso y retrasar la convergencia.

Operadores de Selección – Otras Características

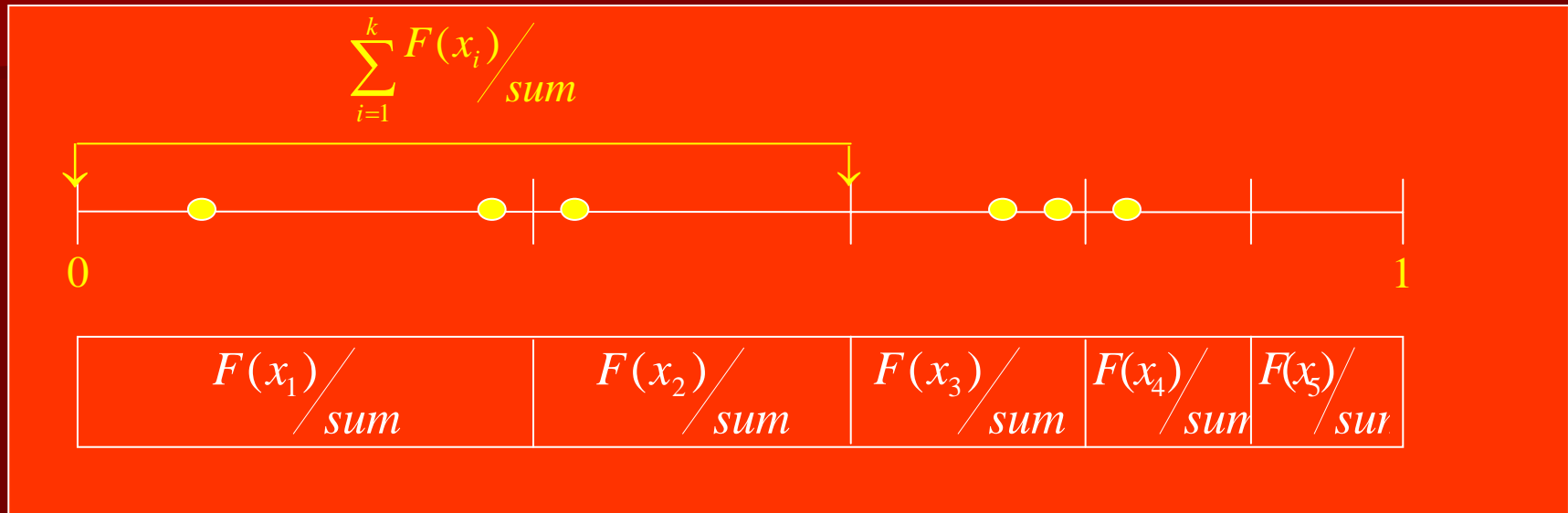
- **Genetic Diversity:** (Diversidad Genética) Proporción de la población que no son copias uno de otro. Específicamente, se refiere a que la población no ha llegado al punto del Crowding
- **Loss of Genetic Diversity:** Proporción de la población que no pasa sus características genéticas a la siguiente generación, por diversos motivos (exceso de mutación, Crowding, excesiva presión de Selección (Selective Pressure), etc.

Operadores de Selección – Ruleta

- **Ruleta con pesos o Stochastic Sampling With Replacement (SSWR):**
Stochastic Sampling = Muestreo Estocástico (al azar), With replacement por la posibilidad de elegir el mismo valor más de una vez (similar a extraer bolitas de color de una caja)



Operadores de Selección – Ruleta



donde $sum = \sum_{i=1}^n F(x_i)$. Se tiran entonces números aleatorios, entre 0 y 1 en este caso, o entre 0 y sum en el caso anterior (computacionalmente es preferible el primero), utilizando $sum * rand$ (generando de a uno), o $sum.*rand(1,n)$ en forma vectorial.

Operadores de Selección – Ruleta

Stochastic Sampling With Replacement (SSWR) – Ruleta

- **Bias:** En general **grande**, debido a errores del generador aleatorio, especialmente en baja cantidad de tiros. Sin embargo aunque se note este efecto, puede funcionar de manera eficiente en grandes poblaciones (más de 100 individuos)
- **Spread:** $[0, \text{número de Tiros}]$ (**el máximo posible**), ya que un dado subintervalo puede tanto ser evitado por el generador aleatorio, como ser elegido todas las veces, aunque esto sea poco probable. De este modo se llenaría la población con copias de un mismo individuo, en una sola generación.
- **Eficiencia:** Se debe generar un número aleatorio por cada tiro que se efectúe. $O(n \log(n))$.

Operadores de Selección – Ruleta Modificada

Stochastic Sampling With Partial Replacement (SSWPR)

- **With Replacement (Con reemplazo):** La probabilidad de cada individuo de ser elegido (subintervalos del $[0, 1]$ o de $[0, \text{sum}]$) no se modifica entre tiro y tiro: **Un individuo elegido tendrá exactamente la misma probabilidad de ser elegido en el siguiente tiro.** Corresponde a volver a poner las bolitas en el sombrero.
- **Without Replacement (Sin reemplazo):**, La probabilidad de un individuo de ser elegido se modifica entre tiro y tiro: **un individuo elegido no puede volver a ser elegido en lo sucesivo** (y tengo por lo tanto un máximo de tiros posibles, tantos como individuos en la población). Corresponde a que el subintervalo correspondiente a la solución es eliminado, o a que una vez elegida una bolita de un color, todas las bolitas de ese color son sacadas del sombrero y descartadas.

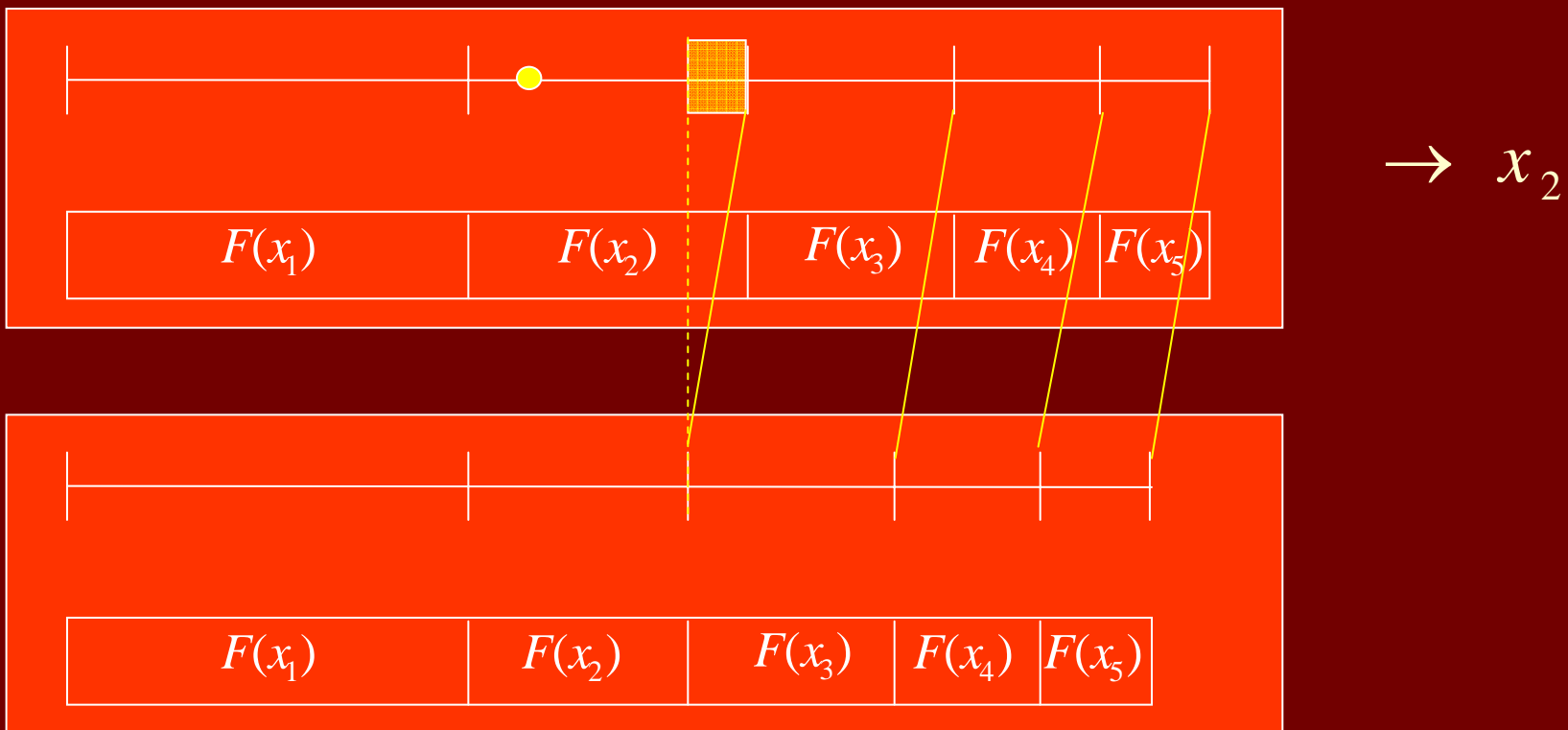
Operadores de Selección – Ruleta Modificada

Stochastic Sampling With Partial Replacement (SSWPR)

- **With Partial Replacement (Con reemplazo parcial):** La probabilidad de un individuo de ser elegido se modifica entre tiro y tiro: **un individuo elegido puede volver a ser elegido en lo sucesivo, pero la probabilidad disminuye en una cantidad fija** cada vez que la elección ocurre (y nuevamente existe un máximo de tiros posibles, ya que eventualmente llega a cero). Corresponde a que el subintervalo correspondiente a la solución es achicado en tamaño, o a que las bolitas sacadas del sombrero son descartadas (pero quedan bolitas del mismo color en el sombrero).
- En el caso de los segmentos, el subintervalo correspondiente a la solución elegida es achicado en tamaño una cantidad fija, antes de tirar de nuevo.

Operadores de Selección – Ruleta Modificada

- En el caso de los segmentos, el subintervalo correspondiente a la solución elegida es achicado en tamaño una cantidad fija, antes de tirar de nuevo.



Probabilidad de reelección de x_2 disminuida en ε

Operadores de Selección – Ruleta Modificada

Stochastic Sampling With Partial Replacement (SSWPR)

- **Bias: Mayor que SSWR (ruleta)**, dado que, a pesar de que el valor de la función de Fitness para el individuo, y por lo tanto la probabilidad original de selección, no se modifican entre tiro y tiro, la probabilidad efectiva de elección va disminuyendo, y por lo tanto es progresivamente más difícil acertarle a esa solución al azar (o sea, la diferencia entre probabilidad y elección efectiva será grande).
- **Spread: Limitado y menor que SSWR**, ya que un dado subintervalo no puede ser elegido demasiadas veces.
- **Eficiencia: Similar a SSWR**

Menor Spread pero Mayor Bias que Ruleta

Operadores de Selección – Remainder Selection Methods

Remainder Selection Methods – Métodos de Selección por Restos

Clase de métodos de Selección, que utilizan otro de los métodos mencionados como parte integral de su algoritmo.

El algoritmo general es el siguiente:

1. Asignar probabilidades a los individuos, entre 0 y sum
2. Elegir Determinísticamente los individuos a reproducir, por parte entera de la probabilidad
3. En general no se alcanza el número requerido de individuos a seleccionar (por problemas de decimales). Se utiliza entonces otro método, como Ruleta o Ruleta modificada, para completar los tiros faltantes, decidiendo entre los individuos con parte fraccionaria no nula.

Remainder Selection Methods - Ejemplo

$n = 10$ Tiros

$$S_1 \rightarrow 0.48 \quad [p_1] = [10 * 0.48] = [4.8] = 4$$

$$S_2 \rightarrow 0.33 \quad [p_2] = [10 * 0.33] = [3.3] = 3$$

$$S_3 \rightarrow 0.12 \quad [p_3] = [10 * 0.12] = [1.2] = 1$$

$$S_4 \rightarrow 0.07 \quad [p_4] = [10 * 0.07] = [0.7] = 0$$

8 Tiros Determinísticamente asignados
Se completan los 2 faltantes con otro método

Operadores de Selección – Remainder Selection Methods

Remainder Selection Methods – Métodos de Selección por Restos

- **Bias: Cero o casi cero**, debido a que la parte determinística del proceso asegura que todos los individuos van a cumplir con una proporción grande de la probabilidad asignada.
- **Spread: Acotado Inferiormente, pero No Superiormente**. O sea, el método asegura que cada individuo será elegido $[np_i]$ veces, lo que asegura una cota inferior al spread (el individuo será elegido como mínimo ese número de veces). Pero cualquier individuo con parte fraccionaria no nula, puede ser elegido todas las veces, de acuerdo al método utilizado para esta parte.
- **Eficiencia**: Alta, debido a la simplicidad de la parte determinística del método, y de la baja cantidad de individuos a los que se debe aplicar el método accesorio.

Operadores de Selección – Remainder Selection Methods

Remainder Selection Methods – Métodos de Selección por Restos

- **Remainder Stochastic Sampling With Replacement (RSSR)** (Método por restos asociado al método de la Ruleta). Se completa el método por restos con el método de la Ruleta. Cualquier individuo con parte fraccional no nula puede ganar todas las veces, pero el número de copias que consigue está limitado por el número de individuos no asignados determinísticamente, que suele ser bajo.
- **Remainder Stochastic Sampling Without Replacement (RSSWR)** Se complementa el Método de los restos por el método de Ruleta sin reemplazo. O sea, cualquier elemento elegido en la parte probabilística del método, es eliminado para los posteriores tiros. Esto limita el Spread, pero aumenta la probabilidad de las fracciones pequeñas de ser elegidas (las grandes son rápidamente eliminadas)

Mínimo Spread, pero Bias no nulo

Selección – Stochastic Universal Sampling

1 – Elijo Un Solo número al azar, en el intervalo $[0, 1/n]$, n = número de tiros (orden $\text{rand} * \text{sum}/n$)

2 – Divido el intervalo total según las probabilidades asignadas

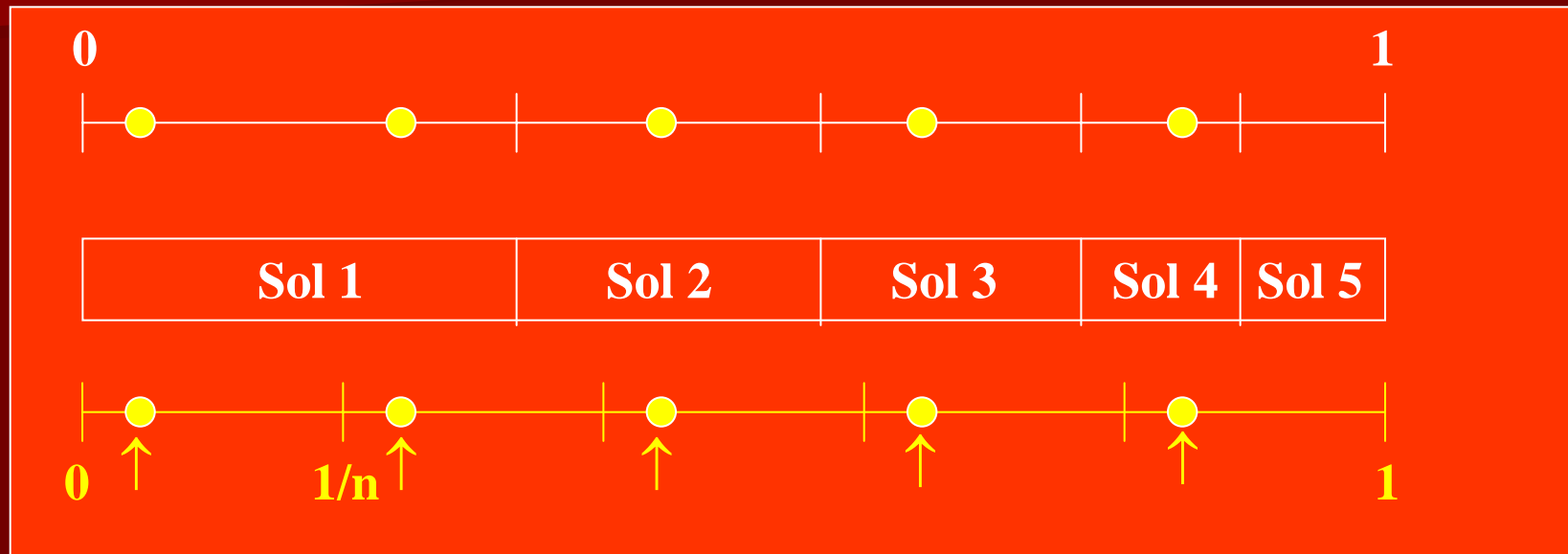
3 – Genero n puntos para elección, $\left\{ ptr, ptr + \frac{1}{n}, ptr + \frac{2}{n}, \dots, ptr + \frac{n-1}{n} \right\}$

$$\begin{aligned} \text{Resulta que } 0 \leq ptr \leq ptr + \frac{1}{n} \leq ptr + \frac{2}{n} \leq \dots \leq ptr + \frac{n-1}{n} \\ \leq \frac{1}{n} + \frac{n-1}{n} \\ = \frac{n}{n} = 1 \end{aligned}$$

por lo que se han generado n puntos equiespaciados, al azar, entre 0 y 1

4 – Elijo según los subintervalos indicados por estos puntos

Selección – Stochastic Universal Sampling



$$\left\{ ptr, ptr + \frac{1}{n}, ptr + \frac{2}{n}, \dots, ptr + \frac{n-1}{n} \right\}$$

Operadores de Selección – Stochastic Universal Sampling

Stochastic Universal Sampling (SUS)

- **Bias: Cero**, debido a que la parte determinística del proceso asegura que todos los individuos van a cumplir con la probabilidad asignada con bastante aproximación, incluso en poblaciones pequeñas.
- **Spread: $[\text{floor}(np_i), \text{ceil}(np_i)] = [[np_i], [np_i] + 1]$ (mínimo)**. El método asegura que cada individuo será elegido como mínimo $[np_i]$, y no más de $[np_i] + 1$ veces, y no hay mala suerte posible que aumente este spread.
- **Eficiencia: Alta**, debido a la simplicidad de la parte determinística del método, y de que se tiene que generar sólo un número al azar, no importa la cantidad de tiros a efectuar. $O(n)$.

Cero Bias, Mínimo Spread, Altamente Eficiente