

# Trabajo practico 1

## The Multiarmed Bandit

Integrantes: Tomas Anderson, Alejo Vaschetti, Manuel Moresi

Profesores: Tomas Capretto, Ignacio Evangelista Fecha de entrega: 17/04/2024

## Introduccion

To do

## Desarrollo

### Situacion ideal

La situacion ideal seria conocer de antemano cual maquina es la que tiene la mayor probabilidad de ganar y jugar siempre en esta. Para que una estrategia sea buena, tiene que asemejarse a dicha situacion.

Para ver la cantidad de dinero que gana en promedio este “metodo” se simulan 1000 años utilizandolo. Estos fueron los resultados:

/histograma/

En promedio, siguiendo este metodo, se espera que ganen “valor” unidades monetarias en un año.

El 95% de las ganancias al final del año estuvieron entre “valor” y “valor”.

## Estrategia 1: Completamente al azar

Esta estrategia consiste en elegir al azar la maquina a jugar en el dia, teniendo las tres la misma probabilidad de ser elegidas. Esta es la estrategia mas basica, ya que no se actualiza ni requiere ningun tipo de informacion previa.

Se juega un año con esta estrategia y se obtienen los siguientes resultados:

/3 lineas/

Se jugo mas en la maquina 2, pero esto se debe a puro azar y no una decision meditada de la estrategia

/1 linea/

Las ganancias son bastante lineales

/posterior 1/ /posterior 2/ /posterior 3/

A traves del paso de los dias, el centro de las distribuciones a posteriori de las maquinas 1, 2 y 3 se van acercando a los valores “valor 1”, “valor 2” y “valor 3”

Para ver la cantidad de dinero que gana en promedio este “metodo” se simulan 1000 años utilizandolo. Estos fueron los resultados:

/histograma/

En promedio, siguiendo este metodo, se espera que ganen “valor” unidades monetarias en un año.

El 95% de las ganancias al final del año estuvieron entre “valor” y “valor”.

## Estrategia 2: Greedy con tasa observada

Esta estrategia seleccion la maquina a jugar en el dia a aquella que tenga la mayor probabilidad de ganar observada hasta el momento. Como al empezar el año no se tiene alguna informacion sobre la probabilidad de ganar de cada maquina, se decide jugar cada maquina por separado hasta que cada una tenga su primer exito.

Se juega un año con esta estrategia y se obtienen los siguientes resultados:

/3 lineas/

Se “jugo” casi exclusivamente la maquina “x” y se ignora a las otras 2

/1 linea/

Las ganancias son bastante lineales

/posterior 1/ /posterior 2/ /posterior 3/

A través del paso de los días, el centro de la distribución a posteriori de la máquina “x” se va acercando al valor “valor 1” y las otras máquinas tienen las mismas distribuciones a posteriori en todos los días observados porque desde antes del día 61 se dejaron de usar.

Para ver la cantidad de dinero que gana en promedio este “método” se simulan 1000 años utilizándolo. Estos fueron los resultados:

/histograma/

En promedio, siguiendo este método, se espera que ganen “valor” unidades monetarias en un año.

El 95% de las ganancias al final del año estuvieron entre “valor” y “valor”.