

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

Cosas que se corren siempre

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2     3.5.0     v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.1
v purrr       1.0.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(ggplot2)
library(gridExtra)
```

Attaching package: 'gridExtra'

The following object is masked from 'package:dplyr':

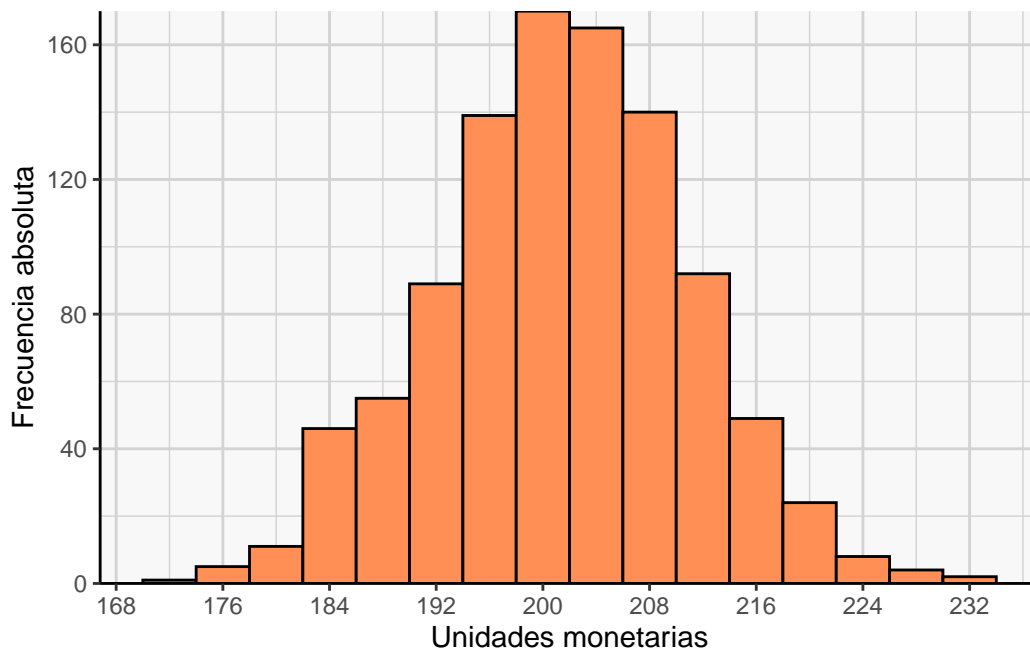
combine

```
source("Funciones.R")
prob=c(.3,.55,.45)
tema = theme(panel.grid = element_line(color = "lightgrey"),
              panel.background = element_rect(fill = "#f8f8f8"),
              axis.line = element_line(color = "black"))
```

Situación ideal

La situación ideal sería conocer de antemano cual máquina es la que tiene la mayor probabilidad de ganar y jugar siempre con ella. Esta sería la mejor forma de ganar dinero, por lo que una buena estrategia tiene que asemejarse a dicha situación.

Para ver la cantidad de dinero que gana en promedio este método se simulan a 1000 personas jugando 1 año utilizándolo. Estos fueron los resultados:



/histograma/

En este modo de operación una persona que juegue un año se espera que va a ganar alrededor de 201.986 unidades monetarias. En el 95% de los años simulados las ganancias totales se encontraron entre 184 y 219 unidades monetarias. Dicho intervalo, y los que serán presentados en las próximas estrategias, se encuentran entre el 2,5% y el 97,5% de las ganancias totales del método ordenadas de menor a mayor.

Para ver que tan beneficioso es cada estrategia, se comparan la cantidad de veces que se gana en promedio esperando que se acerque lo máximo posible a esta forma de jugar. Además, se contrastan los límites de los intervalos propuestos.

Estrategia 1: Completamente al azar

Esta estrategia consiste en elegir cada día al azar qué máquina jugar, teniendo las tres la misma probabilidad de ser selectas. Esta es la estrategia más básica, ya que no requiere ningún tipo

de informacion previa ni evoluciona en el tiempo.

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

```
set.seed(69)

datos_random_dia = actualizar(366, 1)

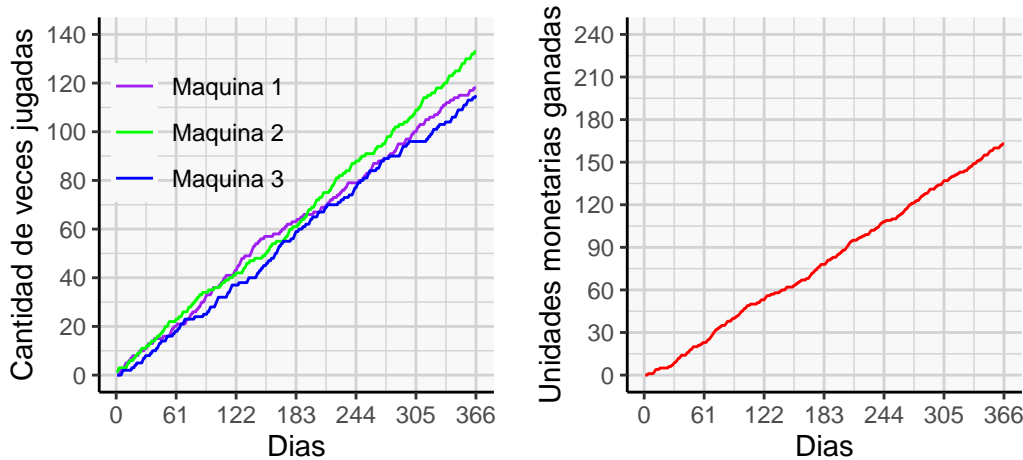
plot_1_random_dia = ggplot(datos_random_dia)+
  geom_line(aes(x = Dia, y = Maquina1Total, colour = "Maquina 1"))+
  geom_line(aes(x = Dia, y = Maquina2Total, colour = "Maquina 2"))+
  geom_line(aes(x = Dia, y = Maquina3Total, colour = "Maquina 3"))+
  tema+
  scale_y_continuous(name = "Cantidad de veces jugadas", breaks = c(seq(0, 140, 20), 366), limits = c(0, 366))+
  scale_x_continuous(name="Dias",breaks = seq(0, 366, 61)) +
  scale_color_discrete(breaks=c("Maquina 1", "Maquina 2", "Maquina 3"),type = c("purple", "green", "red"))+
  theme(plot.margin = margin(1.5,.3,1.5,.3, "cm"))+
  theme(legend.position = c(.25,.75),legend.background = element_blank())
```

Warning: A numeric `legend.position` argument in `theme()` was deprecated in ggplot2 3.5.0.

i Please use the `legend.position.inside` argument of `theme()` instead.

```
plot_2_random_dia = ggplot(datos_random_dia)+
  geom_line(aes(x = Dia, y = Maquina1 + Maquina2 + Maquina3), color = "red")+
  scale_y_continuous(name="Unidades monetarias ganadas", breaks = c(seq(0, 240, 30)),limits = c(0, 240))+
  scale_x_continuous(name="Dias", breaks = seq(0, 366, 61))+tema+
  theme(plot.margin = margin(1.5,.3,1.5,.3, "cm"))

grid.arrange(plot_1_random_dia, plot_2_random_dia, ncol=2)
```



/3 líneas/ /1 línea/

La cantidad de veces que se juegan las 3 máquinas son similares, y la diferencia entre ellas se debe puramente al azar. Este hecho impacta en la evolución de la credibilidad sobre la probabilidad de éxito de cada máquina, como se ve a continuación.

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

Al jugarse una cantidad similar de veces en cada máquina la credibilidad se concentra en igual medida alrededor de los valores de las probabilidades de éxito de cada una. Estos son “valor 1”, “valor 2” y “valor 3” para la máquina 1, 2 y 3 respectivamente.

Las simulaciones de los 1000 años arrojaron los siguientes resultados:

/histograma/

En promedio, siguiendo esta estrategia, se espera ganar “158.609” unidades monetarias en un año.

El 95% de las ganancias al final del año estuvieron entre “140” y “178”.

El promedio de ganancias al final del año es sustancialmente mas chico que el de la situación ideal y los intervalos siquiera se superponen. Se puede pensar que esta estrategia no es recomendable al diferir tanto de un resultado ideal.

Estrategia 2: Greedy con tasa observada

La maquina que esta estrategia selecciona en un cierto dia es aquella que tenga la mayor probabilidad de éxito observada hasta el momento. Como al empezar el año no se tiene ninguna informacion sobre la probabilidad de éxito de cada maquina, se decide jugar cada una por separado hasta todas tengan su primer éxito.

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

/3 lineas/ /1 linea/

En este año se usa casi exclusivamente la maquina “x” y apenas se utilizan las otras dos. Esto pasa porque las probabilidades de exitos iniciales observadas son muy variables, por lo que la primera maquina en tener varios fracasos va a tener una probabilidad muy chica y se dejara de usar. Esto impacta a la credibilidad sobre la probabilidad de éxito de cada máquina de la siguiente manera.

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

La credibilidad sobre la probabilidad de exito de la maquina “x” va a ser muy precisa, al tener muchas muestras de ella, y las demas seran poco precisas y muy parecidas a nuestra creencia inicial.

Las simulaciones de los 1000 años arrojaron los siguientes resultados:

/tabla/ promedio de cant de veces que se usa cada maquina

/histograma/

En promedio, siguiendo esta estrategia, se espera ganar “valor” unidades monetarias en un año.

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

Si bien es un metodo que puede generar mucha ganancia, al usar la mejor maquina mayoritariamente, es arriesgado porque hay una probabilidad de usar por el año entero la peor maquina posible.