

TAREA # 6

Sistema multiagente

Natalia Berenice Pérez López

6 de octubre de 2021

1. Objetivo

El objetivo de esta práctica es primero vacunar con probabilidad p_v a los agentes al momento de crearlos, de tal forma que estén desde el inicio en el estado R , lo cual indica que ya no podrían contagiarse ni propagar la infección. Después se debe estudiar el efecto estadístico del valor de p_v en (de cero a uno en pasos de 0,1) el porcentaje máximo de infectados durante la simulación y el momento (iteración) en el cual se alcanza ese máximo.

2. Desarrollo

Para generar el código objetivo de esta práctica se realizaron algunas ideas y pruebas iniciales del código, las cuales se encuentran en [mi repositorio](#) en GitHub. Se inició tomando como base el código revisado en clase [3] para calcular el movimiento de los agentes, los contagios y los agentes recuperados. Las modificaciones que se le realizaron a este código fueron; agregar un ciclo `for` para variar la probabilidad de vacunación (p_v) en pasos de 0,1 desde cero a uno, vacunar desde el inicio a los agentes de acuerdo a la probabilidad p_v , agregar otro ciclo `for` para hacer 12 réplicas para cada p_v , agregar un condicional `if` para generar las gráficas del movimiento de los agentes solamente para la primer réplica de cada p_v y generar un `data.frame` que contiene la probabilidad de vacunación, el número de réplica, el número máximo de agentes infectados por réplica y el número de iteración en el que sucede este máximo. A continuación se muestra el código modificado para esta práctica:

```
1 library(dplyr)
2 l <- 1
3 n <- 50
4 pi <- 0.05
5 pr <- 0.02
6 v <- 1 / 20
7 pv = seq(0, 1, 0.1)
8 replicas = 12
9 datos = data.frame()
10 infect <- c()
11
12 for (va in pv){
13   for (rep in 1:replicas){
14     agentes <- data.frame(x = double(), y = double(),
15                           dx = double(), dy = double(),
16                           estado = character())
17     for (i in 1:n) {
18       e <- "S"
19       if (runif(1) < pi) {
20         e <- "I"
21       }else {
22         if(runif(1) < va){
23           e = "R"
24         }
25       }
26       agentes <- rbind(agentes, data.frame(x = runif(1, 0, 1),
27                                             y = runif(1, 0, 1),
28                                             dx = runif(1, -v, v),
29                                             dy = runif(1, -v, v),
30                                             estado = e, vacuna = va))
31     }
```

```

32 print(agentes)
33 levels(agentes$estado) <- c("S", "I", "R")
34 epidemia <- integer()
35 r <- 0.1
36 tmax <- 100
37 digitos <- floor(log(tmax, 10)) + 1
38 for (tiempo in 1:tmax) {
39   infectados <- dim(agentes[agentes$estado == "I",,])[1]
40   epidemia <- c(epidemia, infectados)
41   if (infectados == 0) {
42     break
43   }
44   contagios <- rep(FALSE, n)
45   for (i in 1:n) { # posibles contagios
46     a1 <- agentes[i, ]
47     if (a1$estado == "I") { # desde los infectados
48       for (j in 1:n) {
49         if (!contagios[j]) { # aun sin contagio
50           a2 <- agentes[j, ]
51           if (a2$estado == "S") { # hacia los susceptibles
52             dx <- a1$x - a2$x
53             dy <- a1$y - a2$y
54             d <- sqrt(dx^2 + dy^2)
55             if (d < r) { # umbral
56               p <- (r - d) / r
57               if (runif(1) < p) {
58                 contagios[j] <- TRUE
59               }
60             }
61           }
62         }
63       }
64     }
65   }
66   for (i in 1:n) { # movimientos y actualizaciones
67     a <- agentes[i, ]
68     if (contagios[i]) {
69       a$estado <- "I"
70     } else if (a$estado == "I") { # ya estaba infectado
71       if (runif(1) < pr) {
72         a$estado <- "R" # recupera
73       }
74     }
75     a$x <- a$x + a$dx
76     a$y <- a$y + a$dy
77     if (a$x > 1) {
78       a$x <- a$x - 1
79     }
80     if (a$y > 1) {
81       a$y <- a$y - 1
82     }
83     if (a$x < 0) {
84       a$x <- a$x + 1
85     }
86     if (a$y < 0) {
87       a$y <- a$y + 1
88     }
89     agentes[i, ] <- a
90   }
91   aS <- agentes[agentes$estado == "S",]
92   aI <- agentes[agentes$estado == "I",]
93   aR <- agentes[agentes$estado == "R",]
94   resultado = c(va, rep, tiempo)
95   datos = rbind(datos, resultado)
96   names(datos) = c("Vacuna", "Replica", "Iteracion")
97   t1 <- paste(tiempo, "", sep="")
98   v1 = paste(va, "", sep="")
99   while (nchar(t1) < digitos) {
100     t1 <- paste("0", t1, sep="")
101     v1 = paste("0", v1, sep="")
102   }
103   if (rep == 1){
104     salida <- paste("p6_t", v1, t1, ".png", sep="")

```

```

105     tiempo <- paste("Paso", tiempo)
106     png(salida)
107     plot(1, type="n", main=c(tiempo, va), xlim=c(0, 1), ylim=c(0, 1), xlab="x", ylab="y")
108     if (dim(aS)[1] > 0) {
109         points(aS$x, aS$y, pch=15, col="chartreuse3", bg="chartreuse3")
110     }
111     if (dim(aI)[1] > 0) {
112         points(aI$x, aI$y, pch=16, col="firebrick2", bg="firebrick2")
113     }
114     if (dim(aR)[1] > 0) {
115         points(aR$x, aR$y, pch=17, col="goldenrod", bg="goldenrod")
116     }
117     graphics.off()
118 }
119 }
120 #salida <- paste("Vacuna=", va, ".png", sep="")
121 #png(salida)
122 #plot(1:length(epidemia), 100 * epidemia / n, xlab="Tiempo", ylab="Porcentaje de infectados",
123 #main=paste("Vacunados:", va * 100, "%"))
124 #graphics.off()
125 infect <- c(infect, epidemia)
126 }
127 }
128
129 datos2 <- data.frame(Cantidad_infectados = c(infect))
130 datos2 <- filter(datos2, Cantidad_infectados > 0)
131 New = cbind(datos, datos2)
132 MaxInfectados <- New %>%
133     group_by(Vacuna, Replica) %>%
134     filter(Cantidad_infectados == max(Cantidad_infectados))
135
136 MaxInfectados2 <- MaxInfectados %>%
137     group_by(Vacuna, Replica) %>%
138     filter(Iteracion == min(Iteracion))
139
140 library(ggplot2)
141 MaxInfectados2$Vacuna = as.factor(MaxInfectados2$Vacuna)
142 ggplot(MaxInfectados2, aes(x= Vacuna, y= Iteracion, fill = Vacuna)) +
143     geom_boxplot()+
144     labs(x = "Probabilidad de vacuna", y = "Iteracion")
145
146 ggplot(MaxInfectados2, aes(x= Vacuna, fill = Vacuna, y = ((Cantidad_infectados/50) * 100))) +
147     geom_boxplot()+
148     labs(x = "Probabilidad de vacuna", y = "Mayor porcentaje de infectados")

```

Listing 1: Código objetivo de la práctica para obtener la cantidad máxima de infectados por replica y la iteración en la que ocurre.

Para analizar los resultados obtenidos primero se graficó en un diagrama caja-bigote la cantidad máxima de agentes infectados en porcentaje para cada p_v (Ver figura 1). En este diagrama podemos ver que entre mayor sea la probabilidad de vacunación de los agentes se tienen menos infectados, lo cual es un dato coherente ya que cuando el agente está vacunado no puede contagiarse ni propagar la infección.

Otro dato importante de analizar en el experimento es el tiempo o iteración en el cual se tiene la mayor cantidad de agentes infectados, ya que en la vida real esta iteración podría representar por ejemplo el *pico* de una pandemia, por lo tanto nos interesaría retardar la presencia de este *pico* para preparar adecuadamente las instalaciones hospitalarias y así poder enfrentar la pandemia. Para analizar las iteraciones se realizó otro diagrama caja-bigote, el cual se muestra en la figura 2, en este diagrama podemos ver que al aumentar la probabilidad de vacunación de los agentes se consigue alcanzar un número de iteración mayor, esto se mantiene hasta la p_v de 0,8 ya que después de esta se alcanzan iteraciones menores. Es importante mencionar que en la figura 2 se está graficando la primera iteración en la que se tiene el mayor porcentaje de agentes infectados, ya que este mayor porcentaje puede continuar varias iteraciones más y después disminuir su valor.

Para analizar si existe una relación entre la variación de la probabilidad de vacunación de los agentes y la cantidad mayor de agentes infectados se realizó una prueba estadística. Primero se eligió utilizar la prueba estadística ANOVA de una vía, pero debido a los resultados obtenidos al revisar la normalidad de los datos, se eligió realizar la prueba estadística Kruskal Wallis [1].

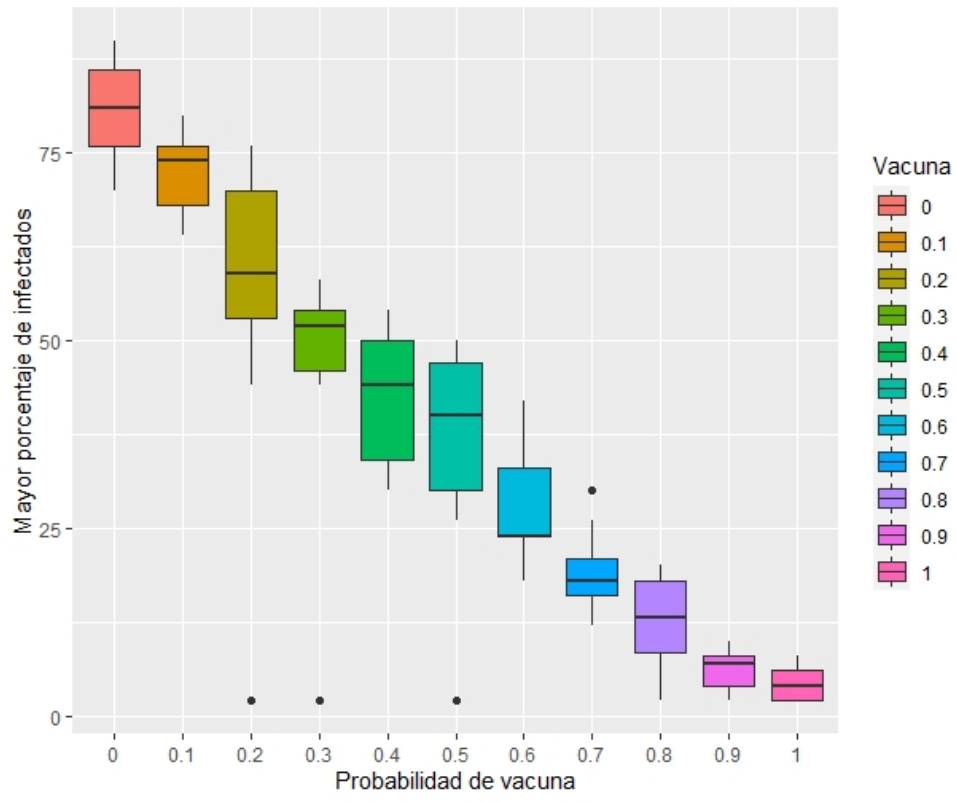


Figura 1: Mayor porcentaje de agentes infectados por replica para cada probabilidad de vacunación p_v .

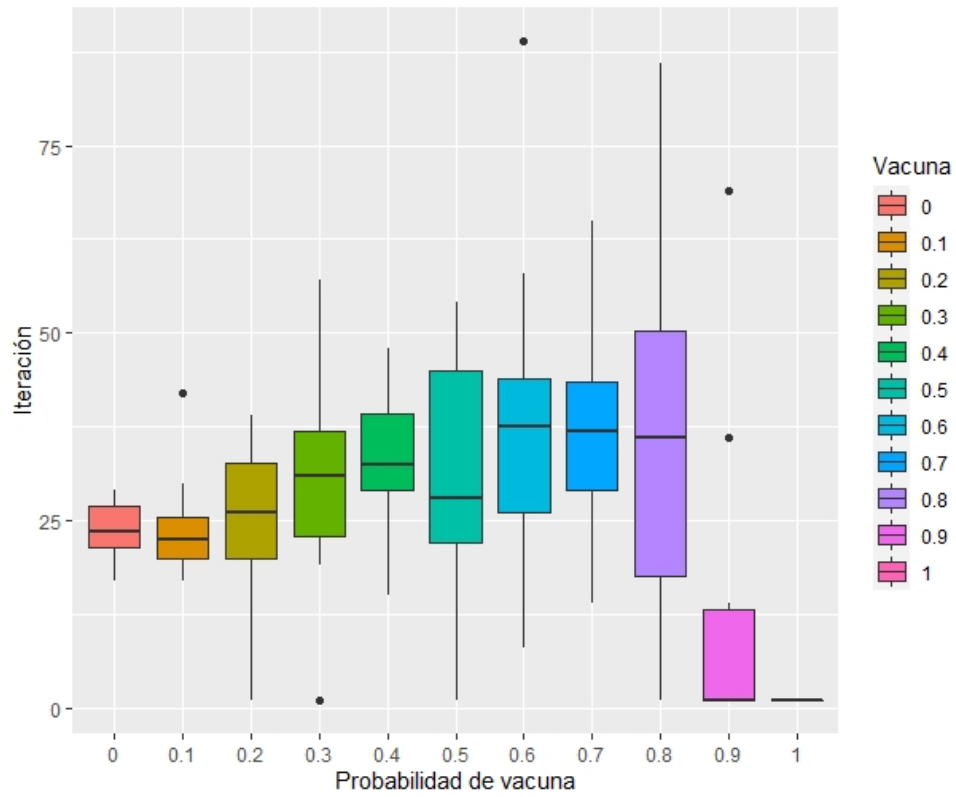


Figura 2: Iteración donde se presenta el máximo número de agentes infectados para cada p_v .

En el cuadro 1 se resumen los resultados de la revisión de los supuestos para poder aplicar la prueba estadística. El supuesto outliers se refiere a la cantidad de valores atípicos que existen en los grupos, la normalidad por grupos se obtuvo con la prueba de **Shapiro Wilk**.

Cuadro 1: Resultados del los supuestos para aplicar la prueba estadística.

Outliers	4
Normalidad por grupo	0: $p = 0,550$ / 0,1: $p = 0,449$ / 0,2: $p = 0,007$ / 0,3: $p = 4,31 \times 10^{-5}$ 0,4: $p = 0,085$ / 0,5: $p = 0,040$ / 0,6: $p = 0,146$ / 0,7: $p = 0,626$ / 0,8: $p = 0,453$ / 0,9: $p = 0,126$ / 1: $p = 0,033$
Normalidad	$p = 3,64 \times 10^{-6}$

En los resultados se observa que para la normalidad por grupos existen valores de p menores a 0,05, y la normalidad general de todo el conjunto de datos es de $p < 0,05$, por lo tanto no se tiene normalidad, así que es necesario realizar la prueba estadística **Kruskal Wallis** ya que ésta es útil para cuando no se tiene normalidad en los grupos de datos.

Al realizar la prueba **Kruskal Wallis** se obtienen los resultados mostrados en el cuadro 2.

Cuadro 2: Resultados al aplicar la prueba estadística **Kruskal Wallis**.

Chi cuadrada	Valor de p
105,96	$2,2 \times 10^{-16}$

Hipótesis nula : Las medias son iguales en todos los grupos.

Hipótesis alternativa: Debido a que $p < 0,05$ se rechaza la hipótesis nula, es decir que si existen diferencias significativas entre las medias de los grupos.

Se entiende entonces que la variación de la probabilidad de vacunación de los agentes si tiene un efecto significativo en la cantidad máxima de infectados que se tendrán en cada réplica.

También podemos realizar la prueba de suma de rangos de **Wilcoxon** por pares [2] para observar los resultados de p y determinar si existen diferencias al comparar entre ellas cada una de las probabilidades de vacunación (Ver cuadro 3).

Cuadro 3: Resultados al aplicar la prueba **Wilcoxon**.

Valor de p	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0,1	0,0624	-	-	-	-	-	-	-	-	-
0,2	0,0056	0,0499	-	-	-	-	-	-	-	-
0,3	0,0022	0,0022	0,1512	-	-	-	-	-	-	-
0,4	0,0017	0,0017	0,0494	0,1849	-	-	-	-	-	-
0,5	0,0022	0,0022	0,0309	0,1094	0,3056	-	-	-	-	-
0,6	0,0017	0,0017	0,0140	0,0194	0,0252	0,1120	-	-	-	-
0,7	0,0022	0,0022	0,0178	0,0239	0,0024	0,0309	0,0580	-	-	-
0,8	0,0028	0,0028	0,0194	0,0252	0,0028	0,0252	0,0053	0,1512	-	-
0,9	0,0017	0,0017	0,0126	0,0178	0,0017	0,0178	0,0017	0,0021	0,1094	-
1	0,0017	0,0017	0,0087	0,0131	0,0017	0,0131	0,0017	0,0021	0,0252	0,1848

En los resultados de la prueba podemos observar que la mayoría de los valores de p son menores a 0,05, solo existen algunas excepciones en las cuales se tiene una $p > 0,05$ como por ejemplo en la $p_v = 0,5$ comparada con la $p_v = 0,4$, lo cual indica que entre ellas no existen diferencias significativas en sus medias.

Posteriormente, para analizar estadísticamente los resultados mostrados en la figura 2 también se realizó la prueba estadística **Kruskal Wallis**. En el cuadro 4 se resumen los resultados de la revisión de los supuestos para poder aplicar la prueba estadística. El supuesto outliers se refiere a la cantidad de valores atípicos que existen en los grupos, la normalidad por grupos se obtuvo con la prueba de **Shapiro Wilk**.

Cuadro 4: Resultados del los supuestos para aplicar la prueba estadística.

Outliers	5
Normalidad	$p = 1,82 \times 10^{-5}$

En los resultados se observa que el valor de p es menor a 0,05, por lo tanto no se tiene normalidad. Al realizar la prueba **Kruskal Wallis** se obtienen los resultados mostrados en el cuadro 5.

Cuadro 5: Resultados al aplicar la prueba estadística **Kruskal Wallis**.

Chi cuadrada	Valor de p
49,99	$2,67 \times 10^{-7}$

Hipótesis nula : Las medias son iguales en todos los grupos.

Hipótesis alternativa: Debido a que $p < 0,05$ se rechaza la hipótesis nula, es decir que si existen diferencias significativas entre las medias de los grupos.

Se entiende entonces que la variación de la probabilidad de vacunación de los agentes si tiene un efecto significativo en el número de iteración en la cual se presenta la cantidad máxima de agentes infectados.

Realizando la prueba de suma de rangos de **Wilcoxon** podemos observar los resultados de p y determinar si existen diferencias al comparar entre ellas cada una de las probabilidades de vacunación (Ver cuadro 6).

Cuadro 6: Resultados al aplicar la prueba **Wilcoxon**.

Valor de p	0	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0,1	1	-	-	-	-	-	-	-	-	-
0,2	1	1	-	-	-	-	-	-	-	-
0,3	1	1	1	-	-	-	-	-	-	-
0,4	0,4393	0,8569	1	1	-	-	-	-	-	-
0,5	1	1	1	1	1	-	-	-	-	-
0,6	0,5930	0,7066	1	1	1	1	-	-	-	-
0,7	0,4073	0,8909	1	1	1	1	1	-	-	-
0,8	1	1	1	1	1	1	1	1	-	-
0,9	0,2295	0,1959	0,4308	0,4073	0,0963	0,4073	0,0963	0,0914	0,8569	-
1	0,0005	0,0005	0,0018	0,0025	0,0005	0,0025	0,0005	0,0006	0,0131	0,5930

En los resultados de la prueba podemos observar que al comparar $p_v = 1$ con todas las demás p_v siempre se obtienen valores de $p < 0,05$, lo cual indica que solamente con la probabilidad de vacunación de 1 se tienen diferencias significativas en el número de iteración en la cual se presenta el máximo porcentaje de agentes infectados.

A continuación se muestra el código utilizado para realizar las pruebas estadísticas **Kruskal Wallis**:

```

1 library(tidyverse)
2 library(ggpubr)
3 library(rstatix)
4 library(rapportools)
5 library(readr)
6
7 #PRUEBA ESTADISTICA PARA CANTIDAD DE INFECTADOS
8 #Estadísticas descriptivas
9 MaxInfectados2 %>%
10   group_by(Vacuna) %>%
11   get_summary_stats(Cantidad_infectados, type = "mean_sd")
12
13 #SUPUESTOS PARA ANOVA
14 #1:Outliers
15 MaxInfectados2 %>%
16   group_by(Vacuna) %>%
17   identify_outliers(Cantidad_infectados)
18
19 #2:Normalidad por Shapiro
20 MaxInfectados2 %>%

```

```

21 group_by(Vacuna) %>%
22 shapiro_test(Cantidad_infectados)
23
24 MaxInfectados2
25 str(MaxInfectados2)
26 names(MaxInfectados2)
27 shapiro.test(MaxInfectados2$Cantidad_infectados)
28
29 #PRUEBA ESTADISTICA KRUSKAL WALLIS
30 kruskal.test(Cantidad_infectados ~ Vacuna, data = MaxInfectados2)
31
32 #PRUEBA WILCOXON
33 pairwise.wilcox.test(MaxInfectados2$Cantidad_infectados, MaxInfectados2$Vacuna)
34
35 #PRUEBA ESTADISTICA PARA INTERACION
36 #Estadísticas descriptivas
37 MaxInfectados2 %>%
38   group_by(Vacuna) %>%
39   get_summary_stats(Iteracion, type = "mean_sd")
40
41 #SUPUESTOS PARA ANOVA
42 #1:Outliers
43 MaxInfectados2 %>%
44   group_by(Vacuna) %>%
45   identify_outliers(Iteracion)
46
47 #2:Normalidad por Shapiro
48 MaxInfectados2 %>%
49   group_by(Vacuna) %>%
50   shapiro_test(Iteracion)
51
52 MaxInfectados2
53 str(MaxInfectados2)
54 names(MaxInfectados2)
55 shapiro.test(MaxInfectados2$Iteracion)
56
57 #PRUEBA ESTADISTICA KRUSKAL WALLIS
58 kruskal.test(Iteracion ~ Vacuna, data = MaxInfectados2)
59
60 #PRUEBA WILCOXON
61 pairwise.wilcox.test(MaxInfectados2$Iteracion, MaxInfectados2$Vacuna)

```

Listing 2: Código para las pruebas estadísticas Kruskal Wallis y Wilcoxon.

3. Conclusión

Con base en los diagramas caja-bigote y los resultados obtenidos de las pruebas estadísticas **Kruskal Wallis** y **Wilcoxon** puedo concluir que la variación de la probabilidad de vacunación de los agentes si afecta significativamente el máximo porcentaje de agentes infectados que se tendrá, ya que a mayor probabilidad p_v se disminuye el porcentaje de agentes infectados. Además considero que la tasa más baja de vacunación para obtener la máxima cantidad de infectados en una iteración mayor es con una $p_v = 0,7$ o $p_v = 0,8$, ya que con estas probabilidades de vacunación se tiene un mayor porcentaje de infectados menor al 25 % de los agentes y además este porcentaje se puede presentar hasta iteraciones mayores a 50, es decir que se podría posponer el *pico* de la pandemia hasta la mitad de su duración total.

En general ésta práctica fue de mi agrado porque nos muestra una aplicación en epidemiología, con la cual se puede representar adecuadamente la situación que se está viviendo en la actualidad con la pandemia del *Covid* – 19.

Referencias

- [1] José Antonio: Estadística Aplicada. Kruskal-wallis en rstudio, 2020. URL https://www.youtube.com/watch?v=WEjudFpbCcE&t=996s&ab_channel=Jos%C3%A9Antonio%3AEstad%C3%ADsticaAplicada.
- [2] RDocumentation. Wilcox.test: Wilcoxon rank sum and signed rank tests. URL <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/wilcox.test>.
- [3] Elisa Schaeffer. Multiagent, 2021. URL <https://github.com/satuelisa/Simulation/blob/master/MultiAgent/SIR.R>.