Analisis de Plants vs Zombies

Ari Castillo

Contents

Analisis de variables aleatorias: "Lanzamaiz vs Dr. Zomboss"

Introduccion

Estaba jugando el primer plantas vs zombies cuando me puse a pensar, cuanto daño hacen las plantas realmente, dado que si bien se pueden contabilizar por ejemplo, los aciertos ante un zombie normal, resulta complicado cuantificar que tan eficaces son estas ante zombies mas resistentes, dado asi, me puse a pensar en la que es personalmente mi planta favorita: 'Lanzamaíz'.

Lanzamaíz

Esta planta se obtiene en el escenario final del juego, el tejado, su habilidad consiste en poder lanzar proyectiles que no chocan con el desnivel del tejado, estos proyectiles tienen una probabilidad del 25% de que sean una mantequilla, la cual provoca el doble de daño y paraliza al zombie durante 5s, recordaremos estas estadisticas para el analisis posterior.

```
shot <- 20
lucky_shot <- 40
lucky_shot_prob <- 0.25
```

Dr. Zomboss

Por otro lado, el nemesis del tejado, y el enemigo principal de la primera entrega, el Dr. Zomboss. Un zombie que creo un robot gigante con la capacidad de escupir bolas de fuego y hielo que aplastan todas las plantas a su paso. Segun la wikipedia, este posee 40,000 puntos. Un desocupado completamente. Sin embargo, nos servira de punto de referencia para el analisis.

```
zomboss_hp <- 40000
```

Analisis

Analisis simple

Dentro de este analisis recopilare los datos mas sencillos que podrias sacar, dados los datos de la wiki, sera una vision sencilla de esta, pero abarca puntos que no deben ser dejados sin tocar, tratare de realizar una serie de preguntas para hacer de esta parte algo breve y conciso.

```
n_simple_shots <- zomboss_hp / shot
n_simple_shots
```

Cuantos disparos de Lanzamaiz tomarian para derrotar al Dr. Zomboss?

```
## [1] 2000
```

Es una division sencilla, dado que sabemos la vida del Dr. Zomboss y la vida que quita cada disparo del Lanzamaiz, esto se responde con una division.

Cuanto tiempo le tomaria al Lanzamaiz derrotar al Dr. Zomboss? Introducimos una nueva variable a nuestro analisis, la cual es el tiempo de intervalo entre disparos de la Lanzamaiz, el cual segun fuentes confiables (Copilot), es de 3s.

```
shot_interval <- 3
n_simple_shots * shot_interval
```

```
## [1] 6000
```

El resultado lo podemos saber a partir de la multiplicacion de la cantidad de disparos y el tiempo que toma cada uno. Simple, verdad?

EL Analisis

Ya me aburri de hacer la otra onda y la verdad que tampoco lo queria hacer tan largo, vamos sobre, sabemos que el Lanzamaiz tiene una probabilidad de 25% de acertar un mantecazo con el doble de daño, que la verdad por eso se usa, que sino quien sabe.

Volviendo a la estadistica, esto responde al comportamiento de una variable aleatoria, cosa que podriamos calcular diversos escenario usando las herramientas de R, para crear, por ejemplo, simulaciones de montecarlo que nos ayuden a responder a los comportamientos de esta variable en una distribucion normal, pero dado que eso es muy aburrido y terminaria el analisis en 2 viñetas de codigo, nos vamos a alargar con explicaciones, porque andamos sin nada que hacer.

Simulaciones de Montecarlo

Para los que no las conozcan, las simulaciones de Montecarlo son las expresiones que decias con tus amigos de "Si peleo 100 veces con un gorila, cuantas veces les gano", si no quedo del todo claro, a medida que avancemos, entenderan a lo que me refiero. Esta herramienta la usaremos en el analisis para simular los disparos de la planta, por ejemplo "De 100 disparos, cuantos son mantecas?".

Manteca

Anteriormente mencione que dividir la vida del Dr. Zomboss por el daño que hace el Lanzamaiz es aburrido, porque lo es, si bien es un escenario posible, muuuuy improbable, pero posible, es algo que acorta mucho el resultado, y tampoco es muy preciso, la razon de esto es que la manteca juega un rol importante en el daño de el lanzamaiz, a continuacion la demostracion de esto.

```
set.seed(42)
 x<- sample(c(shot, lucky_shot),
            2000,
            prob = c(1-lucky_shot_prob, lucky_shot_prob),
            replace=TRUE
```

```
           )
 head(x)
```

Demostracion

```
## [1] 40 40 20 40 20 20
```

Lo que ves arriba es una muestra de la simulacion de los 2000 disparos que dijimos que le tomaria al lanzamaiz derrotar al Dr. Zomboss. Lo sumaremos para saber exactamente cuanta vida quitaria en esta simulacion

```
sum(x)
```

```
## [1] 50040
```

Aproximadamente 10,000 mas de lo calculado inicialmente, por que se produce este resultado? Podemos calcularlo? Con estas preguntas nos introducimos la estadistica de variables aleatorias.

Variables Aleatorias

Antes de hacer grande este nombre de "Las variables aleatorias", dejame introducirte a este concepto de una manera que un jugador de PvZ promedio entenderia, imagina que lanzas una moneda al aire, esta moneda puede caer en cara o cruz, y es un evento que no puede ser predecido con certeza, dado a esto se le llama "aleatorias".

En matematicas y estadistica, esta es la forma elegante a la que llamamos a un valor que obtenemos del azar, y por que lo estamos usando aca? El disparo del lanzamaiz cumple con estas caracteristicas, es un valor (20 o 40 de daño), que obtenemos de una probabilidad (25%). Y dado que ya hubo gente mas desocupada que yo en un pasado que desengloso todas las caracteristicas de estas variables, las usaremos para describir este comportamiento.

Valor Esperado

El **valor esperado** de una variable aleatoria es una idea bastante intuitiva. Imagina que estás jugando a un juego de azar. Podrías preguntarte: "Si juego este juego muchas veces, ¿cuánto puedo esperar ganar o perder en promedio?" Esa respuesta, entenado, es el valor esperado y se calcula de la siguiente manera:

Expected Value (Mean) of a Random Variable

$$\mu = \sum x\, P(x)$$

```
e_value <- shot*(1-lucky_shot_prob) + lucky_shot * lucky_shot_prob
e_value
```

```
## [1] 25
```

Dicho esto, podriamos decir de que esta es la media de danio por disparo de un lanzamaiz, siendo asi un punto de referencia mas exacto al momento de calcular cuantos disparos hacen falta para matar al Dr. Zomboss. Si vemos detenidamente este valor esperado es casi el mismo comparado a la media obtenida por la simulacion de anteriormente

```
avgs <- c(mean(x), e_value)
avgs
```

```
## [1] 25.02 25.00
```

RESPUESTA RAPIDA

Dado que ya tenemos el valor esperado, podemos calcular de mejor medida cuantos disparos tomaria y ya salimos de esto de una vez

```
real_n_shots <- zomboss_hp / e_value
real_n_shots
```

```
## [1] 1600
```

1600 bombazos y pal lobby.

Por el modico tiempo de...

```
time <- real_n_shots * shot_interval
time
```

```
## [1] 4800
```

4800 segundos, unos 80 minutitos para que la mejor planta del mundo gane.

RESPUESTA DE VERDAD

Si bien la respuesta anterior es valida, debemos de tener en cuenta que este es producto de la media de una simulacion de variables aleatorias, lo que lo convierte en una variable aleatoria por si misma, y incluso en ese caso podriamos decir de que si bien, probablemente el Dr. Zomboss haya muerto con esa cantidad de disparos, puede tambien no haberlo hecho, y aqui es donde entra mi querida simulacion de montecarlo.

Cual es el rango de disparos con el cual el Dr. Zomboss puede morir?

Al inicio les mencione el video que use como referencia dado que particularmente ese valor que dio por respuesta es el valor maximo de disparos con el cual el Dr. Zombos moriria, a continuacion la simulacion que lo demuestra

```
#Un millon de operaciones, pq podemos
B <- 1000000

distribution <- replicate(B, {
  life <- zomboss_hp

  shots <- sample(c(shot, lucky_shot),
                  size = 2000,
                  prob = c(1 - lucky_shot_prob, lucky_shot_prob),
                  replace = TRUE)

  cumulative_damage <- cumsum(shots)

  num_shots <- which(cumulative_damage >= life)[1]
  num_shots
})
```

Como se puede ver, si bien nuestro valor esperado usado es 1600, en un escenario real podemos tener diversos resultados que no calquen exactamente con estos, pero si a un aproximado. Para comprender esto, necesitamos entender dos cosas importantes: la **media** y la **desviación estándar**.

- La **media** ya la expliqué antes, es básicamente el promedio de los datos.

- La **desviación estándar** (para hacerla corta) es una medida de qué tanto se dispersan o se separan los datos respecto a la media. Es como decir **"más o menos X cantidad"** alrededor del valor promedio.

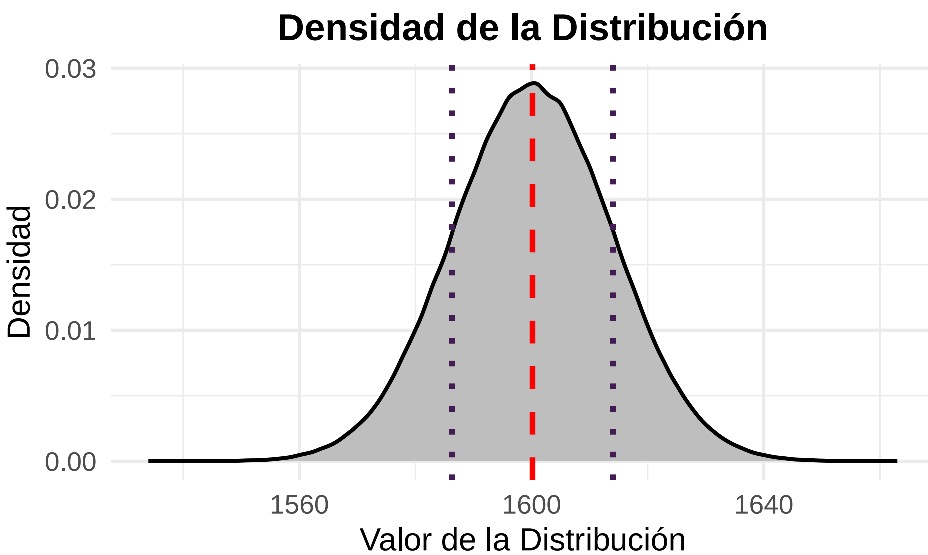
```
dis_mean <- mean(distribution)
dis_sd <- sd(distribution)

specs <- c(dis_mean, dis_sd)
names(specs) <- c("mean", "standard_deviation")
specs
```

```
##              mean standard_deviation
##        1600.15894           13.84892
```

Visualizacion

La linea roja es la media, y las lineas moradas son la desviacion estandar.



No se te hace conocida esta grafica? Como que ya la viste en algun lado? Alguna pesadilla? No? Balurde.

Esta grafica es conocida como la campana de gauss o la distribucion normal, que a como puedes ver, tiene por caracteristicas estar trazada centrada simetricamente alrededor de la media de los valores, el obtener esta grafica nos ayuda a comprender la naturaleza de los datos en estudio, y dado asi, podriamos incluso calcular la probabilidad de que X caso suceda, y para no dejar con las ganas, vamos a probar que tal con un par de casos hipoteticos.

Probabilidades

El calculo de la probabilidad de una funcion normal se define como la densidad del area bajo la curva El cálculo de la probabilidad de una función normal se define como la densidad del área bajo la curva de la distribución normal, delimitada por un intervalo específico, y se obtiene mediante la integración de la función de densidad de probabilidad. Osea la siguiente formula:

La función de densidad de probabilidad de una distribución normal es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Para calcular la probabilidad de que X esté entre a y b, se utiliza la integral de la función de densidad:

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$

Tuani, pero estamos en R, entonces vamos a usar una **funcioncita lindita** que nos ahorrará ese cálculo manual.

Con estos los valores de la media y la desviacion estandar, podemos calcular probabilidades fácilmente usando una función en R

Jugando con valores

Probemos primero la probabilidad de que el Dr. Zomboss haya muerto con los 2000 disparos que menciona el video

```
pnorm(2000, mean= dis_mean, sd= dis_sd)
```

```
## [1] 1
```

Segun la distribucion normal, esta muertisimo, no chances, pero podriamos decir lo mismo de que se muere con 21389128 disparos, no? Queremos datos reales, cuanto es el maximo que ha aguantado en un millon de simulaciones?

```
max(distribution)
```

```
## [1] 1663
```

En UN MILLON de intentos, no ha pasado ni de los 1700 disparos, ya ven a lo que me referia al inicio con que era una respuesta muy distopica?

Ahora hablando serial, con que intervalos de confianza te puedo decir que el Dr Zomboss esta muerto

```
lower_bound <- quantile(distribution, 0.0005)
upper_bound <- quantile(distribution, 0.9995)

c(lower_bound, upper_bound)
```

```
##  0.05% 99.95%
##   1555   1646
```

TE LO JURO, te lo prometo dogs, que en un 99.9% de las ocasiones te va a tomar entre 1554 a 1646 disparos y por lo tanto te digo que te va a tomar

```
times <- shot_interval * c(lower_bound, upper_bound)
times/60
```

```
##  0.05% 99.95%
##  77.75  82.30
```

entre 77.7 y 82.3 minutitos doctor.

Conclusion

Bueno, con esta la dejamos mi gente, el Lanzamaíz y el Dr. Zomboss en una batalla épica, pasamos del Lanzamaíz necesitando unos 2000 disparos para llevarse al gigante, a comprender que la mayoria de veces con 1600 serian suficientes. Ya se la saben, pongan atencion en clase, tomen bastante agua, y recuerden que las mates estan en todos lados.

Muchas gracias por leer mi articulo, espero que te haya entretenido tanto como me entretuve haciendolo.