

# Tarea 4

## Distribución de Poisson

**Erick Cervantes Mendieta**  
Matrícula: 2032430

*Modelos Probabilistas Aplicados*

Septiembre 2020

### 1. Descripción de la distribución de Poisson

La distribución de Poisson es otra de las distribuciones “especiales”, es decir, de las más comunes en la vida real. Sirve para representar el número de eventos de poca frecuencia (a veces se les llama eventos raros) que ocurren en el tiempo o en el espacio.

La función de probabilidad de una distribución de Poisson es:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

donde  $P(X = x)$  es la probabilidad de  $x$  ocurrencias ( $x = 0, 1, 2, \dots$ ), dado  $\lambda$  igual al número promedio de eventos que ocurren por unidad de tiempo o de espacio.

Una forma común de encontrar una variable aleatoria con una distribución Poisson es por medio de un modelo llamado *proceso Poisson*, si observamos dicho proceso con  $\lambda$  igual al número medio de sucesos *por unidad* (longitud, área, etc.), entonces  $Y =$  número de sucesos en  $a$  unidades tiene una distribución Poisson con media  $a\lambda$ . Una suposición clave en el desarrollo de la teoría del proceso Poisson es la independencia de los números de sucesos en intervalos inconexos (áreas, etc.).

### 2. Resumen de generadores

Los generadores propuestos a lo largo de los años se pueden clasificar en varios grupos:

1. Generadores basadas en la conexión con procesos homogéneos de Poisson (Knuth, 1969). Estos generadores son muy simples, pero funcionan en un tiempo esperado proporcional a  $\lambda$ .

2. Métodos de inversión, la inversión por búsqueda secuencial iniciada en 0 se ejecuta en el tiempo esperado proporcional a  $\lambda$ . Si la búsqueda secuencial comienza en la moda, entonces el tiempo esperado es  $O(\sqrt{\lambda})$  (Fishman, 1976).
3. Generadores basados en propiedades recursivas de la distribución (Ahrens y Dieter, 1974). Se sabe que uno de estos generadores tarda un tiempo esperado proporcional al  $\log(\lambda)$ .
4. Métodos de rechazo, los métodos de rechazo parecen conducir a los algoritmos uniformemente rápidos más simples (Atkinson, 1979; Ahrens y Dieter, 1980; Devroye, 1981; Schmelser y Kachitvichyanukul, 1981)
5. Método de aceptación-complemento con la distribución Normal como distribución inicial (Véase Ahrens y Dieter (1982)). Este enfoque conduce a algoritmos eficientes uniformemente rápidos, pero los programas de computadora son bastante largos.

En la literatura podemos encontrar algunos métodos para generar variables aleatorias de Poisson que se basaban típicamente en la proximidad de la distribución Normal para valores grandes de  $\lambda$ . Pero los algoritmos que realmente importan son los que son simples, rápidos o ambos. La definición de “rápido” puede incluir o no el tiempo de configuración. Además, dado que nuestras comparaciones no pueden basarse en implementaciones reales, es importante diferenciar entre modelos computacionales [1] (Devroye, 1986).

### 3. Generadores simples

La conexión entre la distribución de Poisson y los tiempos entre llegadas exponenciales en un proceso puntual homogéneo es la siguiente: si  $E_1, E_2, \dots$  son variables aleatorias exponenciales y  $X$  es el número entero más pequeño tal que

$$\sum_{i=1}^{X+1} E_i > \lambda$$

entonces  $X$  es Poisson( $\lambda$ ).

Usando el hecho de que una variable aleatoria uniforme se distribuye como  $e^{-E}$ , entonces se puede deducir lo siguiente: sean  $U_1, U_2, \dots$  variables aleatorias uniformes  $(0, 1)$ , y sea  $X$  el número entero más pequeño tal que

$$\prod_{i=1}^{X+1} U_i < e^{-\lambda}$$

entonces  $X$  es Poisson( $\lambda$ ).

Con la información anterior, se implementó un algoritmo en lenguaje R [2], haciendo variar el valor de  $\lambda$  a un valor menor a seis, cuya gráfica se visualiza con sesgo a la izquierda, a un valor igual a seis, cuya gráfica se supone un poco simétrica, y con un valor grande,

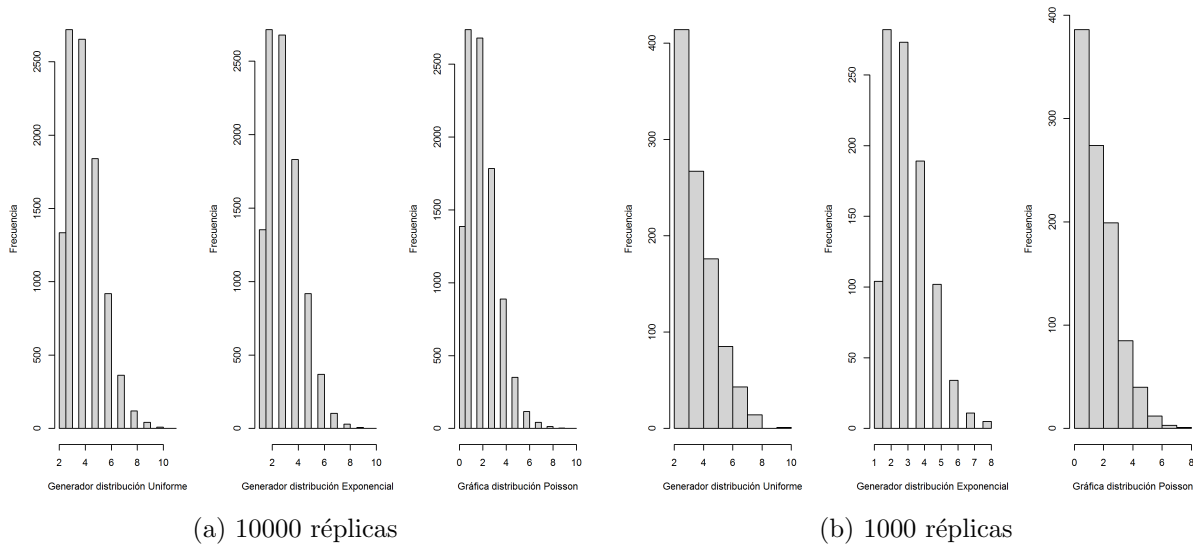


Figura 1: Gráficas utilizando  $\lambda = 2$ .

para monitorear el comportamiento de la gráfica. Por otro lado, se modificó de igual forma el número de réplicas en el experimento, en este caso solo se tomaron en cuenta dos valores, uno pequeño (1000) y uno grande (10000).

En la figura 1 se puede observar que el número de réplicas es un factor importante para que las gráficas se parezcan. En la figura 2 se puede observar este mismo comportamiento, sin embargo en la figura 3, se puede apreciar que al aumentar tanto la  $\lambda$  como el número de réplicas las gráficas parecen bastante similares. Es necesario otro tipo de herramienta estadística que nos permita concluir acerca del comportamiento de los parámetros para poder inducir cuáles son los valores que hacen que los generadores utilizando la distribución uniforme y la exponencial sean cuasi iguales a la distribución de Poisson.

Otra propuesta fue aproximar mediante una distribución Normal, sin embargo, la variación de parámetros no estuvo muy clara, al parecer si se maneja un valor de  $\lambda$  menor a 6, las gráficas muestran cierto parecido, sin embargo al aumentar el valor se van distorsionando aún más, por lo que este análisis se dejó aparte y no se consideró como parte del análisis (ver figura 4).

Al iniciar la tarea se produjo una idea errónea de lo que se quería realizar en la tarea, y el intento que se hizo fue aproximar a la distribución Poisson utilizando las distribuciones uniforme y exponencial, en una parte se generó un número aleatorio uniforme y luego se dividió el intervalo  $[0, 1]$  en subintervalos de modo que el  $j$ -ésimo subintervalo tenía longitud  $p_j$ . Por otro lado, se utiliza el hecho de que el número de eventos en algún intervalo  $[0, t]$  tiene una distribución de Poisson cuando los tiempos transcurridos entre los eventos son exponenciales, es decir, la variable a considerar es el número máximo de variables aleatorias exponenciales. En la figura 5 se puede apreciar el comportamiento de las gráficas, en las cuales

## Tarea 4

---

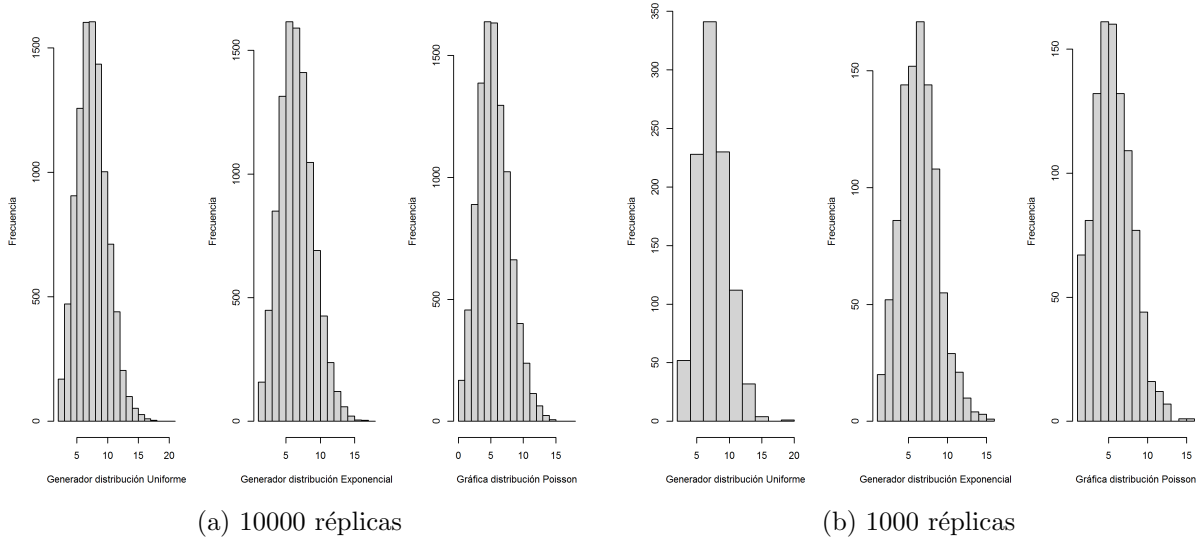


Figura 2: Gráficas utilizando  $\lambda = 6$ .

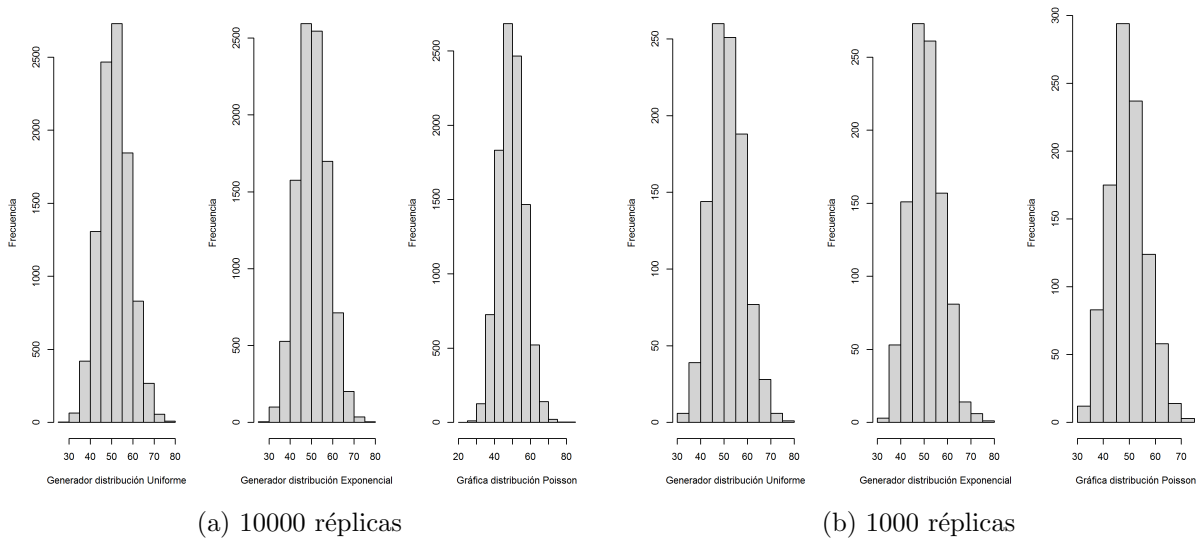


Figura 3: Gráficas utilizando  $\lambda = 50$ .

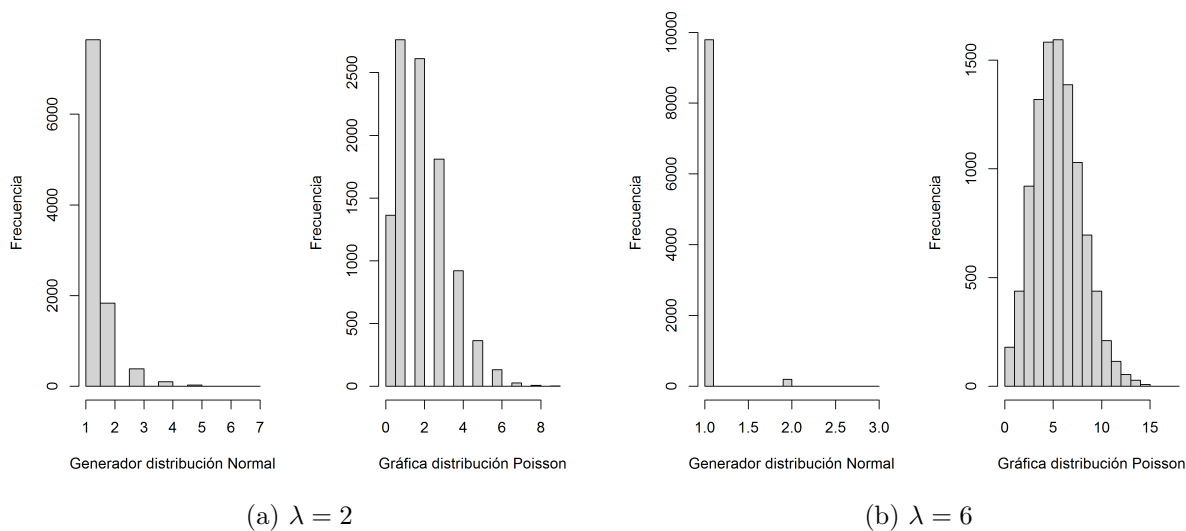


Figura 4: Gráficas utilizando 10000 réplicas.

podemos observar que no hay algún cambio drástico en cuanto al cambio de parámetros.

## Referencias

- [1] Luc Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, 175 Fifth Avenue, New York, 1986.
- [2] R Core Team. R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>, 2020.

## Tarea 4

---

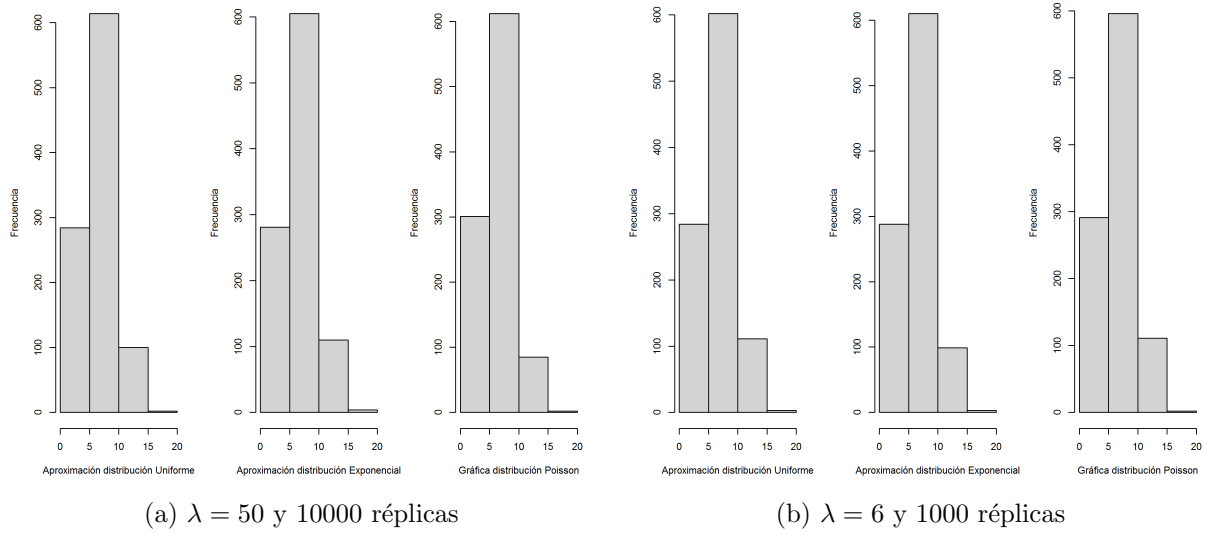


Figura 5: Gráficas utilizando diferentes valores de  $\lambda$  y número de réplicas.