

Proyecto de Simulación de eventos discretos Simulación de Reparaciones

Javier Rodríguez Sanchez C-411
María de Lourdes Choy Fernández C-412
Luis Alejandro Rodríguez Otero C-411

February 26, 2024

Contents

1	Introducción	3
1.1	Breve introducción del proyecto	3
1.2	Objetivos y metas	3
1.3	Variables que describen el problema	3
2	Detalles de implementación	4
3	Modelo Matemático	5
3.1	Descripción del modelo de simulación	5
3.2	Supuestos y restricciones	5
4	Resultados y experimentos	7
4.1	Hallazgos de la simulación	7
4.2	Interpretación de los resultados	7
4.3	Necesidad de realizar el análisis estadístico de la simulación . .	8
4.4	Análisis de parada de la simulación	8
5	Cálculo probabilístico para la estimación de resultados	10
5.1	Ejemplos de la efectividad de la estimación	11

1 Introducción

1.1 Breve introducción del proyecto

El proyecto consiste en analizar un problema de simulación de eventos discretos, implementar dicha simulación y analizar los resultados obtenidos por esta. El problema concreto consiste en un sistema que necesita n máquinas para funcionar y tiene m máquinas de repuesto, cada vez que una máquina se rompe es sustituida por una de las de repuesto y se envía a reparar, el tiempo de reparación y de duración de las máquinas siguen una distribución de variable aleatoria específica, el sistema falla cuando se rompe una máquina y no se cuenta con un reemplazo para esta.

1.2 Objetivos y metas

El objetivo del proyecto es analizar distintos casos en los que llevar a cabo esta simulación, variar las distribuciones con que se rompen y se arreglan las máquinas, la cantidad de máquinas disponibles, así como analizar la posibilidad de implementar más de un modelo que se ajuste al problema a simular. Principalmente el proyecto consiste en encontrar el o los escenarios en que el tiempo en que se rompe el sistema es mayor, es decir, que sea más duradero el sistema.

1.3 Variables que describen el problema

- **T** \longrightarrow Tiempo de la simulación.
- **F** \longrightarrow Distribución que sigue el tiempo de duración de las máquinas antes de romperse.
- **G** \longrightarrow Distribución que sigue el tiempo de reparación de las máquinas.
- **n** \longrightarrow Máquinas que necesita el sistema para funcionar.
- **m** \longrightarrow Máquinas que se tienen de repuesto inicialmente.

2 Detalles de implementación

Pasos seguidos para la implementación:

1. Selección del lenguaje de programación: Python.
2. Esbozo del problema a simular previo a la implementación.
3. Implementación de las distribuciones de variables aleatorias a utilizar: distribuciones.py
4. Implementación de 2 modelos diferentes que se ajustan al problema a analizar (Uno asumiendo la pérdida de memoria y otro asumiendo lo contrario): reparaciones.py
5. Implementación de funciones auxiliares para ejecutar múltiples simulaciones y hallar la media y varianza de los tiempos de simulación.
6. Implementación de métodos para ver los resultados de las simulaciones.
7. Realizar múltiples simulaciones y pruebas para analizar la correctitud de los algoritmos y encontrar posibles bugs.

3 Modelo Matemático

3.1 Descripción del modelo de simulación

Para simular el problema en cuestión implementamos 2 modelos de simulación que nos parecían en un principio equivalentes a dicho problema.

Modelo 1: Asumimos un modelo sin pérdida de memoria, calculamos desde un principio los tiempos en que se romperán todas las máquinas y guardaremos esa información en un heap, para siempre tener acceso a la próxima máquina que se romperá, cuando esto ocurra generamos el tiempo que esta tardará en repararse, añadiremos una de las máquinas de repuesto y se generará su tiempo de duración para añadirlo al heap. El tiempo de simulación se actualiza cada vez que una máquina se rompe o una se repara, lo que ocurra primero. La simulación termina cuando una máquina se rompe y no se tienen máquinas de repuesto disponibles, devolviendo el tiempo final de simulación, el cual se corresponde con el tiempo de vida del sistema.

Modelo 2: Asumimos un modelo con pérdida de memoria. Tiene un funcionamiento similar al modelo 1, solo que en este asumimos la pérdida de memoria, es decir, no generamos todos los tiempos de rotura de las máquinas sino que generamos solo el de la próxima máquina que se va a romper, y calculamos el tiempo de rotura de la siguiente solo cuando esta se rompa. No llevamos un heap sino una única variable que contiene el tiempo con el que se romperá la siguiente máquina. El tiempo de simulación se actualiza cada vez que una máquina se rompe o una se repara, lo que ocurra primero. La simulación termina cuando una máquina se rompe y no se tienen máquinas de repuesto disponibles, devolviendo el tiempo final de simulación, el cual se corresponde con el tiempo de vida del sistema.

3.2 Supuestos y restricciones

- Una vez que una máquina se rompe es sustituida por un repuesto instantáneamente, sin lapso de tiempo, así mismo, de ser posible, comienza a ser reparada de forma instantánea.
- El reparador (que asumimos es una persona) no tiene conocimiento previo de cuanto tiempo le tomará reparar una máquina, por lo que no aplica ninguna heurística de, por ejemplo, reparar siempre la que menos tiempo le tome.

- Varias distribuciones de variables aleatorias fueron tomadas en cuenta para las simulaciones. Finalmente decidimos que las más ajustadas a la realidad eran la distribución exponencial para el tiempo de rotura de las máquinas y la distribución normal para el tiempo en estas se reparan, por algunas consideraciones del tipo, las máquinas tienen más posibilidad de romperse mientras mas avanza el tiempo, el reparador siempre tarda un tiempo entre reparación y reparación y es poco probable que repare dos máquinas instantáneamente, etc. Para lograr varios pares de distribuciones para simulaciones distintas lo que se hizo fue variar el parámetro que estas reciben
- Implementamos ambos modelos ya que nos parecía que el problema era equivalente si tenia o no pérdida de memoria, pero aun asi, quisimos dejar que las simulaciones tuvieran la última palabra.
- Asumimos todas las máquinas como iguales, es decir, tener rotas las máquinas 2 y 9 lo interpretamos como tener 2 máquinas rotas, sin importar cuales.
- El tiempo de simulación se da en unidades de tiempo, dependiendo del contexto en que se quiera analizar el problema este puede ser interpretado como segundos, minutos, horas, días, etc.
- Intercambiamos la sección 3 con la sección 4 como venían indicadas en la orden del problema a la hora de redactar este informe pues nos parece más adecuado describir los modelos antes de analizar sus resultados

4 Resultados y experimentos

4.1 Hallazgos de la simulación

Como dijimos anteriormente, implementamos 2 modelos: Modelo 1 (sin pérdida de memoria) y modelo 2 (con pérdida de memoria).

Las simulaciones del modelo 1 resultaron (la mayoría) en una media de tiempo relativamente corta, es decir, el sistema solía fallar en poco tiempo. En el caso del modelo 2 las simulaciones duraban más tiempo, tenían una media más alta. Respecto a la varianza, el modelo 2 resultó tener una varianza notablemente superior a la del modelo 1 en la mayoría de los casos. Aquí se muestran los resultados para algunos ejemplos de ejecución:

Modelo	F	G	n	m	Media	Varianza
1	Exp(4)	Norm($6, \frac{6}{10}$)	10	20	8.7	1.5
2	Exp(4)	Norm($6, \frac{6}{10}$)	10	20	230	67
1	Exp(10)	Norm($15, \frac{15}{10}$)	20	30	15.7	2.3
2	Exp(10)	Norm($15, \frac{15}{10}$)	20	30	875.3	215
1	Exp(20)	Norm($30, \frac{30}{10}$)	10	10	22.1	5.4
2	Exp(20)	Norm($30, \frac{30}{10}$)	10	10	554.3	214.7

nota: Los valores de la media y varianza son aproximaciones. Para estos resultados se hicieron un total de 1000 simulaciones con cada conjunto de parámetros. Estos resultados se pueden comprobar ejecutando el archivo Analisis de resultados.py y probar los parámetros de la tabla u otros deseados.

4.2 Interpretación de los resultados

La razón por la que las simulaciones duran poco tiempo en el modelo 1 se debe a que todos los tiempos de rotura se calculan a la vez y siguen la misma distribución, por lo que son valores que oscilan alrededor de un mismo valor, esto significa que existe una probabilidad no pequeña de que todas tengan tiempos de rotura similares, y por tanto se rompan a la vez, dejando desprotegido el sistema. En el caso del modelo 2 ocurre lo contrario, al calcular los tiempos de rotura de una máquina a la vez es poco probable que se rompan dos máquinas simultáneamente.

Sin embargo también podemos observar que los resultados del modelo 2 son más impredecibles que los del 1, al tener una varianza mayor los resultados independientes de cada simulación abarcaron un conjunto más amplio de valores que los valores de las simulaciones del modelo 1, la razón de que

esto ocurra se basa en lo mismo, los valores de las simulaciones del modelo 1 siempre son pequeños por lo explicado anteriormente, abarcan un conjunto de números pequeños, pero en el modelo 2 tenemos simulaciones que duraron mucho tiempo, y otras que duraron poco, debido a que la característica de generar solo un tiempo de rotura a la vez vuelve más impredecible al modelo.

4.3 Necesidad de realizar el análisis estadístico de la simulación

El análisis estadístico puede ser útil a la hora de tener una mejor interpretación de los resultados dados por una simulación. Por ejemplo puede ser útil para tomar decisiones informadas basadas en los datos y no en la intuición, en nuestro caso particular puede resultarle útil al jefe de la empresa del sistema que estamos analizando, a raíz de los resultados de las simulaciones este puede analizar cuáles son las condiciones que debe crear en su sistema para alargar el tiempo de duración de este. El análisis estadístico también puede servir a la hora de validar el modelo de simulación y ver que tan cercano está de la situación real. También puede ser útil para tener un mejor conocimiento de datos dispersos, por ejemplo, lo que se explicó anteriormente de los resultados obtenidos del modelo 2, que son resultados muy variados, como refleja la varianza de las simulaciones del modelo, y por esto se hace imprescindible analizar los datos de este modelo a partir de la media. Justamente Varianza y Media fueron las medidas estadísticas que tuvimos en cuenta para analizar los datos de nuestra simulación.

4.4 Análisis de parada de la simulación

El criterio de parada del problema a simular es que el sistema falle, es decir, que se rompa una máquina y no haya un repuesto para esta. Después de numerosas simulaciones hay algunos aspectos a tener en cuenta en ambos modelos.

En el modelo 1 se hace difícil tener simulaciones largas, por lo explicado anteriormente. Para casi cualquier conjunto de parámetros la simulación dura mucho menos tiempo que el que es útil analizar resultados, por lo que podemos concluir que para este modelo no tiene sentido analizar su criterio de parada ya que este se comporta de manera similar para casi todos los escenarios.

En el modelo 2 la situación cambia. Primeramente se deben elegir la distribución del tiempo de reparación y la distribución del tiempo de rotura de manera cuidadosa, ya que si el reparador repara las máquinas a un ritmo

mayor del que estas se rompen el tiempo de simulación es potencialmente infinito. Para solucionar esto siempre intentamos que la escala de la distribución exponencial (distribución con la que se rompen las máquinas) sea menor que el valor de la media de la distribución normal (distribución con la que se reparan las máquinas), de esta manera las máquinas se rompen a un ritmo mayor que con el que son reparadas, por lo que el sistema irremediablemente fallará en un momento u otro. Otro escenario en el que el tiempo de simulación de este modelo es potencialmente infinito es cuando el valor de n (máquinas en el sistema) es notablemente menor que el valor de m (máquinas de reemplazo), es decir, cuando necesito pocas máquinas para funcionar y tengo muchas de repuesto, lógicamente, es muy improbable que el sistema falle incluso si el ritmo de rotura es mayor que el ritmo de reparación. En los casos análogos a estos las simulaciones pueden terminar prematuramente, es decir, con un número de n muy superior al de m o con un ritmo de reparación muy inferior al de rotura, de esta forma el sistema se queda sin repuestos muy rápidamente y el sistema no tarda en fallar. Por lo que de este análisis podemos concluir que necesitamos valores de n y m relativamente cercanos (preferiblemente $n \geq m$) y elegir las distribuciones F y G de manera correcta (como se explicó más arriba) para obtener tiempos de simulación adecuados para analizar datos.

5 Cálculo probabilístico para la estimación de resultados

Sea T el tiempo estimado que tendrá el sistema para dejar de funcionar. Luego $T = k * E(F)$ donde $E(F)$ es el tiempo estimado que le tomará romperse a una máquina, mientras que k es el número de máquinas que al romperse harán que el sistema colapse. Luego k está acotado por la cantidad de reemplazos. $k = m + \frac{T-E(F)}{E(G)}$, donde m es el número inicial de máquinas a reemplazar y $E(G)$ es el tiempo esperado que le tome reparar la máquina al mecánico. $T - E(F)$ es el tiempo estimado de la duración del proceso menos el tiempo que le tome romperse a la primera máquina. Luego despejando k :

$$\frac{T}{E(F)} = m + \frac{T - E(F)}{E(G)}$$

$$T * E(G) = m * E(F) * E(G) + T * E(F) - E(F)^2$$

$$T = (m * E(G) - E(F)) * \frac{E(F)}{E(G) - E(F)}$$

Aquí asumimos que $E(G)$ es mayor que $E(F)$, ya que de lo contrario el valor esperado sería potencialmente infinito.

Por ejemplo si $m = 10$, $E(G) = 5$ y $E(F) = 1$ el tiempo esperado sería:

$$T = (10 * 5 - 1) * \frac{1}{5 - 1} = 12.25$$

Importante destacar que esta estimación fue concebida a partir del análisis del comportamiento del modelo 2, por lo que las estimaciones son acertadas solo para las simulaciones hechas por este Modelo. Para el modelo 1 no encontramos mecanismo de estimación de resultados.

5.1 Ejemplos de la efectividad de la estimación

Modelo	F	G	n	m	Media real	Media estimada
2	Exp(4)	Norm($6, \frac{6}{10}$)	10	20	230	232
2	Exp(10)	Norm($15, \frac{15}{10}$)	20	30	875.3	880
2	Exp(20)	Norm($30, \frac{30}{10}$)	10	10	554.3	560
2	Exp(2)	Norm($4, \frac{4}{10}$)	30	15	59.1	58
2	Exp(1)	Norm($5, \frac{5}{10}$)	50	30	37.8	37.25
2	Exp(7)	Norm($9, \frac{9}{10}$)	30	20	590.1	605.5

nota: Los valores de la media y varianza son aproximaciones. Para estos resultados se hicieron un total de 1000 simulaciones con cada conjunto de parámetros. Estos resultados se pueden comprobar ejecutando el archivo Analisis de resultados.py y probar los parámetros de la tabla u otros deseados.