

Productividad de un trabajador, una simulación de eventos discretos

Roger Fuentes Rodríguez
Jackson Vera Pineda
Kevin Manzano Rodríguez

Universidad de la Habana
C312

9 de junio de 2024

Abstract

Las técnicas de Pomodoro, populares por prometer aumentar la productividad mediante descansos regulares, han sido objeto de debate. Nos proponemos evaluar su eficacia mediante simulaciones basadas en la teoría de colas y un modelo de cadenas de Markov con datos similares. Se diseñaron experimentos controlados para comparar la eficiencia de seguir un método de descanso estructurado, o tomar descansos aleatorios. Se demuestra que no hay diferencias significativas en la cantidad de tareas completadas entre ambos enfoques. En la simulación el tiempo de trabajo siguió una distribución exponencial, mientras que la cantidad de tareas completadas siguió una distribución de Poisson. Basándonos en nuestros hallazgos, recomendamos explorar diferentes estrategias de descanso que se adapten mejor a cada individuo, en lugar de adherirse rígidamente a políticas de descanso predefinidas.

Además, nuestro estudio busca realizar comparaciones estadísticas entre diferentes escenarios de trabajo, evaluando cómo varían los resultados bajo distintas condiciones de tareas, descansos e interrupciones. Este análisis comparativo permite ofrecer recomendaciones prácticas para la implementación de políticas laborales que promuevan un ambiente de trabajo más productivo.

2. Metodología

Se diseña una simulación discreta para modelar el comportamiento de un trabajador en un entorno laboral típico. La simulación se diseñó utilizando software de simulación adecuado, eligiendo entre opciones como **Arena**, **AnyLogic** o **SimPy**, dependiendo de las características específicas requeridas para nuestro modelo.

1. Introducción

La eficiencia laboral es un factor crítico en cualquier organización, siendo la productividad del trabajador uno de los pilares fundamentales para alcanzar objetivos empresariales. En este contexto, la gestión de tiempo y recursos se convierten en una tarea esencial, donde los descansos y las interrupciones juegan un papel significativo en la determinación de la productividad diaria. Conscientes de la importancia de optimizar estos aspectos, nuestro proyecto se enfoca en simular el comportamiento de un trabajador mediante eventos discretos, buscando modelar de manera realista las tareas, descansos e interrupciones que ocurren durante un día laboral típico.

Nuestra propuesta se centra en desarrollar un modelo que refleje con precisión las dinámicas de trabajo, permitiendo analizar cómo estos eventos afectan la productividad del trabajador. Con la simulación, pretendemos identificar patrones y tendencias que puedan ser aprovechados para mejorar la eficiencia laboral. Específicamente, nos interesan explorar estrategias que permitan minimizar los efectos negativos de las interrupciones y maximizar los períodos de descanso de manera óptima, con el objetivo de incrementar la productividad general.

2.1. Modelado de Eventos Laborales

Dentro de la simulación, se modelan tres tipos principales de eventos:

1. **Tareas:** Representan las actividades productivas que realiza el trabajador, con variabilidad en duración y prioridad.
2. **Descansos:** Reflejan los intervalos de tiempo destinados al descanso del trabajador, considerando la duración y la frecuencia de estos descansos.
3. **Interrupciones:** Modifican el flujo normal de tareas, representando distracciones o interrupciones externas.

Cada evento es modelado con parámetros probabilísticos que reflejan la variabilidad real en un entorno laboral.

2.2. Análisis Estadístico

Una vez definida la simulación, procedemos a ejecutar múltiples instancias de la simulación bajo diferentes escenarios para recopilar datos suficientes para el análisis estadístico. Los escenarios varían en términos de duración de las tareas, frecuencia de descansos e intensidad de interrupciones.

Utilizamos herramientas estadísticas para analizar los resultados, incluyendo medidas de centralidad (como la media y mediana), dispersión (varianza y desviación estándar) y distribuciones de probabilidad para los tiempos de tareas, descansos e interrupciones. También realizamos pruebas estadísticas para comparar la productividad entre los diferentes escenarios.

2.3. Recomendaciones Basadas en Resultados

Finalmente, basándonos en los hallazgos de nuestro análisis estadístico, derivamos recomendaciones prácticas para mejorar la productividad laboral. Estas recomendaciones pueden incluir estrategias para gestionar mejor los descansos y minimizar las interrupciones, así como sugerencias para la implementación de políticas laborales que promuevan un ambiente de trabajo más eficiente.

3. Detalles de la implementación

3.1. Herramientas y tecnologías utilizadas

1. Lenguaje: **Python**, debido a su amplio soporte para operaciones matemáticas y manipulación de arrays, así como su comunidad activa que ofrece múltiples bibliotecas útiles para la simulación y análisis de datos.
2. **SimPy** para llevar a cabo simulaciones complejas. **NumPy**, para el manejo eficiente de matrices y operaciones matemáticas, esencial para la implementación de la cadena de Markov, entre otras tenemos **Random**, **Scipy.stats**, **Pandas**.

3.2. Descripción del Proceso de Implementación

3.2.1. Teoría de colas

1. Creación de los diferentes generadores de variables aleatorias.
2. Creación de la clase **Person** con sus atributos, procesos y métodos: Se especifica cómo interactúa por medio de los métodos que serán tratados como procesos, y a la persona como un recurso; por la librería **SimPy**.
3. Realización de la simulación: Se genera un *Environment*(ambiente) de **SimPy**, luego se ejecuta la simulación.
4. Recogida de datos.
5. Se calcula una aproximación de la cantidad de veces que se está en cada estado para luego compararla con el otro modelo propuesto.

3.2.2. Cadena de Markov

1. Recogida de datos producido por el modelo de teoría de colas.
2. Modelado de la Cadena de Markov: Se definen los estados del sistema (trabajando, descansando, interrumpido) y se establecen las probabilidades de transición entre estos estados basándose en los datos proporcionados.
3. Creación de la Matriz de Transición: Se construye una matriz de transición 3×3 que representa las probabilidades de cambio de estado.
4. Simulación de Pasos de Tiempo: Se implementa un bucle que simula el paso de tiempo, donde en cada iteración se genera un número aleatorio para determinar el próximo estado del sistema basándose en las probabilidades de transición.
5. Registro de Resultados: Durante la simulación, se registra el estado del sistema en cada paso de tiempo. Al final de la simulación, se analizan estos resultados para obtener información sobre el comportamiento del sistema.

Se puede consultar la implementación realizada desde este repositorio de GitHub. Repositorio de GitHub

4. Resultados y Experimentos

4.1. Principales interrogantes a tratar

1. Comparar qué estrategia de administración de tiempo es más productiva:
 - a) $24 - 6$, 24 minutos de trabajo y 6 minutos de descanso.
 - b) $12 - 3$, 24 minutos de trabajo y 6 minutos de descanso.
 - c) Free(aleatorio)
2. ¿Cuál es la distribución de la cantidad de tareas completadas en una ventana de tiempo de 480 minutos?
3. ¿Cómo afecta la duración de los descansos a la cantidad de tareas completadas?
4. ¿Cómo afectan las interrupciones a la cantidad de tareas completadas?
5. ¿Cómo se relacionan el número de descansos y el número de tareas completadas?

4.2. Primer acercamiento

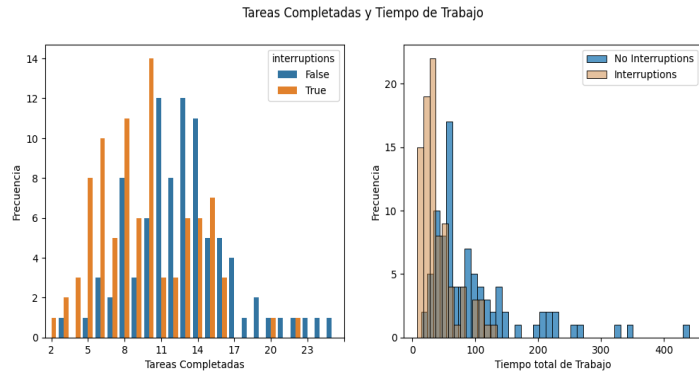


Figura 1: Tareas completadas y Tiempo de trabajo, con o sin interrupciones

Tenemos dos tipos de resultados obtenidos, evidentes y no-evidentes. Entre los no-evidentes tenemos que no es importante la metodología seguida con respecto a los descansos, se puede ser libre con el horario; el tiempo de trabajo distribuye exponencial y la cantidad de tareas distribuye poisson. Mientras en los evidentes encontramos que el tiempo de trabajo y la cantidad de tareas son proporcionales, y la cantidad de tareas realizadas se ve afectadas si se simula alguien propenso a interrupciones, observar 1.

Se llevaron a cabo simulaciones de 480 minutos (8 horas) con distintas metodologías, descansos cada 24, 12, o una cantidad aleatoria de minutos siguiendo una distribución exponencial con parámetro $\lambda = 24$; la duración de los descanso respectivamente queda : 6, 3, o una cantidad aleatoria de minutos dada por una distribución exponencial con parámetro $\lambda = 6$. Todas estas variaciones se simularon con o sin interrupciones.

A partir de las simulaciones con el modelo de colas, se recopilan los datos para conformar el modelo de cadenas de Markov, con el cual se llevan a cabo comparaciones, para ello se utilizan funciones que arrojan la diferencia modular entre los valores que devuelven ambos modelos ($f(x) = |y_1 - y_2|$ siendo x, y_1, y_2 el tiempo de simulación, el valor que devuelve el modelo de colas, el modelo que devuelve el modelo de cadenas, respectivamente).

Los tests utilizados para validar las hipótesis, son: "KS statistic", "Chi-square statistic", "U statistic", "D statistic", "H statistic".

Para el criterio de parada se usa qué tanto se aleja el estimador muestral de la población, y esto se comprueba cuando $\frac{\sigma}{\sqrt{k}} < c$ donde c es el error que se está dispuesto a cometer (la mitad de la longitud de confianza) con 95 % de confianza.

4.3. El efecto de las interrupciones

Es intuitivo pensar que las interrupciones pueden provocar una disminución en el desempeño de una persona. Aunque en nuestra simulación no se tengan en cuenta los efectos psicológicos de dichas perturbaciones, estas toman tiempo útil, por tanto podríamos preguntar-

nos ¿Cómo afectan las interrupciones a la cantidad de tareas completadas?

Para responder esta interrogante hemos realizado las siguientes pruebas:

1. Prueba Kolmogorov-Smirnov: test no paramétrico de Bondad de Ajuste para verificar si las muestras provienen de poblaciones significativamente diferentes con respecto al tiempo trabajado.
2. Prueba Chi-Square se comprueba que la cantidad de tareas realizadas en ambas muestras es significativamente diferente.
3. Prueba de Mann-Whitney U (Wilcoxon Rank): prueba no paramétrica sobre las medianas de ambas muestras con hipótesis alternativa "menor que" (parámetro utilizado *alternative = 'less'*) para verificar cuál de las dos es la menor. Este test es usado (y no T-test por ejemplo) dando que no podemos asumir que la variable en cuestión está normalmente distribuida, ni que las condiciones que garantizan el *TCL* (Teorema Central del Límite) se cumplen.

En la prueba Kolmogorov-Smirnov dado que él $p\text{-value} < \alpha = 0,05$ existe evidencia suficiente para rechazar la hipótesis nula, por tanto consideraremos que las dos muestras (el tiempo de trabajo con y sin interrupción) no siguen la misma distribución.

En la prueba Chi-Square dado que él $p\text{-value} < \alpha = 0,05$ existe evidencia suficiente para rechazar la hipótesis nula, por tanto consideraremos que las dos muestras (la cantidad de tareas completadas y sin interrupción) no siguen la misma distribución.

En la prueba de Mann-Whitney U (Wilcoxon Rank) dado que el $p\text{-value} < \alpha = 0,05$ existe evidencia suficiente para rechazar la hipótesis nula, por tanto consideraremos la hipótesis alternativa: es significativamente mayor la cantidad de tareas completadas si no la persona no es interrumpida a si lo es.

4.4. Distribución de variables de interés

Determinar la distribución que sigue un conjunto de datos en general no es una tarea sencilla y en muchos de los casos no es posible realizarse. Luego de observar los gráficos de los datos de la simulación, pueden notarse no muy evidentes, pero sí rasgos distintivos de algunas distribuciones conocidas, como exponencial en el caso del tiempo total trabajado, y poisson en el caso de la cantidad de tareas resueltas.

Se comprueban dichas hipótesis:

1. Realizando prueba Kolmogorov-Smirnov con la variable de tiempo de trabajo, primero ajustando a los posibles valores exponenciales, y luego comparándola con datos de la distribución teórica.
2. Realizando prueba Chi-square con la cantidad de tareas completadas.

En la prueba de Kolmogorov-Smirnov dado que el $p\text{-value} < \alpha = 0,05$ existe evidencia suficiente para rechazar la hipótesis nula, por tanto no es posible afirmar que el tiempo trabajado sin interrupciones sigue una distribución exponencial. Pero dado que el $p\text{-value} > \alpha = 0,05$ no existe evidencia suficiente para rechazar la hipótesis nula, por tanto es posible considerar que el tiempo de trabajo con interrupciones sigue una distribución exponencial según los datos arrojados por nuestra simulación. Más aun dado que el $p\text{-value} > \alpha = 0,05$ no existe evidencia suficiente para rechazar la hipótesis nula, por tanto es posible considerar que el tiempo de trabajo en general sigue una distribución exponencial según los datos arrojados por nuestra simulación.

En la prueba de Chi-Square dado que el $p\text{-value} > \alpha = 0,05$ no existe evidencia suficiente para rechazar la hipótesis nula, por tanto es posible considerar que la cantidad de tareas en general sigue una distribución Poisson según los datos arrojados por nuestra simulación.

4.5. Mejor estrategia de administración

Se propone verificar si alguna, y cuál de las estrategias "Pomodoro", tiene un resultado considerable en la cantidad de tareas resueltas de nuestra simulación.

Para esta cuestión se realizó:

1. Crear gráficos de dispersión para visualizar el comportamiento del trabajo y el descanso, entre los grupos que siguen diferentes metodologías de planificación
2. Prueba de Kruskal-Wallis no paramétrica para determinar si hay diferencias estadísticamente significativas entre medianas de los grupos de interés (es equivalente al ANOVA), dado que no podemos garantizar la normalidad de nuestros datos.

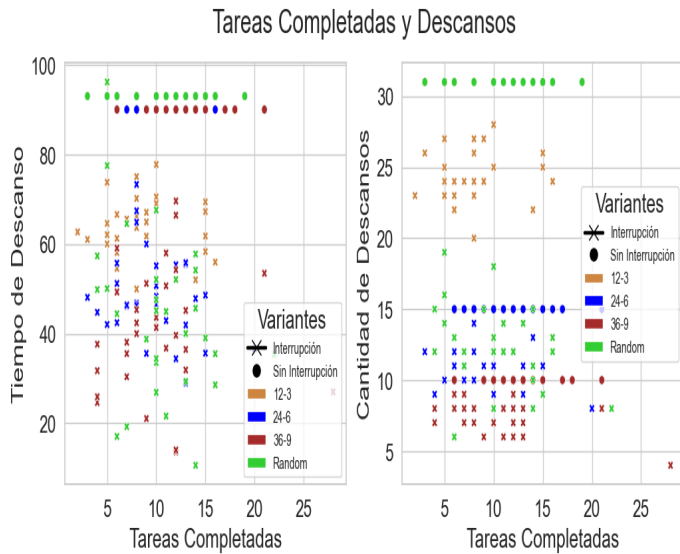


Figura 2: Tareas Completadas y Descansos

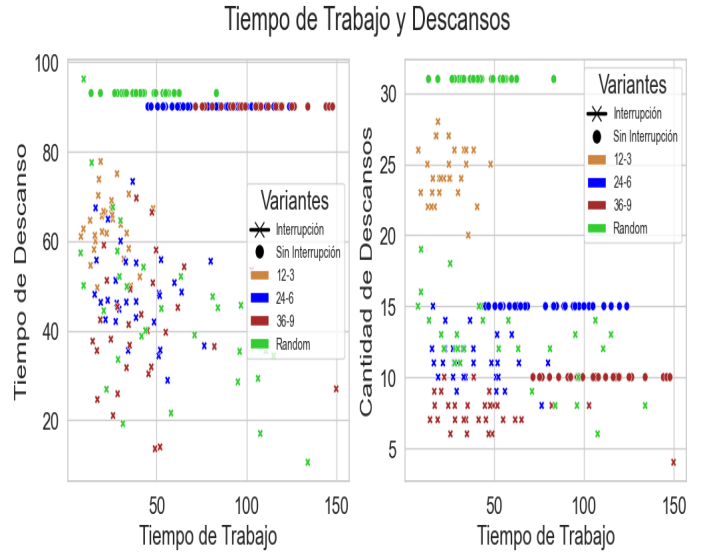


Figura 3: Tiempo de Trabajo y Descansos

En 2 se puede observar que ningún grupo sobresale más que otro en la cantidad de tareas completadas, puede notarse porque ningún color (grupos de interés) se posiciona significativamente más a la derecha que otro, en su lugar se posicionan aleatoriamente en el plano. Esto nos sugiere que puede no existir diferencia o mejora entre aplicar una estrategia de planificación u otra, ni siquiera analizando los grupos según las interrupciones (puntos y cruces).

Al hacer un análisis similar, se observa ligeramente un mejor desempeño en el tiempo de trabajo en estrategias con descansos más prolongados y más espaciados (ver figura 3 gráfica inferior derecha), pero no se asegura que sea significativo.

5. Modelo Matemático

El modelo esta conformado por un conjunto de estados y transiciones entre estos con ciertos valores, es decir un grafo dirigido ponderado. Se plantean los estados:

$$\{trabajando = W, descansando = B, interrumpido = I\}$$

La función de transición:

$$\delta(q_i) = \{(q_0, p_0), (q_1, p_1), (q_2, p_2)\}$$

donde p_j representa la probabilidad de que estando en q_i se vaya a q_j .

Se toma un conjunto de restricciones y supuestos de forma severa:

- No se está en dos estados a la vez
- cuando se entra a un estado no se sale hasta que se termina
- La determinación del siguiente estado no depende de información anterior.
- la suma de las probabilidades que salen de un estado suman 1

Para el cálculo de las probabilidades, se utiliza la probabilidad clásica tomando en cuenta los datos que brindó el modelo basado en teoría de colas. Se cuentan la cantidad de veces que se pasa de estar en el estado q_i a estar en el estado q_j denotemosla como c_j entonces,

$$p_j = \frac{c_j}{\sum_{i=1}^3 c_i}$$

De antemano es notable que las restricciones tomadas son muy fuertes por lo que los resultados no deberían ser muy justos. Llevamos a cabo varias comparaciones, dada la representación como cadena de Markov, se contó la cantidad de veces que se estaba en cada estado, para simulaciones tomando como entrada la cantidad de iteraciones; por su contrapuesto tenemos al modelo por colas y se quiere ver qué tan próximo resulta la cadena a este, entonces tomamos como aproximación para cada parámetro a $\frac{\text{Tiempo parámetro}}{\text{cantidad parámetro}}$ o sea el tiempo que está en los diferentes estados dividido por la cantidad de veces que se complete dicho estado.

Para todas las gráficas se toma como las curvas rojas a la realizada por teoría de colas, y a la azul a la realizada por cadena de Markov.

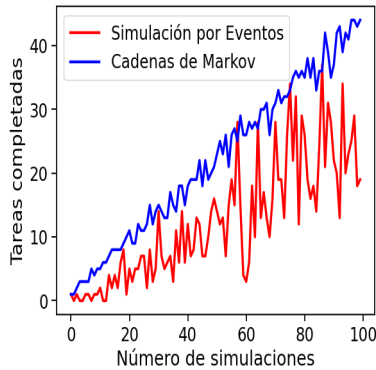


Figura 4: Parámetro=tareas

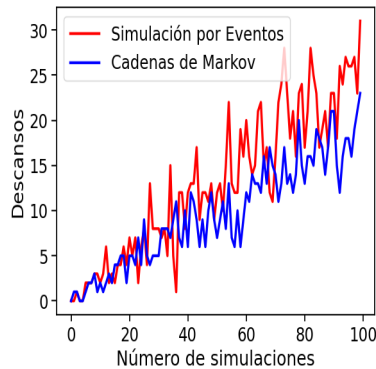


Figura 5: Parámetro=descansos

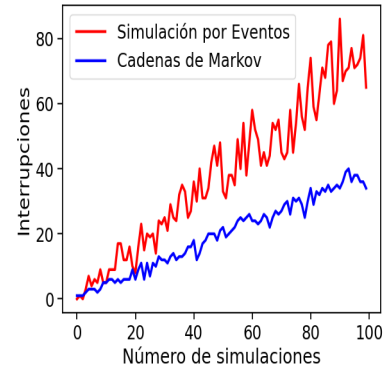


Figura 6: Parámetro=interrupciones

Comparaciones

Notemos que aunque se ven diferentes guardan relación, pues siguen su crecimiento a lo largo del tiempo aunque con un alto grado de ruido en particular cuando se trata de las tareas 5, pues recordemos que se tomo ciertas asunciones con respecto a estas, veamos como con las interrupciones 6 si se tiene más proximidad porque su comportamiento en ambos modelos es bastante similar

Referencias

- [1] Biwer, F., Wiradhany, W., oude Egbrink, M. G. A., & de Bruin, A. B. H. (2023). *Understanding effort regulation: Comparing 'Pomodoro' breaks and self-regulated breaks..* British Journal of Educational Psychology, 93(Suppl. 2), 353-367. <https://doi.org/10.1111/bjep.12593>
- [2] Léon G. Faber, Natasha M. Maurits, Monique M. Lorist. *Mental Fatigue Affects Visual Selective Attention.* Article in PLOS ONE · October 2012 DOI:10.1371/journal.pone.0048073
- [3] Blaz Kos. *Interruptions at work Why are they a problem and the best ways to handle them.* Spica. <https://www.spica.com/blog/interruptions-at-work>
- [4] Hannah Ross. *The Impact of Interruptions on Productivity & How to Combat Them.* Fellow. <https://fellow.app/blog/productivity/the-impact-of-interruptions-on-productivity-how-to-combat-them/>