

Minería de Datos

Lab. #4

Ivan Saavedra, Ph.D.

saavedrai@uninorte.edu.co

Universidad del Norte
División de Ingenierías
Dpto. Ingeniería de Sistemas



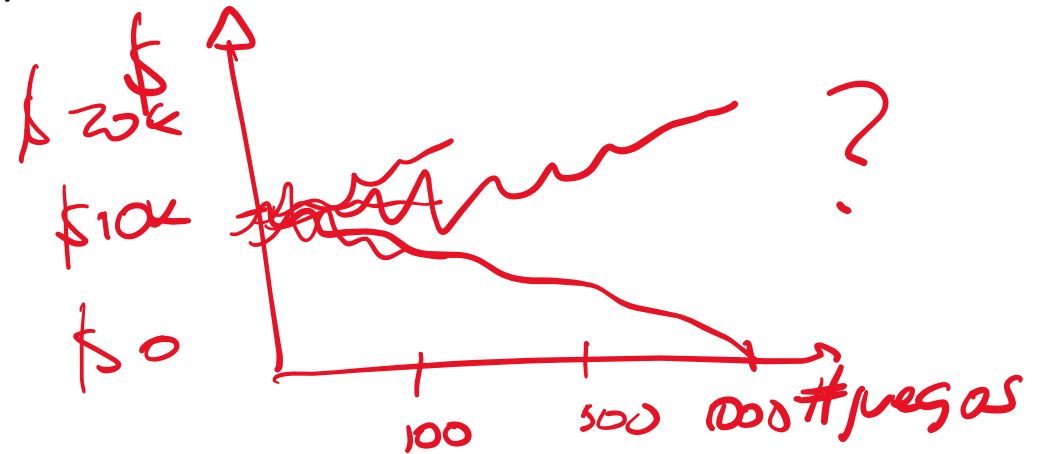
202130

1 Monte Carlo simulations

Asuma un juego en la cual un jugador realiza lanzamientos de un dado personalizado para obtener un resultado de 1 a 100. Si el jugador obtiene un resultado entre 1-51, la casa gana, pero si el resultado se encuentra entre 52-100, el jugador gana. Se necesita encontrar la distribución del monto resultante después de que el jugador realice un número x de apuestas. ¿Será que el jugador podrá perderlo todo?

Metodología

- Crear simulador del lanzamiento del dado
 - Devuelve verdadero si el jugador gana
- Crear función que simula la apuesta
 - Dinero Total: El jugador comienza con \$10,000 ← *initial*
 - Monto de la Apuesta: el monto a apostar en cada apuesta es de \$100
 - Juegos Totales: El número de veces que el jugador juega.
- Crear el ciclo que llame a la función que simule las apuestas
- Visualizar el resultado de los montos resultantes al finalizar un número de apuestas (número de juegos vs monto restante)



2 Monte Carlo simulations

Usted ha sido seleccionado para pronosticar cuanto dinero debe ser presupuestado para las comisiones de venta para el siguiente año. Usted deberá utilizar el conjunto de datos suministrados en la tabla para realizar esta predicción utilizando simulaciones de Monte Carlo.

Metodología

- **Commision Ammount** = Actual Sales * Commission Rate
- **The commission rate** esta basado en el porciento a planificar "the percent to plan", ver tabla.
- De datos históricos, fue concluido que el porcentaje histórico para la distribución objetivo "the historical % to target distribution" sigue una distribución normal con media 100% y desviación estándar de 10%.
- Crear simulador de porcentajes que sigan el histórico de distribución normal
- Crear función que calcule la tasa de comisión
- Calcule la comisión
- Realizar 1000 simulaciones
- Simulando **500 sales reps** con distribución histórico de objetivo de ventas "Historical Sales Target Distribution" presentados en la tabla
- Visualizar el resultado de las comisiones resultantes al finalizar un numero de simulaciones
- Cual seria el promedio del total de las comisiones resultante?

Sales Rep	Sales Target	Actual Sales	Percent to Plan	Commission Rate	Commission Amount
1	\$ 100,000	\$ 88,000	88%	2%	\$ 1,760
2	\$ 200,000	\$ 202,000	101%	4%	\$ 8,080
3	\$ 75,000	\$ 90,000	120%	4%	\$ 3,600
4	\$ 400,000	\$ 360,000	90%	2%	\$ -
5	\$ 500,000	\$ 350,000	70%	2%	\$ -
	\$ 1,275,000	\$ 1,090,000			\$ 13,440

Rate Schedule	
0-90%	2%
91-99%	3%
>=100%	4%



Best

f

Support

Thank you ... support END = #

[]

(I)

&

Generador de texto y autocompletad

- Utilizar el conjunto de datos de ejemplo para construir la **matriz de probabilidad de transición**

- Crear el ciclo que llama a la función anterior un numero de veces n

- un numero de veces n
"Thanks" y "Thank you",

4. MCMC

Asuma que necesita realizar muestras de una distribución gamma definida como $f(x) = \frac{x^{c-1}}{\Gamma(c)b^c} e^{-\frac{x}{b}}$ en el intervalo de 0 a 20.

b, c
 \propto
 $\Gamma(c)$
 $gamma(c)$

Metodología

- Crear función para encontrar $f(x)$ ✓
- Crear función de MCMC ✓
 - Utilizar la **distribución uniforme** para generar los valores de los estados
 - Visualice **distribución real vs simulada** para una corrida del MCMC
- Crear ciclo que llame a la función MCMC un numero de **x veces para promediar el valor máximo** de la función en cada corrida del MCMC

MCMC

Entregable:

- Un archivo de Jupyter Notebook con el desarrollo del análisis
- Las conclusiones y respuestas al objetivo del análisis deben ser contestadas en el mismo notebook.
- Se sugiere que comente las secciones de manera adecuada para una mejor interpretación de su análisis.
- La fecha de entrega es el Lunes 13 de Septiembre del 2021 vía catalogo web enlace de laboratorios.