Pontificia Universidad Católica de Chile
Departamento de Ingeniería Industrial y Sistemas
ICS3105 - Optimización Dinámica

Tarea 2

2º semestre 2021 - Matías Klapp

Grupo 5

Anaís Martínez, Alexandra Ovalle, Shun Wei Rao

Preguntas

- Se modela el problema que planifica una política de decisión que minimiza el costo total esperado como un MDP.

  Etapas $\{1, 2, ..., 52\}$

  Estados $y_t \in \{0, 1, ..., 200\}$

  Acciones $\mathbb{X}(y) = \{0, 1, ..., 200 - y\}$

  Costo inmediato:

$$r(y, x, D) = \underbrace{1500 \cdot \mathbb{I}_{x>0} + 25 \cdot x}_{g(x)} + \underbrace{90 \cdot (D - y - x)^+}_{q} + \underbrace{6 \cdot (y + x - D)^+}_{h} \ \forall t$$

  En promedio:

$$r(y, x) = 1500 \cdot \mathbb{I}_{x>0} + 25 \cdot x + 90 \cdot \mathbb{E}_{D_t}(D - y - x)^+ + 6 \cdot \mathbb{E}_{D_t}(y + x - D)^+ \ \forall t$$

  Probabilidad de transición de estado:

$$p(y_{t+1}|y + x) = \begin{cases} f(y + x - y_{t+1}) & \text{si } 0 < y_{t+1} \leq y + x \\ \sum_{k \geq y+x} f(k) & \text{si } y_{t+1} = 0 \\ 0 & \text{e.o.c} \end{cases} \tag{1}$$

  Además, dado que D sigue una distribución Uniforme Discreta: $X \sim U(\mu_{m(t)} - \delta_{m(t)}, \mu_{m(t)} + \delta_{m(t)})$. La probabilidad de que haya demanda x y la esperanza vienen dadas respectivamente por las Ecuaciones (2) y (3)

$$P\left[X = x\right] = \frac{1}{2 \cdot \delta_{m(t)} + 1} \tag{2}$$

$$\mathbb{E}(X) = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \mu_{m(t)} \tag{3}$$

Por lo tanto, el costo inmediato en promedio y la probabilidad de estado quedán dadas por las Ecuaciones (4) y (5)

$$r(y, x) = 1500 \cdot \mathbb{I}_{x>0} + 25 \cdot x + 96 \cdot \mu_{m(t)} \tag{4}$$

$$p(y_{t+1}|y + x) = \begin{cases} \frac{1}{2 \cdot \delta_{m(t)}+1} & \text{si } 0 < y_{t+1} \leq y + x \\ 1 - \frac{y+x}{2 \cdot \delta_{m(t)}+1} & \text{si } y_{t+1} = 0 \\ 0 & \text{e.o.c} \end{cases} \tag{5}$$

- Se corre el modelo para obtener la decisión óptima en cada periodo para cada posible estado, para ello, el código fue basado en lo visto al final de la clase 2-Inventario estocástico multiperiodo. A este código se le hicieron algunas modificaciones estructurales, dado que la instancia era mucho más compleja que la del ejemplo (ver Detalles en archivo Python adjunto).

  La optimización se basa en revisar todas las posibles combinaciones de estados y acciones en cada periodo dadas las respectivas probabilidades. Todo esto recursivamente, de adelante hacia atrás. En consecuencia, el mínimo costo esperado del sistema a lo largo de los 12 meses se encuentra al observar aquel costo asociado a la primera semana y con un estado inicial de 30 unidades (enunciado). De modo que, dicho costo resulta ser $59358.217. El código demora en correr 248.93 segundos, además considera dentro del tiempo la impresión de t, s, x óptimo y costo acumulado desde la etapa final asociado a cada combinación.

  Se verifica que la política óptima sigue una política(s, S) a partir de una pequeña muestra del input (Ver Cuadro 1). Se puede apreciar que para cada semana se tiene una política (s, S) distinta, donde el s es aquel estado desde el que no se encarga nada de producto. En el Cuadro 1, para la semana 1: s = 10; en la semana 2: s = 11 y en la semana 7: s = 9. Posterior a estos estados, es decir, con niveles de inventario mayores, la decisión será no pedir nada. Por otra parte, el S está dado por y+x, en otras palabras, el nivel de inventario que queda una vez que llega lo que se pidió al inicio de cada semana. En el Cuadro 1, para la semana 1: S = 122; en la semana 2: S=118 y en la semana 7: S = 100.

Cuadro 1: Resultados de optimización (tipo política (s, S))

Semana	y	x	y+x
1	0	122	122
1	1	121	122
1	2	120	122
1	3	119	122
1	4	118	122
1	5	117	122
1	6	116	122
1	7	115	122
1	8	114	122
1	9	113	122
1	10	0	10
2	0	118	118
2	1	117	118
2	2	116	118
2	3	115	118
2	4	114	118
2	5	113	118
2	6	112	118
2	7	111	118
2	8	110	118
2	9	109	118
2	10	108	118
2	11	0	11
9	0	122	100
9	1	121	100
9	2	120	100
9	3	119	100
9	4	118	100
9	5	117	100
9	6	116	100
9	7	115	100
9	8	92	100
9	9	0	9

- Sabiendo que la política óptima es del tipo (s, S), la iteración será solamente por los períodos. El valor S_t es la que minimiza la función pérdida más la esperanza del costo futuro y el valor s_t es el mayor valor posible menor que S_t, tal que la pérdida de s_t es mayor a la pérdida en S_t más el costo fijo de pedir. (Ver código MDP.py). Luego la política a aplicar en los T periodos corresponden a los mismos deducidos en la parte anterior. En la figura 1 se muestra el valor de s y S según el período.

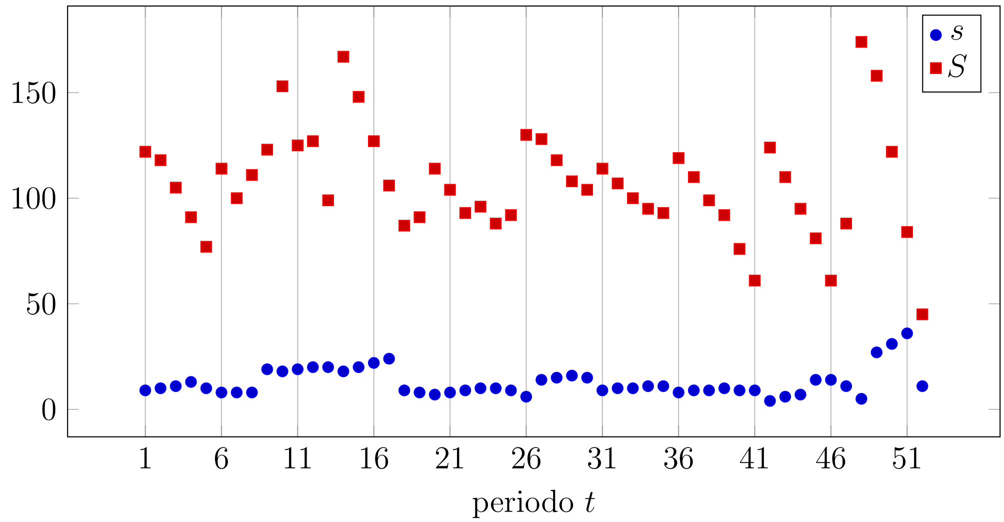

Figura 1: Política (s, S)

- Se implementa a continuación una serie de políticas heurísticas para comparar sus desempeños.

Conservadora

Esta heurística se basa en encargar hasta la capacidad, 200 unidades. Además, las demandas que hay en cada periodo están aproximadas como la esperanza de las distribuciones respectivas, es decir, $\mu_{m(t)}$.

Newsvendor

Con esta heurística se optimiza considerando solo 1 período que es el actual y con un inventario inicial. El óptimo en este caso también corresponde a aplicar una política de forma (s, S). Se asume que cada semana se parte con un inventario inicial de $y_1 = 30$ para aplicar la metodología.

Futuro promedio

Esta heurística tiene la misma lógica que el apartado dos, donde se calcula el óptimo. La diferencia es que las demandas son deterministas, al ser consideradas como conocidas, son la media de las distribuciones mensuales.

Dicho lo anterior, se obtienen los siguientes resultados:

Cuadro 2: Resultados de Heurísticas

Heurística	Costo Esperado	Valor de la Opt. Est. (VOE)	VOE (%)	Tiempo de ejec. [seg]
Conservadora	$165861	$106502.783	179.423	0.001
Newsvendor	$47255	-$12103	25.6	1.29
Futuro prom.	$ 57131	-$2227.217	3.75	1.801

Del Cuadro 2 se puede observar que la heurística más alejada del costo óptimo es la conservadora, lo cual es lógico, dado que en cada periodo existe demanda y se desencadenan nuevos pedidos. Así, solo con el costo fijo acumulado de todas las semanas se llega $ 78000. Además, esta es la heurística que corre en menor tiempo, debido a su simplicidad, ya que no revisa todas las posibles decisiones, solo escoge aquel pedido que llena la capacidad de la bodega.

Con respecto al problema del vendedor de diarios se utilizó una aproximación del valor s que debería ser el óptimo debido a que en cada iteración no existía un valor tal que $G(s)$ fuera mayor a $G(S) + K$. Por esta razón se modificó el problema para que s fuera exactamente una unidad menor que S.

Por último, la heurística del futuro promedio resultó ser la mejor, dado que exploraba muchas más opciones al considerar el futuro, sin embargo, por la misma razón demora más tiempo que las anteriores. Además, simplifica más el problema que el algoritmo que optimiza, debido a que le quita la estocasticidad a los datos, al asumir una demanda determinista (el promedio de la distribución) y no cada caso con su respectiva probabilidad.

Por otra parte, se puede apreciar que en el caso de las heurísticas de Newsvendor y futuro promedio, resultaron costos esperados menores que los del algoritmo óptimo, esto se puede deber a que se trabajó directamente con la demanda aproximada como $\mu_{m(t)}$ y no se revisaron todos los casos con las respectivas probabilidades. De modo que, no se alcanzara a captar el efecto de costos grandes bajo dicha simplificación.

- Se corre 1000 simulaciones de demandas anuales y se evalúa el desempeño en un caso donde se conoce la demanda de antemano y aprovechando esta información se toman las decisiones y por otro lado aplicando la política (s, S).

Política (s, S)

Si se aplica la política (s, S) con las mismas demandas se tiene otro resultado. El proceso es el siguiente.

Algorithm 1 Evaluación de política (s, S)

Input: $y_1 \leftarrow 30$
  for escenario $\in$ escenarios **do**
    for periodo $\in$ periodos **do**
      if $y_t < s_t$ **then**
        $y'_t = S_t - y_t$
      else if $y_t > s_t$ **then**
        $y'_t = y_t$
      end if
      $y_{t+1} = (y'_t - D_t)^+$
      $C_i = K\mathbb{I}_{y'_t > y_t} + c(y'_t - y_t) + q(D_t - y'_t)^+ + h(y'_t - D_t)^+$
    end for
  end for

En la figura 2 se muestra la distribución de costos finales aplicando política (s, S) dados las demandas simuladas en las 1000 instancias.

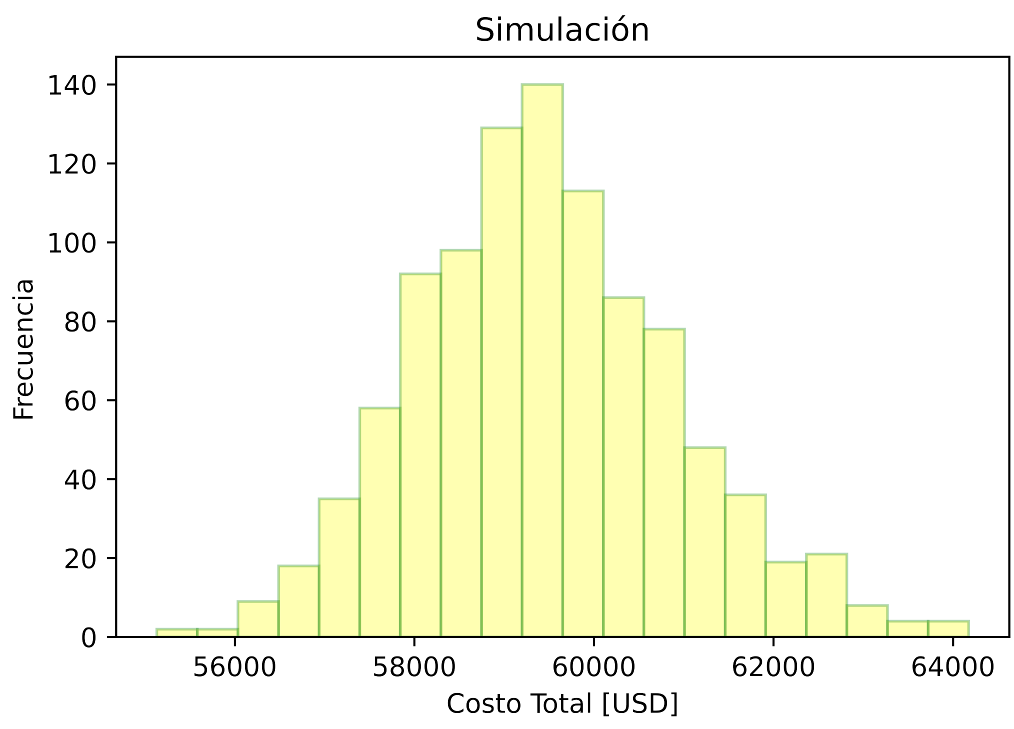

Figura 2: Caption

Información Perfecta

En cambio, un planificador con información perfecta (PI) planificaría minimizando la suma de los costos de cada período eligiendo la cantidad a pedir de cada periodo.

$$\min_{y'_t \geq y_t} \left\{ \sum_{t=1}^{T} r_t(y'_t|y_t) \right\}$$

Donde:

$$\begin{aligned} y_1 &= 30 \\ r_t(y'_t|y_t) &= K\mathbb{I}_{y'_t > y_t} + c(y'_t - y_t) + q(D_t - y'_t)^+ + h(y'_t - D_t)^+ \\ y_{t+1} &= \max\{y'_t - D_t, 0\} \end{aligned}$$

Dicho lo anterior Δ^m no puede ser negativo, la política óptima (s, S) funciona bien en promedio pero no puede ganarle a aquella optimización que considera información perfecta anticipando el futuro (ajustando la solución al escenario). Así, $C^m \geq C^m_{PI}$, este último funciona como una cota inferior y Δ^m es la pérdida de objetivo asociado a no saber el futuro.

- Considerando Lead Time:

  Si se toma en cuenta la posibilidad de tener un Lead Time, entonces al momento de tomar las decisiones, ya no se tiene que considerar un solo periodo, sino que la probabilidad de que lo que se necesita llegue en uno, dos o tres periodos más adelante.

  La consideración principal tiene que ver con que, con el inventario que se tiene, existe un 50 % de probabilidad de que se tenga que considerar 2 periodos de demanda seguidos, un 35 % de probabilidad de que se tengan que considerar 3 periodos de demanda seguidos o, finalmente, un 15 % de probabilidad de que se tengan que considerar 4 periodos de demanda seguidos. Lo que se ve reflejado en que en cada periodo se debe considerar la probabilidad de que llegue algún pedido previo.

  Modelo

  El nuevo estado post decisión sería:

  $$y'_t = y_t + \mathbb{I}_{\text{Si es que llega en } t}(x_{t-1} + x_{t-2} + x_{t-3})$$

  Se modela el problema que planifica una política de decisión que minimiza el costo total esperado considerando el *lead time.*

  Etapas $\{1, 2, ..., 52\}$

  Estados $y_t \in \{0, 1, ..., 200\}$

  Acciones $\mathbb{X}(y) = \left\{0, 1, ..., 200 - y - \sum_{i=1}^{3} \mathbb{X}_{t-i}(t_{t-i})\right\}$

Probabilidad de que llegue en i semanas mas P_i

Costo inmediato:

$$r(y,x,D) = \underbrace{1500 \cdot \mathbb{I}_{x>0} + 25 \cdot x}_{g(x)} + \underbrace{90 \cdot (D - y - \sum_{i=1}^{3} x_{t-i} P_i)^+}_{q} + \underbrace{6 \cdot (y + \sum_{i=1}^{3} x_{t-i} P_i - D)^+}_{h}$$

En promedio:

$$r(y,x) = \underbrace{1500 \cdot \mathbb{I}_{x>0} + 25 \cdot x}_{g(x)} + \underbrace{90 \cdot \mathbb{E}_{D_t}(D - y - \sum_{i=1}^{3} x_{t-i} P_i)^+}_{q} + \underbrace{6 \cdot \mathbb{E}_{D_t}(y + \sum_{i=1}^{3} x_{t-i} P_i - D)^+}_{h}$$

Probabilidad de transición de estado:

$$p(y_{t+1}|y + \sum_{i=1}^{3} x_{t-i}P_i) = \begin{cases} f(y + \sum_{i=1}^{3} x_{t-i}P_i - y_{t+1}) & \text{si } 0 < y_{t+1} \leq y + x \\ \sum_{k \geq y + \sum_{i=1}^{3} x_{t-i}P_i} f(k) & \text{si } y_{t+1} = 0 \\ 0 & \text{e.o.c} \end{cases} \tag{6}$$

Además, dado que D sigue una distribución Uniforme Discreta: $X \sim U(\mu_{m(t)} - \delta_{m(t)}, \mu_{m(t)} + \delta_{m(t)})$. La probabilidad de que haya demanda x y la esperanza vienen dadas respectivamente por las Ecuaciones (2) y (3)

$$P\left[X = x\right] = \frac{1}{2 \cdot \delta_{m(t)} + 1} \tag{7}$$

$$\mathbb{E}(X) = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \mu_{m(t)} \tag{8}$$

Por lo tanto, el costo inmediato en promedio y la probabilidad de estado quedán dadas por las Ecuaciones (4) y (5)

$$r(y,x) = 1500 \cdot \mathbb{I}_{x>0} + 25 \cdot x + 96 \cdot \mu_{m(t)} \tag{9}$$

$$p(y_{t+1}|y + \sum_{i=1}^{3} x_{t-i}P_i) = \begin{cases} \frac{1}{2 \cdot \delta_{m(t)} + 1} & \text{si } 0 < y_{t+1} \leq y + \sum_{i=1}^{3} x_{t-i}P_i \\ 1 - \frac{y + \sum_{i=1}^{3} x_{t-i}P_i}{2 \cdot \delta_{m(t)} + 1} & \text{si } y_{t+1} = 0 \\ 0 & \text{e.o.c} \end{cases} \tag{10}$$