PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
Departamento de Ingeniería Industrial y de Sistemas
Departamento de Ingeniería de Transporte y Logística

Curso: ICS3105
Semestre: 02-2021
Profesor: Mathias A. Klapp

Tarea 2 - Optimización Dinámica

Fecha de entrega: 15 de Octubre 2021

Instrucciones: Se recomienda usar Latex. Evite respuestas largas y sólo responda lo que se pregunta. El corrector puede solicitar revisar su código fuente, pero su respuesta debe ir completamente en el reporte. Se prohíbe discutir la tarea fuera del grupo.

Inventario estocástico

Considere un problema de inventario estocástico como el desarrollado en clase, donde un tomador de decisiones (*i.e.*, usted) debe ordenar productos al comienzo de cada semana (después de observar inventario) a lo largo de un horizonte de ejecución discreto de $T = 52$ semanas. Específicamente, suponga que inicia en la semana 1 con $y_1 = 30$ unidades almacenadas y que su bodega posee capacidad para $Q = 200$ unidades. La dinámica de eventos en cada semana $t \in \{1, \dots, T\}$ es la siguiente:

1. Usted observa el inventario del producto en bodega $y_t \in \{0, 1, 2, \dots, Q\}$.
2. Luego, decide ordenar $x_t \in \{0, 1, \dots, Q - y_t\}$ unidades de producto a un costo de ordenamiento $\$1500 + 25 \cdot x_t$. El inventario ordenado es repuesto instantáneamente hasta $y'_t = y_t + x_t$.
3. Observa la demanda D_t de la semana t, donde D_t es una variable aleatoria independiente. El cliente recibe todo el producto demandado disponible (es decir, $\min(D_t, y'_t)$) y el sobrante se almacena para la siguiente semana. El costo por unidad-semana en inventario es $h = \$6$ (a pagar si un producto pasa en inventario de una semana a la otra) y el costo por unidad demandada no satisfecha es $q = \$90$.

Asuma que la demanda semanal D_t es discreta y distribuida uniforme en $\left\{\mu_{m(t)} - \delta_{m(t)}, \dots, \mu_{m(t)} + \delta_{m(t)}\right\}$, donde $m(t)$ es el mes asociado a esa semana t, μ_m es un parámetro de demanda esperada y δ_m es un error. Los valores de los parámetros se encuentran tabulados a continuación para cada mes:

mes m	μ_m	δ_m	Semanas en mes
1	20	4	1–4
2	18	8	5–8
3	30	5	9–13
4	32	7	14–17
5	17	5	18–22
6	18	8	23–26
7	25	7	27–30
8	19	10	31–35
9	18	6	36–39
10	15	2	40–44
11	20	1	45–48
12	40	15	49–52

Preguntas:

- **(5 puntos)** Modele el problema que planifica una política de decisión que minimiza el costo total esperado como un Proceso de Decisión Markoviana (MDP).

  La formulación MDP se expone a continuación:

  Etapas: 52 etapas, una por cada semana.

  Decisión (x_t): cantidad de producto a pedir en la etapa t.

  $$\mathbb{X}_t(s_t) = \{x_t \in \mathbb{Z}^+ : x_t \leq Q - s_t\} \quad \forall t \in \{1, ..., 52\}$$

  Con Q la capacidad máxima de la bodega. En este caso en particular $Q = 200$.

  Estados: nivel de inventario al inicio de la etapa t.

  $$\mathbb{S}_k^t = \{0, ..., Q\} \quad \forall t \in \{1, ..., 52\}$$

  Función de transición: la probabilidad de que dada una decisión x_t en el estado s_t llegue al estado s_{t+1}.

  - Si $s_{t+1} = 0 \rightarrow \mathbb{P}(s_{t+1}|x_t, s_t) = \mathbb{P}(D_t \geq x_t + s_t)$
  - Si $0 < s_{t+1} \leq x_t + s_t \rightarrow \mathbb{P}(s_{t+1}|x_t, s_t) = \mathbb{P}(D_t = s_{t+1} - (x_t + s_t))$
  - Si $s_{t+1} > x_t + s_t \rightarrow \mathbb{P}(s_{t+1}|x_t, s_t) = 0$

  Costo inmediato: costo en la etapa t de tomar la decisión x_t en el estado s_t.

  $$r_t(x_t, s_t) = 1500 \cdot \mathbb{I}_{x_t > 0} + 25 \cdot x_t + 90 \cdot \mathbb{E}_{D_t}[(D_t - (x_t + s_t))^+] + 6 \cdot \mathbb{E}_{D_t}[(x_t + s_t - D_t)^+]$$

  - $\mathbb{E}_{D_t}[(D_t - (x_t + s_t))^+] = \sum_{d_t = \min\{\mu_t - \delta_t, x_t + s_{t+1}\}}^{\mu_t + \delta_t} (d_t - (x_t + s_t)) \cdot \frac{1}{2 \cdot \delta_t + 1}$
  - $\mathbb{E}_{D_t}[(x_t + s_t - D_t)^+] = \sum_{d_t = \mu_t - \delta_t}^{\max\{\mu_t - \delta_t, x_t + s_t - 1\}} (x_t + s_t - d_t) \cdot \frac{1}{2 \cdot \delta_t + 1}$
  - El valor de $d_t = x_t + s_t$ queda implícito en las relaciones debido a que no aporta en ninguno de los dos costos, pero es parte de las probabilidades totales.

  Objetivo: minimizar el costo esperado desde la etapa 1, donde $s_1 = 30$.

  $$\min_{x_t \in \mathbb{X}_t(s_t) \forall t \in T} \sum_{t=1}^{T} r_t(x_t, s_t)$$

- **(10 puntos)** Obtenga el mínimo costo esperado del sistema a lo largo de los 12 meses (optimice sin asumir una estructura (s, S) de antemano). Calcule el tiempo de cómputo requerido para evaluar el mínimo costo esperado y obtener un política de decisión óptima.

  Verifique que efectivamente la política óptima obtenida posee una estructura (s, S) para cada semana.

  Para resolver el problema se utilizó lo siguiente:

  - **Periodo terminal**: $V_{53}(y) = 0, \quad \forall y \in \{0, ..., Q\}$
  - **Recursión**: $V_t(y) = \min_{x_t \in \mathbb{X}_t}\{r(x_t, y) + \mathbb{E}(V_{t+1}(y + x_t))\}, \quad \forall t \in T, \forall y \in \{0, ..., Q\}$

  El código demora cerca de 11.5 segundos en ejecutarse y entrega una política con un costo esperado de 59358.22.

- **(10 puntos)** Optimice el problema de nuevo, pero asuma de antemano que busca una política (s, S) para reducir al máximo el tiempo de cómputo y el espacio de búsqueda. Explique que hizo y compare el tiempo de cómputo con respecto al anterior.

  Para obtener la política (s, S) óptima, se utiliza el procedimiento visto en clases:

  - Para todo $y \in \{0, ..., Q\} : C_(T + 1) = 1$

- Para todo $t = 1, ..., T$:
  - Calcular S_t que minimiza $G_t(y) = G(y) + \mathbb{E}_D[C_{t+1}(y - D_t)^+]$, donde $G(y) = 25 \cdot y + 90 \cdot \mathbb{E}_D[(D - y)^+] + 6 \cdot \mathbb{E}_D[(y - D)^+]$.
  - Calcular $s_t \leq S_t$ que cumpla con $G_t(s_t) = G_t(s_t) + 1500$
  - Para todo $y = 0, .., Q$ calcular

$$C_t(y) = \begin{cases} G_t(S_t) + K - cy & \text{si } y \leq s_t \\ G_t(y) - cy & \text{si } y > s_t \end{cases} \tag{1}$$

Al comparar el desempeño de ambos métodos, se observa que llegan al mismo costo esperado para el inventario inicial del enunciado y retornan políticas equivalentes, ya que MDP retorna las combinaciones para cada par inventario-semana mientras que (s, S) entrega las reglas de decisión. Sin embargo, el método (s, S) demora un tiempo considerablemente menor que el método MDP en ejecutarse.

El código demora cerca de 0.26 segundos en ejecutarse y entrega una política con un costo esperado de 59358.22

- **(15 puntos)** Evalúe el costo esperado de las siguientes políticas heurísticas y compare porcentualmente cada valor con el mínimo costo esperado.
  - Política conservadora que llena la bodega en cada etapa.

    Para esta política la función de costos (r_t) se calcula de la misma manera. Sin embargo, al momento de tomar la decisión no se busca la decisión que para un tiempo e inventario minimice dicha función, sino que se obliga a que la decisión (x) siempre sea ordenar $x = Q - inv$, es decir, llenar la bodega.

    El código demora cerca de 0.14 segundos en ejecutarse y entrega una política con un costo esperado de 165861.0
  - Política *Newsvendor*, es decir, una política miope que optimiza el costo inmediato en cada etapa sin mirar a futuro.

    En la política de Newsvendor al momento de tomar la decisión el cost-to-go (r_t) es míope al futuro, es decir, no considera el V_{t+1}.

    El código demora cerca de 15.12 segundos en ejecutarse y entrega una política con un costo esperado de 94576.62
  - Política óptima del futuro promedio, es decir, una política que decide cuanto ordenar en cada etapa t resolviendo un problema determinista asumiendo valores promedio de demanda futura.

    Siguiendo la política de futuro promedio, al momento de tomar la decisión en el cost-to-go (r_t) se considera el promedio de la demanda y no todo el soporte de su distribución. No obstante, al evaluar el costo de la decisión si se considera la aleatoriedad de la demanda en el cost-to-go.

    Por ello, se crearon dos funciones de cost-to-go:
    - **Deterministico**: Aquí el cost-to-go considera el promedio de la demanda. A partir de esta, para cada semana e inventario, se busca la decisión (x) que lo minimiza.
    - **Estocástico**: Aquí el cost-to-go es el mismo del MDP inicial. Es decir, considera todo el soporte de la demanda con sus respectivas probabilidades para tomar la decisión. Usando esta función se evalúa el costo de la decisión (x).

    El código demora cerca de 17.2 segundos en ejecutarse y entrega una política con un costo esperado de 59473.82.

    Finalmente, los errores porcentuales respecto a la política sS.
- **(10 puntos)** Ahora estimaremos el "costo de la incertidumbre" mediante simulación, es decir, la diferencia entre el costo esperado del sistema real con dinamismo de información e incertidumbre y el costo esperado de un sistema análogo donde el planificador conoce toda la demanda de antemano antes de optimizar (*i.e.*, un planificador clarividente). Para ello:

Política	Costo	Variación porcentual
Newsvendor	94576.62	59.33%
Conservadora	165861.0	179.42%
Futuro promedio	59473.82	0.19%

- Simule una muestra de M vectores de demanda $\{\mathbf{D}^1, \ldots, \mathbf{D}^M\}$.
- Para cada m-ésima realización anual de demanda $\mathbf{D}^m = \{D_1^m \ldots, D_T^m\}$ con $m \in \{1, \ldots, M\}$ haga lo siguiente:
  * Aplique la política de decisión óptima obtenida anteriormente y calcule el costo ejecutado C^m del sistema para la realización $\mathbf{D}^m$.
  * Ahora optimice el ordenamiento de productos, pero asumiendo que el planificador conoce $\mathbf{D}^m$ de antemano. Llame a este costo optimizado C_{PI}^m (PI viene de *Perfect Information*).
  * Calcule la diferencia de costos $\Delta^m = C^m - C_{PI}^m$. ¿Puede ser negativo?
- Calcule $\bar{\Delta}$ el promedio de los Δ^m y su desviación muestral S_Δ y estime un intervalo de confianza para el "costo esperado de la incertidumbre" con una precisión relativa al 95% de confianza.

  Discuta y comente su resultado.

  Hay que considerar que Δ^m es no negativo para toda instancia m, porque C_{PI}^m es el menor costo que puede tener cualquier política al ser la decisión óptima, dado un escenario. En el mejor de los casos, la política sS puede tener el mismo valor que el costo optimizado para una instancia en particular, sin embargo, en otros escenarios C^m es mayor, debido a que la política se obtiene a partir minimizar el promedio del costo y no un escenario en particular.

  IC = [2478.57, 2847.42].

  Interpretando el significado del intervalo de confianza, notamos que con probabilidad de 95% el promedio del costo de la incertidumbre se encuentra entre 2478.57 y 2847.42. El promedio fue de 2662.5.

  Nota: No se espera que los intervalos de confianza sean idénticos a los presentados en esta pauta.

- **(10 puntos)** Ahora supongan que cada vez que ordena productos existe un *Lead Time* L_t, es decir, un tiempo de entrega de la órden que tarda una semana con probabilidad 50%, dos semanas con probabilidad 35% y tres semanas con probabilidad 15%. Modele el problema de inventario estocástico con esa nueva consideración y optimice el problema que minimiza el costo total esperado.

  Para este ejercicio se utilizaron los siguientes supuestos.

  - **Supuesto 1**: No se pueden botar unidades a la basura. Para ello, se evita que se sobrepase la demanda, obligando a que $x_t \leq Q - inventario_t$.
  - **Supuesto 2**: No se puede hacer una nueva orden si es que la anterior no ha llegado.
  - **Supuesto 3**: Las ordenes llegan al principio del periodo.

  La formulación MDP se expone a continuación:

  Etapas: 52 etapas, una por cada semana.

  Decisión: cantidad de producto a ordenar (pedir) en la etapa t x_t. $\mathbb{X}_t(s_t) = \{x_t \in \mathbb{Z}^+ : x_t \leq Q - s_t\} \quad \forall t \in \{1, ..., 52\}$. Con Q la capacidad máxima de la bodega. En este caso en particular $Q = 200$.

  Estados: nivel de inventario al inicio de la etapa t, tamaño de la orden en proceso y tiempo que queda para que la orden llegue. $\mathbb{S}_t = \{(inv, O, k) : \forall inv \in \{0, ..., Q\}, O \in \{0, ..., Q - inv\}, t \in \{1, ..., 52\}\}$.

  Función de Transición: La función de transición depende del valor de k y de D_t.

- k = 0 y O = 0:

  Con leadtime 1

  * Si $s_t = (y_1, 0, 0) \rightarrow \mathbb{P}((y_2, 0, 0)|x_t, s_t) = 0,5 \cdot \mathbb{P}(y_1 - D_t + x_t = y_2) \quad \forall y_2 \neq 0$
  * Si $s_t = (y_1, 0, 0) \rightarrow \mathbb{P}((0, 0, 0)|x_t, s_t) = 0,5 \cdot \mathbb{P}(y_1 + x_t \leq D_t)$

  Con leadtime 2

- Si $s_t = (y_1, 0, 0) \rightarrow \mathbb{P}((y_2, x_t, 1)|x_t, s_t) = 0,35 \cdot \mathbb{P}(y_1 - D_t = y_2) \quad \forall y_2 \neq 0$
- Si $s_t = (y_1, 0, 0) \rightarrow \mathbb{P}((0, 0, 1)|x_t, s_t) = 0,35 \cdot \mathbb{P}(y_1 \leq D_t)$

  Con leadtime 3

  * Si $s_t = (y_1, 0, 0) \rightarrow \mathbb{P}((y_2, x_t, 2)|x_t, s_t) = 0,15 \cdot \mathbb{P}(y_1 - D_t = y_2) \quad \forall y_2 \neq 0$
  * Si $s_t = (y_1, 0, 0) \rightarrow \mathbb{P}((0, 0, 2)|x_t, s_t) = 0,15 \cdot \mathbb{P}(y_1 \leq D_t)$

- k = 1:
  * Si $s_t = (y_1, O, 1) \rightarrow \mathbb{P}((y_2, 0, 0)|x_t, s_t) = \mathbb{P}(y_1 - D_t + O = y_2) \quad \forall y_2 \neq 0$
  * Si $s_t = (y_1, 0, 1) \rightarrow \mathbb{P}((0, 0, 0)|x_t, s_t) = \mathbb{P}(y_1 + x_t + O \leq D_t)$

- k = 2
  * Si $s_t = (y_1, O, 2) \rightarrow \mathbb{P}((y_2, O, 1)|x_t, s_t) = \mathbb{P}(y_1 - D_t = y_2) \quad \forall y_2 \neq 0$
  * Si $s_t = (y_1, O, 2) \rightarrow \mathbb{P}((0, O, 1)|x_t, s_t) = \mathbb{P}(y_1 \leq D_t)$

Costo Inmediato: costo en la etapa t de tomar la decisión x_t en el estado s_t.

$$r_t(x_t, s_t) = 1600 \cdot \mathbb{I}_{x_t > 0} + 20 \cdot x_t + 75 \cdot \mathbb{E}_{D_t}[(D_t - (x_t + s_t))^+] + 4 \cdot \mathbb{E}_{D_t}[(x_t + s_t - D_t)^+]$$

- $\mathbb{E}_{D_t}[(D_t - (x_t + s_t))^+] = \sum_{d_t = min\{\mu_t + \delta_t, x_t + s_t + 1\}}^{\mu_t + \delta_t} (d_t - (x_t + s_t)) \cdot \dfrac{1}{2 \cdot \delta_t + 1}$
- $\mathbb{E}_{D_t}[(x_t + s_t - D_t)^+] = \sum_{d_t = \mu_t - \delta_t}^{max\{\mu_t - \delta_t, x_t + s_t - 1\}} (x_t + s_t - d_t) \cdot \dfrac{1}{2 \cdot \delta_t + 1}$
- El valor de $d_t = x_t + s_t$ queda implícito en las relaciones debido a que no aporta en ninguno de los dos costos, pero es parte de las probabilidades totales.

Objetivo: minimizar el costo esperado desde la etapa 1, donde $s_1 = 30$.

$$\min_{x_t \in \mathbb{X}_t(s_t) \forall t \in T} \sum_{t=1}^{T} r_t(x_t, s_t)$$

El costo obtenido fue de 62930.23.