Stochastic Scheduling

Models Real World Uncertainty

- processing times
- arrivals
- machine availability
- ...

Our Model:

- Distribution over job data known in advance.
- Realization only known when job arrives/completes or when it can be inferred.

Example:

\[ p_j = \begin{cases} 
1 & \text{Pr} = 1/2 \\
3 & \text{Pr} = 1/2 
\end{cases} \]

After 1 unit of time, if the job doesn’t complete, we know that it will take 3 units.
Example

\[ p_1 = \begin{cases} 1 & \text{Pr} = 1/2 \\ 9 & \text{Pr} = 1/2 \end{cases} \]

\[ E[C_{p_1}] = \frac{1}{2} (1+9) = 5 \]

\[ p_2 = \begin{cases} 4 & \text{Pr} = 1/4 \\ 6 & \text{Pr} = 1/2 \\ 8 & \text{Pr} = 1/4 \end{cases} \]

\[ E[C_{p_2}] = \frac{1}{4} (1+4) + \frac{1}{2} (6) + \frac{1}{4} (8) = 6 \]

Problem: \( 1||\sum C_j \)

Question: What is the right algorithm? Is there still a simple ordering rule

\[
\text{EPT} \quad \text{is the right alg}\]

\[
E[C_j] = \frac{1}{8} (1+5) + \frac{1}{8} (9+13) + \frac{1}{4} (1+7) + \frac{1}{4} (9+15) + \frac{1}{8} (1+9) + \frac{1}{8} (9+17) = 16
\]
Comparing random variables

- Density Function: \( f(x) \)
- Distribution Function: \( F(x) = P(X \leq t) = \int_0^t f(x) \, dx \)

Definitions of \( X_1 \geq X_2 \)
- Larger in Expectation: \( E(X_1) \geq E(X_2) \)
- Stochastically larger: \( \forall t : P(X_1 > t) \geq P(X_2 > t) \)
- Almost surely larger: \( P(X_1 \geq X_2) = 1 \quad \\Rightarrow (X_2 \geq t) \rightarrow \emptyset \)
Another example, $P\|C_{\text{max}}$

Case 1: $p_1 = p_2 = 1$

Case 2: $p_1 = 1$

$$p_2 = \begin{cases} 
0 & \text{Pr} = 1/2 \\
2 & \text{Pr} = 1/2
\end{cases}$$

Case 3:

$$p_1 = p_2 = \begin{cases} 
0 & \text{Pr} = 1/2 \\
2 & \text{Pr} = 1/2
\end{cases}$$

Case 4: $p_1, p_2$ both uniform in $[0, 2]$. 
Another example, $P\|C_{\text{max}}$

Case 1: $p_1 = p_2 = 1$

Case 2: $p_1 = 1$

$$p_2 = \begin{cases} 0 & \Pr = 1/2 \\ 2 & \Pr = 1/2 \end{cases}$$

Case 3:

$$p_1 = p_2 = \begin{cases} 0 & \Pr = 1/2 \\ 2 & \Pr = 1/2 \end{cases}$$

Case 4: $p_1, p_2$ both uniform in $[0, 2]$.

$$C_{\text{max}} = \max(p_1, p_2)$$

$$\mathbb{E}[C_{\text{max}}] = \mathbb{E}\left[\max(p_1, p_2)\right] = \max(\mathbb{E}[p_1], \mathbb{E}[p_2])$$

$$\mathbb{E}[C_{\text{max}}] = \frac{1}{2}(0) + \frac{1}{2}(1) + \frac{1}{2}(2) + \frac{1}{2}(2) = \frac{3}{2}$$
Another example, $P||C_{\text{max}}$

Case 1: $p_1 = p_2 = 1$

Case 2: $p_1 = 1$

$$p_2 = \begin{cases} 
0 \text{ Pr } = 1/2 \\
2 \text{ Pr } = 1/2 
\end{cases}$$

Case 3:

$$p_1 = p_2 = \begin{cases} 
0 \text{ Pr } = 1/2 \\
2 \text{ Pr } = 1/2 
\end{cases}$$

Case 4: $p_1, p_2$ both uniform in $[0, 2]$.
Another example, $P||C_{\text{max}}$

Case 1: $p_1 = p_2 = 1$

Case 2: $p_1 = 1$

\[
p_2 = \begin{cases} 
0 & \text{Pr} = 1/2 \\
2 & \text{Pr} = 1/2 
\end{cases}
\]

Case 3:

\[
p_1 = p_2 = \begin{cases} 
0 & \text{Pr} = 1/2 \\
2 & \text{Pr} = 1/2 
\end{cases}
\]

Case 4: $p_1, p_2$ both uniform in $[0, 2]$.

\[
\mathbb{E}\left[\max\left(p_1, p_2\right)\right] = \frac{4}{3}
\]
Two R.V. $Y_1, Y_2$ chosen from $[0, 2]$  

$$E[\max (Y_1, Y_2)] = \int_0^2 t \cdot \Pr (Z = t) \, dt$$

$$= \int_0^2 (1 - \Pr (Z \leq t)) \, dt$$

$$= \int_0^2 (1 - \frac{t^2}{4}) \, dt = \frac{y}{3}$$
1) Make Dough
   Mixing, melting, knead 7
   3 4 7
   15, sit for 30 min

2) Make filling
   - Cutting onions × 10 min
   - Cooking phet 30 min
   - Ake onions > 10 min
     mixing

3) Roll dough & cut circles ~ 10 min

4) Assemble peoples & fill ~ 10 min

5) Boil in batches of 12 ~ 5 mins/batch

6) Fry in batches of 6 ~ 5 mins/batch
Objective Values

1. $C_{\text{max}} = 1$
2. $C_{\text{max}} = 3/2$
3. $C_{\text{max}} = 3/2$
4. $C_{\text{max}} = 4/3$
Different Models of Stochastic Scheduling

Models of Knowledge

• static: Choose order of jobs based on distribution only
• dynamic: Choose order of jobs based on knowledge gained when running

Also consider Preemption vs. Non-preemption

Example:

• $1||\sum U_j$

• 3 jobs with same distribution:

\[
p_j = \begin{cases} 
2 & \text{Pr} = 1/2 \\
8 & \text{Pr} = 1/2 
\end{cases}
\]

\[
d_j = \begin{cases} 
1 & \text{Pr} = 1/2 \\
5 & \text{Pr} = 1/2 
\end{cases}
\]

What is the expected objective value for:

• static non-preemptive
• dynamic non-preemptive
• dynamic preemptive
static non-preemptive
order jobs based on distribution info
- 1, 2, 3 order

\[ p = (2, 8) \quad d = (1, 5) \]

Cannot execute

\[ d_1 = 1 \quad d_2 = 8 \quad d_3 = 15 \]

Jobs complete on time

\[ \frac{1}{16} + \frac{1}{16} + 2 \left( \frac{1}{16} \right) + \frac{1}{8} = \frac{3}{8} \]
dynamic non-preemptive
(can't preempt, but we can learn)

\[ t = 2 \]

\[ p_1 = 2 \]

\[ p_2 \]

\[ d_1 = 1 \]

or \[ d_1 = 5 \]

\[ d_2 \]

\[ d_3 \]

\[ \frac{d_2}{d_3} = 1 \]

\[ d_2 \sim 5 \]

\[ d_3 \sim 5 \]

(we don't know \((p_2, p_3)\))

run a job with \(d = 5\) if possible

\[ \text{first job to run} \]

\[ \Pr \left( \sum Z_j \right) = \Pr (Z_{10}) + \Pr (Z_{20}) + \Pr (Z_{23}) \]

\[ = \Pr (Z_{10} = 1) + \Pr (Z_{20} = 1) + \Pr (Z_{23} = 1) \]

\[ = \Pr (p_1 = 2) \cdot \Pr (d_1 = 5) \]

\[ + \Pr (p_2 = 2) \cdot \Pr (p_{12} = 2) \cdot \Pr (d_1 = 5) \]

\[ \approx \frac{1}{5} \cdot \frac{1}{5} + \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{3}{9} = \frac{1}{4} + \frac{3}{16} = \frac{7}{16} \]
**Dynamic Preemptive Scheduling**

- For $t = 1$ or $2$: Job 1 not running anymore.
- For $t = 3$ or $4$:
  - Pick a job $m_j$ if possible.

**Example Calculation**

$$E(2) + E(2_{12}) + E(2_3)$$

$$= \Pr(D = 5) \cdot \Pr(P = 2) + \Pr(D_{10} = 5) \cdot \Pr(P_{21} = 2) + \Pr(D_{10} = 1) \cdot \Pr(P_{23} = 2) \cdot \Pr(D_{13} = 5)$$

$$= \frac{1}{2} \cdot \frac{1}{2} + \frac{3}{4} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{4} + \frac{3}{8} + \frac{1}{16} = \frac{11}{16}$$
Another Example

- **Problem:** $1|\text{pmtn}|\sum C_j$
- **Jobs:**

  - $p_1 = \begin{cases} 1 & \text{Pr} = 1/2 \\ 3 & \text{Pr} = 1/2 \end{cases}$
  - $p_2 = \begin{cases} 2 & \text{Pr} = 1/2 \\ 4 & \text{Pr} = 1/2 \end{cases}$
  - $p_3 = \begin{cases} 1 & \text{Pr} = 1/2 \\ 7 & \text{Pr} = 1/2 \end{cases}$