Reading for This Lecture

- C&T Chapter 16
Monte Carlo Methods

• We now consider the following very general stochastic program

\[ \min_{x \in S} \{ f(x) \equiv \mathbb{E}[F(x, \xi)] \}, \quad (1) \]

where \( \xi \) is a random vector on the probability space \((\Omega, P)\), as usual.

• The standard two-stage stochastic program with recourse we have been considering is a special case of (1).

  \begin{itemize}
  \item \( S \equiv \{ x \mid Ax = b, \ x \geq 0 \} \)
  \item \( f(x) \equiv c^T x + Q(x) \)
  \item \( Q(x) \equiv \mathbb{E}[Q(x, \omega)] \)
  \item \( Q(x, \omega) \equiv \min_{y \geq 0} \{ q(\omega)^T y \mid W y = h(\omega) - T(\omega)x \} \)
  \end{itemize}

• The methodology we consider here holds for more general SPs, however.


**Sampling**

- Instead of solving (1), we solve an approximating problem.
- Let $\xi^1, \ldots, \xi^N$ be $N$ independent realizations of the random variable $\xi$:

$$\min_{x \in S} \{ \hat{f}_N(x) \equiv N^{-1} \sum_{j=1}^{N} F(x, \xi^j) \}.$$

- $\hat{f}_N(x)$ is the *sample average* function.
- $\hat{f}_N(x)$ is an unbiased estimator of $f(x)$, i.e.,

$$\mathbb{E}[\hat{f}_N(x)] = f(x)$$
Sample Variance

• Since $\xi^j$ are independent, we can estimate $\text{Var}(\hat{f}_N(x))$.

• This is done using the sample variance:

$$\hat{\sigma}^2(x) = \frac{1}{N(N - 1)} \sum_{j=1}^{N} [(F(x, \xi^j) - \hat{f}_N(x))^2]$$
Statistics Break

- Let \( \chi_1, \chi_2, \ldots, \chi_n \) be independent, identically distributed (iid) random variables.

- Let \( S_n = \sum_{i=1}^{n} \chi_i \).

- Assume \( \mu \equiv \mathbb{E}[|\chi_i|] < \infty \).

**Weak Law of Large Numbers**

\[
\lim_{n \to \infty} P\left( \left| \frac{S_n}{n} - \mu \right| \geq \delta \right) = 0 \quad \forall \delta > 0
\]
Strong Law of Large Numbers

\[ \lim_{n \to \infty} \frac{S_n}{n} \to \mu \quad \text{Almost surely} \]

- *Almost surely* means “with probability 1”, or..

\[ P( \lim_{n \to \infty} \frac{S_n}{n} \neq \mu ) = 0 \]
Central Limit Theorem

Further, assume that $\chi_1, \chi_2, \ldots, \chi_n$ have finite nonzero variance $\sigma^2$:

$$\lim_{n \to \infty} P \left( \frac{S_n - n\mu}{\sigma \sqrt{n}} \leq x \right) = \mathcal{N}(0, 1)$$

$\mathcal{N}(\mu, \sigma^2)$: Normally distributed random variable with mean $\mu$, variance $\sigma^2$.
A More Convenient Form of the CLT

\[ \frac{S_n - n\mu}{\sigma \sqrt{n}} \approx \mathcal{N}(0, 1) \]

\[ \sqrt{n} \left( \frac{\bar{X} - \mu}{\sigma} \right) \approx \mathcal{N}(0, 1) \]

\[ \sqrt{n}(\bar{X} - \mu) \approx \mathcal{N}(0, \sigma^2) \]
Lower Bound on the Optimal Objective Function Value

\[ v^* = \min_{x \in S} \{ f(x) \equiv \mathbb{E}[F(x, \xi)] \} \]

For some sample \( \xi^1, \ldots, \xi^N \), let

\[ \hat{v}_N = \min_{x \in S} \{ \hat{f}_N(x) \equiv N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \}. \]

**Theorem 1.**

\[ \mathbb{E}[\hat{v}_N] \leq v^* \]
Proof

\[ v^* = \min_{x \in S} \mathbb{E}[F(x, \xi)] = \min_{x \in S} \mathbb{E} \left[ N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \right] \]

\[ \min_{x \in S} N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \leq N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \quad \forall x \in S \quad \iff \]

\[ \mathbb{E} \left[ \min_{x \in S} N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \right] \leq \mathbb{E} \left[ N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \right] \quad \forall x \in S \iff \]

\[ \mathbb{E} [\hat{v}_N] \leq \mathbb{E} \left[ N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \right] \quad \forall x \in S \iff \]

\[ \mathbb{E} [\hat{v}_N] \leq \min_{x \in S} \mathbb{E} \left[ N^{-1} \sum_{i=1}^{N} F(x, \xi^i) \right] = v^* \]
Now we need to somehow estimate $\mathbb{E}[\hat{v}_n]$.

The expected value $\mathbb{E}[\hat{v}_N]$ can be estimated as follows.

Generate $M$ independent samples, $\xi^{1,j}, \ldots, \xi^{N,j}, j = 1, \ldots, M$, each of size $N$, and solve the corresponding sample average approximation (SAA) problems

$$
\min_{x \in S} \left\{ \hat{f}_N^j(x) := N^{-1} \sum_{i=1}^{N} F(x, \xi^{i,j}) \right\},
$$

for each $j = 1, \ldots, M$. Let $\hat{v}_N^j$ be the optimal value of problem (2), and compute

$$
L_{N,M} \equiv \frac{1}{M} \sum_{j=1}^{M} \hat{v}_N^j
$$
Lower Bounds

• The estimate $L_{N,M}$ is an unbiased estimate of $\mathbb{E}[\hat{v}_N]$.

• By our last theorem, it provides a statistical lower bound for the true optimal value $v^*$.

• When the $M$ batches $\xi^{1,j}, \xi^{2,j}, \ldots, \xi^{N,j}$, $j = 1, \ldots, M$, are i.i.d., we have by the Central Limit Theorem that

$$\sqrt{M} \left[ L_{N,M} - \mathbb{E}[\hat{v}_N] \right] \to \mathcal{N}(0, \sigma_L^2)$$
Confidence Intervals

- The sample variance estimator of $\sigma_L^2$ is

$$s_L^2(M) \equiv \frac{1}{M-1} \sum_{j=1}^{M} \left( \hat{v}_N^j - L_{N,M} \right)^2.$$ 

Defining $z_\alpha$ to satisfy $P\{N(0,1) \leq z_\alpha\} = 1 - \alpha$, and replacing $\sigma_L$ by $s_L(M)$, we obtain an approximate $(1 - \alpha)$-confidence interval for $\mathbb{E}[\hat{v}_N]$:

$$\left[ L_{N,M} - \frac{z_\alpha s_L(M)}{\sqrt{M}}, L_{N,M} + \frac{z_\alpha s_L(M)}{\sqrt{M}} \right]$$
Example

minimize

\[ Q(x_1, x_2) = x_1 + x_2 + 5 \int_{\omega_1=1}^{4} \int_{\omega_2=1/3}^{2/3} y_1(\omega_1, \omega_2) + y_2(\omega_1, \omega_2) \, d\omega_1 \, d\omega_2 \]

subject to

\[
\begin{align*}
\omega_1 x_1 + x_2 + y_1(\omega_1, \omega_2) & \geq 7 \\
\omega_2 x_1 + x_2 + y_2(\omega_1, \omega_2) & \geq 4 \\
x_1 & \geq 0 \\
x_2 & \geq 0 \\
y_1(\omega_1, \omega_2) & \geq 0 \\
y_2(\omega_1, \omega_2) & \geq 0
\end{align*}
\]
Upper Bounds

\[ v^* = \min_{x \in S} \{ f(x) \equiv \mathbb{E}[F(x, \xi)] \} \]

• From this definition, it is obvious that

\[ f(x) \geq v^* \quad \forall x \in S \]

• How can we estimate \( f(\hat{x}) \) for some \( \hat{x} \in S \)?
Estimating $f(\hat{x})$

- Consider $\hat{x} \in S$.
- We generate $T$ independent batches of samples of size $\bar{N}$, denoted by $\xi^{1,j}, \xi^{2,j}, \ldots, \xi^{\bar{N},j}$, $j = 1, \ldots, T$.
- Each batch has the unbiased property, namely

\[
\mathbb{E} \left[ \hat{f}^{j}_{\bar{N}}(x) \right] = \bar{N}^{-1} \sum_{i=1}^{\bar{N}} F(x, \xi^{i,j}) = f(x), \text{ for all } x \in S.
\]
- We can then use the average value defined by

\[
U_{\bar{N},T}(\hat{x}) \equiv \bar{T}^{-1} \sum_{j=1}^{T} \hat{f}^{j}_{\bar{N}}(\hat{x})
\]

as an estimate of $f(\hat{x})$. 
More Confidence Intervals

By applying the Central Limit Theorem again, we have that

\[
\sqrt{T} \left[ U_{\bar{N},T}(\hat{x}) - f(\hat{x}) \right] \to N(0, \sigma^2_U(\hat{x})), \quad \text{as } T \to \infty,
\]

where \( \sigma^2_U(\hat{x}) \equiv \text{Var} \left[ \hat{f}_{\bar{N}}(\hat{x}) \right] \). We can estimate \( \sigma^2_U(\hat{x}) \) by the sample variance estimator \( s^2_U(\hat{x},T) \) defined by

\[
s^2_U(\hat{x},T) \equiv \frac{1}{T-1} \sum_{j=1}^T \left[ \hat{f}_{\bar{N}}^j(\hat{x}) - U_{\bar{N},T}(\hat{x}) \right]^2.
\]

By replacing \( \sigma^2_U(\hat{x}) \) with \( s^2_U(\hat{x},T) \), we can proceed as above to obtain a \((1 - \alpha)\)-confidence interval for \( f(\hat{x}) \):

\[
\left[ U_{\bar{N},T}(\hat{x}) - \frac{z_\alpha s_U(\hat{x},T)}{\sqrt{T}}, U_{\bar{N},T}(\hat{x}) + \frac{z_\alpha s_U(\hat{x},T)}{\sqrt{T}} \right].
\]
Putting it all together

- $\hat{f}_N(x)$ is the sample average function
  - Draw $\omega^1, \ldots, \omega^N$ from $P$
  - $\hat{f}_N(x) \equiv N^{-1} \sum_{j=1}^{N} F(x, \omega^j)$
  - For stochastic LP w/recourse $\Rightarrow$ solve $N$ LPs.

- $\hat{v}_N \equiv \min_{x \in S} \left\{ \hat{f}_N(x) \equiv N^{-1} \sum_{j=1}^{N} F(x, \omega^j) \right\}$ is the optimal solution value for the sample average function.

- Estimate $\mathbb{E}[\hat{v}_N]$ as $\mathbb{E}[\hat{v}_N] = L_{N,M} = M^{-1} \sum_{j=1}^{M} \hat{v}_N^j$ (solve $M$ stochastic LPs, each of sampled size $N$).
Recapping Theorems

**Theorem 2.** \[ \mathbb{E}[\hat{v}_N] \leq v^* \leq f(x) \quad \forall x \in S \]

**Theorem 3.** \[ U_{\bar{N},\bar{T}}(\hat{x}) - L_{N,M}[\hat{v}_N] \rightarrow f(\hat{x}) - v^*, \text{ as } N, M, \bar{N}, \bar{T} \rightarrow \infty \]

- We are mostly interested in estimating the quality of a given solution \( \hat{x} \). This is \( f(\hat{x}) - v^* \).
- \( \hat{f}_{N'}(\hat{x}) \) computed by solving \( N' = \bar{N}T \) independent LPs.
- \( \mathbb{E}[\hat{v}_N] \) computed by solving \( M \) independent stochastic LPs.
An Experiment

- Solve a stochastic sampled approximation of size $NM$ times (thus obtaining an estimate of $E[\hat{v}_N]$).

- For each of the $M$ solutions $x^i, i = 1, \ldots, M$, estimate $f(x^i)$ by solving $N' = \tilde{N}T$ LPs.

- Test Instances

| Name | Application                  | $|\Omega|$  | $(m_1, n_1)$ | $(m_2, n_2)$ |
|------|------------------------------|------------|--------------|--------------|
| LandS| HydroPower Planning          | $10^6$     | (2,4)        | (7,12)       |
| gbd  | ?                            | $6.46 \times 10^5$ | (?,?)        | (?,?)        |
| storm| Cargo Flight Scheduling      | $6 \times 10^{81}$ | (185, 121)   | (?,1291)     |
| 20term| Vehicle Assignment          | $1.1 \times 10^{12}$ | (1,5)        | (71,102)     |
| ssn  | Telecom. Network Design      | $10^{70}$  | (1,89)       | (175,706)    |
Convergence of Optimal Solution Value

- $9 \leq M \leq 12$, $N' = 10^6$
- Monte Carlo Sampling

<table>
<thead>
<tr>
<th>Instance</th>
<th>$N = 50$</th>
<th>$N = 100$</th>
<th>$N = 500$</th>
<th>$N = 1000$</th>
<th>$N = 5000$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20term</td>
<td>253361 254442</td>
<td>254025 254399</td>
<td>254324 254394</td>
<td>254307 254475</td>
<td>254341 254376</td>
</tr>
<tr>
<td>gbd</td>
<td>1678.6 1660.0</td>
<td>1595.2 1659.1</td>
<td>1649.7 1655.7</td>
<td>1653.5 1655.5</td>
<td>1653.1 1655.4</td>
</tr>
<tr>
<td>LandS</td>
<td>227.19 226.18</td>
<td>226.39 226.13</td>
<td>226.02 226.08</td>
<td>225.96 226.04</td>
<td>225.72 226.11</td>
</tr>
<tr>
<td>storm</td>
<td>1550627 1550321</td>
<td>1548255 1550255</td>
<td>1549814 1550228</td>
<td>1550087 1550236</td>
<td>1549812 1550239</td>
</tr>
</tbody>
</table>