Stochastic Decomposition

For Problems with Continuous Random Outcomes

References

Higle, J. L. and S. Sen (1991). "Stochastic decomposition: An algorithm for two-stage linear programs with recourse." *Mathematics of Operations Research* **16**(3): 650-669.

Higle, J. L. and S. Sen (1996). *Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming*. Dordrecht, Kluwer Academic Publishers.

Consider the **2-stage stochastic LP**:

$$\text{Minimize } z = cx + E\left[\min q(\omega)\, y(\omega)\right]$$

subject to

$$Ax = b$$
$$T(\omega)x + Wy(\omega) = h(\omega),$$
$$x \geq 0,\ y(\omega) \geq 0$$

where

x = first-stage decision

and

$y(\omega)$ = second-stage decision *after* random event ω is observed

where $y(\omega)$ must satisfy the ***second-stage constraints***

$$T(\omega)x + Wy(\omega) = h(\omega),$$

$q(\omega)$, $T(\omega)$ &/or $h(\omega)$ being continuous random variables.

Consider, for example, the case in which only h is random.

A possible computational approach:

- *discretize* the range of each right-hand-side $h_i(\omega)$
- use Benders' decomposition (i.e., the "L-shaped Method") to solve the approximate problem

If the number of right-hand-sides (m_2) and/or the number of discrete values approximating each right-hand-side are large, the number of scenarios is so large as to make this computationally infeasible.

For example, if there are $m_2=10$ constraints, and only 10 discrete values are used for each right-hand-side, the number of scenarios is 10^{10} !

The **Stochastic Decomposition** (SD) method of Higle & Sen is based upon (the *uni-cut* version of) Benders' decomposition, but

- uses only a *finite sample* of the random outcomes
- solves most of the second-stage problems only *approximately*

For both these reasons, therefore, it is an *approximation* scheme.

Stochastic Decomposition Algorithm of Higle & Sen

Step 0. *a.* Determine a *lower* bound L on the optimal value.

b. Set iteration counter t=0.

c. Initialize $\Lambda = \emptyset$ which will store the dual extreme points that are generated during the computations.

Step 1. Increment the iteration counter $t \leftarrow t+1$.

Solve the current Benders' *Master Problem*:

$$\begin{aligned} &\textit{Maximize} \quad cx + \theta \\ &\text{subject to } \; Ax = b, \\ &\qquad \theta \geq \alpha^s x + \beta^s, \quad s = 1,2,\ldots t \\ &\qquad x \geq 0 \end{aligned}$$

to obtain x^t

Step 2. Generate a sample ω^t (of size 1).

Step 3. Solve (optimally) the second-stage **subproblem** problem for the current x^t and ω^t:

$$\begin{aligned} &Min\ \ q(\omega)y(\omega) \\ &s.t.\ \ \ Wy(\omega) = h(\omega) - T(\omega)x^t \\ &y(\omega) \geq 0 \end{aligned}$$

or its dual LP,

$$\begin{aligned} &Max\ \ \lambda[h(\omega)-T(\omega)x^t] \\ &s.t.\ \ \lambda W \leq q(\omega) \end{aligned}$$

to obtain the dual solution λ_t^t, which, if not found previously, is added to the set Λ.

Step 4. Using the current x^t,

for all *previously-generated* scenarios ω^s, $s = 1, \ldots t-1$,

approximately solve the second stage *dual* subproblem,

restricting the search to dual extreme points Λ previously computed:

$$\underset{\lambda \in \Lambda}{Max} \left[\mathrm{h}\left(\omega^s\right) - T\left(\omega^s\right) x^t \right] \lambda$$

to obtain λ_s^t.

Note that this gives an ***under****-estimate of the optimal cost for this scenario, since the maximization is over a* ***subset*** *of all dual extreme points!*

Step 5. Generate the *new* optimality cut, to be added to the Master Problem:

$$\theta \geq \frac{1}{t} \sum_{s=1}^{t} \lambda_s^t \left(h\left(\omega^s\right) - T\left(\omega^s\right) x \right) \equiv \alpha_t^t + \beta_t^t x$$

Step 6. Update each of the *old* optimality cuts, $(s = 1, 2, \ldots t-1)$ by replacing

$$\theta \geq a_s^{t-1} + \beta_s^{t-1} x$$

with

$$\theta \geq \frac{t-1}{t}\left(\alpha_s^{t-1} + \beta_s^{t-1} x\right) + \frac{1}{t} L$$

and return to **Step 1**.

Updating the Optimality Cuts

- The effect of updating the old optimality cuts in step 6 is to "fade out" the cuts as more information becomes available.

- The lower bound ***L*** is often zero, or it may be an estimate of the expected value with perfect information, computed using a sample of random outcomes.

Convergence Properties:

Let $\left\{x^t\right\}_{t=1}^{\infty}$ be the sequence of solutions of the Master Problems.

Then there exists a ***subsequence***, $\left\{x^{t_n}\right\} \subseteq \left\{x^t\right\}$ such that

every limit point of $\left\{x^{t_n}\right\}$ solves the stochastic

programming problem with probability 1.

Incumbent Solution

One difficulty in the basic method is that convergence to an optimum may occur only on a *subsequence*. For this reason, Higle & Sen suggest retaining an ***incumbent*** solution.

This incumbent solution is updated whenever there is a "sufficient" decrease in cost compared to the current incumbent.

Furthermore, in **step 6**, no update is performed for the cut generated in the iteration at which the current incumbent was found.

Termination

In practice, the SD algorithm is terminated if

- the improvement in the objective is small,
- no new dual extreme points are found, and
- the incumbent has not changed

for a specified number of iterations,

EXAMPLE: Randomly-generated problem

Dimensions:

- n_1 = # first-stage variables = 4
- m_1 = # first-stage constraints = 3
- n_2 = # second-stage variables = 12 (including 2 "simple recourse" variables per constraint)
- m_2 = # second-stage constraints = 4

First-stage data:

A,B=

¯2	1	8	0	>	14
3	¯3	9	7	>	32
1	1	1	1	<	16

i	variable	cost
1	X[1]	5
2	X[2]	1
3	X[3]	7
4	X[4]	2

Objective: Minimize

Second-stage data

(Only the right-hand-side vector is random!)

Right-hand-sides in second stage =

i	mean	std dev
1	¯13	1.4
2	¯7	0.6
3	11	0.5
4	24	1.9

Second-stage Costs:

i	variable	q
1	Y[1]	10
2	Y[2]	10
3	Y[3]	10
4	Y[4]	7
5	Surplus1	99
6	Surplus2	99
7	Surplus3	99
8	Surplus4	99
9	Short1	99
10	Short2	99
11	Short3	99
12	Short4	99

Technology matrix T

(coefficients of X in 2nd stage) =

¯4	0	¯3	¯1
¯1	5	¯4	¯4
2	¯2	4	0
4	¯1	5	1

Technology matrix W

(coefficients of Y in 2nd stage) =

1	¯1	¯2	5	1	0	0	0	¯1	0	0	0
0	¯3	5	¯1	0	1	0	0	0	¯1	0	0
¯1	0	2	2	0	0	1	0	0	0	¯1	0
1	2	1	2	0	0	0	1	0	0	0	¯1

Solving the Certainty-Equivalent Problem

Found by solving certainty equivalent problem,
i.e., replacing all random parameters by their expected values.

Total objective function: 46.1403
Stage One: nonzero variables:

i	variable	value
1	X[1]	2.85221
2	X[2]	2.93628
3	X[3]	2.09602
4	X[4]	2.26327
6	surplus_2	2.45487
7	slack_3	5.85221

Second Stage: nonzero variables

i	variable	value
4	Y[4]	1.39204

Stochastic Decomposition Algorithm

Iteration #1

Trial X for primal subproblems (#1) is

i	Variable	Value
1	X[1]	2.85221
2	X[2]	2.93628
3	X[3]	2.09602
4	X[4]	2.26327

(found by solving problem with expected values of right-hand-sides)

Solve subproblem with new trial x (#1) :
Primal Subproblem Result: nonzero elements of X (#1):

i	X[i]
1	2.85221
2	2.93628
3	2.09602
4	2.26327

RHS = ¯12.4758 ¯8.23344 10.544 24.9054 *(first scenario)*

Second-stage cost: 78.4487
Optimal dual vector: 48.2273 ¯85.4091 ¯60.7727 ¯99

Newly-generated optimality cut at iteration 1

s	i	beta	x[1]	x[2]	x[3]	x[4]
1	1	¯3004.89	625.045	206.5	541.136	¯194.409

s is scenario #, i is dual solution #, beta is constant

Aggregate cut:

beta	X[1]	X[2]	X[3]	X[4]
¯3004.89	625.045	206.5	541.136	¯194.409

Primal subproblems summary

First stage cost: 36.396

Second stage costs:

s	Lambda#	cost
1	1	78.4487

Average second stage cost: 78.4487

Total: 114.845

Solution of Master Problem

```
X= 2.85221 2.93628 2.09602 2.26327

       First-stage cost= 40.75
Estimated second-stage cost Q(X) = ¯4828.23
Total (estimated) expected value: ¯4787.48
```

Iteration #2

Trial X for primal subproblems (#2) is

i	Variable	Value
1	X[1]	0.00
2	X[2]	0.00
3	X[3]	1.75
4	X[4]	14.25

(found by previous master problem)

Solve subproblem with new trial x (#2) :

Primal Subproblem Result:

RHS = ¯15.0969 ¯6.55505 11.2261 21.3609 *(second scenario)*

Second-stage cost: 4060.6

Optimal dual vector: 69.7714 65.4 ¯39.2286 ¯99

Solve subproblem with incumbent solution (#1) :

Primal Subproblem Result:

i	X[i]
1	2.85221
2	2.93628
3	2.09602
4	2.26327

RHS = ¯15.0969 ¯6.55505 11.2261 21.3609

Second-stage cost: 289.983
Optimal dual vector: ¯2.34783 ¯18.7391 99 ¯99

Newly-generated optimality cut at iteration 2

s	i	beta	x[1]	x[2]	3]	x[4]
1	2	¯1238.2	169.87	192.696	17	21.6957
2	2	¯845.065	169.87	192.696	17	21.6957

s is scenario #, i is dual solution #, beta is constant

Aggregate cut:

beta	X[1]	X[2]	3]	X[4]
¯1041.63	169.87	192.696	17	21.6957

Primal subproblems summary

First stage cost: 40.75
Second stage costs:

s	Lambda#	cost
1	2	¯899.283
2	2	289.983

Average second stage cost: ¯304.65
Total: ¯263.9

Solution of Master Problem

X= 0 0 1.75 14.25
First-stage cost: 24.8889
Estimated second-stage cost Q(X) = ¯981.186
Total (estimated) expected value: ¯956.298

Iteration #3

Trial X for primal subproblems (#3) is

i	Variable	Value	
3	X[3]	3.55556	(found by Master Problem)

Solve subproblem with new trial x (#3) :
Primal Subproblem Result:

RHS = ¯11.7763 ¯6.8984 11.2903 25.526 *(third scenario)*

Second-stage cost: 376.236
Optimal dual vector: ¯76.2917 13.625 ¯99 ¯12.7083

Solve subproblem with incumbent solution (#2) :
Primal Subproblem Result:

nonzero elements of X (#2):

i	X[i]
3	1.75
4	14.25

RHS = ¯11.7763 ¯6.8984 11.2903 25.526
Second-stage cost: 3854.96
Optimal dual vector: 69.7714 65.4 ¯39.2286 ¯99

Newly-generated optimality cut at iteration 3

s	i	beta	x[1]	x[2]	x[3]	x[4]
1	3	¯4288.18	818.943	¯504.457	1122.83	430.371
2	3	¯4037.14	818.943	¯504.457	1122.83	430.371
3	3	¯4242.78	818.943	¯504.457	1122.83	430.371

s is scenario #, i is dual solution #, beta is constant

Aggregate cut:

beta	X[1]	X[2]	X[3]	X[4]
¯4189.37	818.943	¯504.457	1122.83	430.371

Primal subproblems summary

First stage cost: 24.8889

Second stage costs:

s	Lambda#	cost
1	3	¯44.8642
2	3	¯295.9024
3	3	3854.9594

Average second stage cost: 1171.4

Total: 1196.29

That is, the 3^{rd} dual solution in the list was optimal for all three scenarios.

Solution of Master Problem

```
X= 0 0 3.55556 0
First-stage cost: 18.906
Estimated second-stage cost Q(X) = ¯966.468
Total (estimated) expected value: ¯947.562
```

Iteration #4

Trial X for primal subproblems (#4) is

i	Variable	Value
3	X[3]	2.20457
4	X[4]	1.73698

(found by Master Problem)

Solve subproblem with new trial x (#4) :
Primal Subproblem Result:

RHS = ¯14.1861 ¯7.00585 10.8897 24.0418 *(fourth scenario)*

Second-stage cost: 216.109
Optimal dual vector: ¯76.2917 13.625 ¯99 ¯12.7083

Solve subproblem with incumbent solution (#2) :
Primal Subproblem Result:

i	X[i]
3	1.75
4	14.25

RHS = ¯14.1861 ¯7.00585 10.8897 24.0418
Second-stage cost: 3842.45
Optimal dual vector: 69.7714 65.4 ¯39.2286 ¯99

Newly-generated optimality cut at iteration 4

s	i	beta	x[1]	x[2]	x[3]	x[4]
1	3	¯4288.18	818.943	¯504.457	1122.83	430.371
2	2	¯845.065	169.87	192.696	17	21.6957
3	3	¯4242.78	818.943	¯504.457	1122.83	430.371
4	3	¯4255.29	818.943	¯504.457	1122.83	430.371

s is scenario #, i is dual solution #, beta is constant

Aggregate cut:

beta	X[1]	X[2]	X[3]	X[4]
¯3407.83	656.675	¯330.169	846.371	328.202

Primal subproblems summary

First stage cost: 18.906

Second stage costs:

s	Lambda#	cost
1	3	¯1019.882
2	2	¯769.903
3	3	¯1065.280
4	3	3842.451

Average second stage cost: 246.846

Total: 265.752

Solution of Master Problem

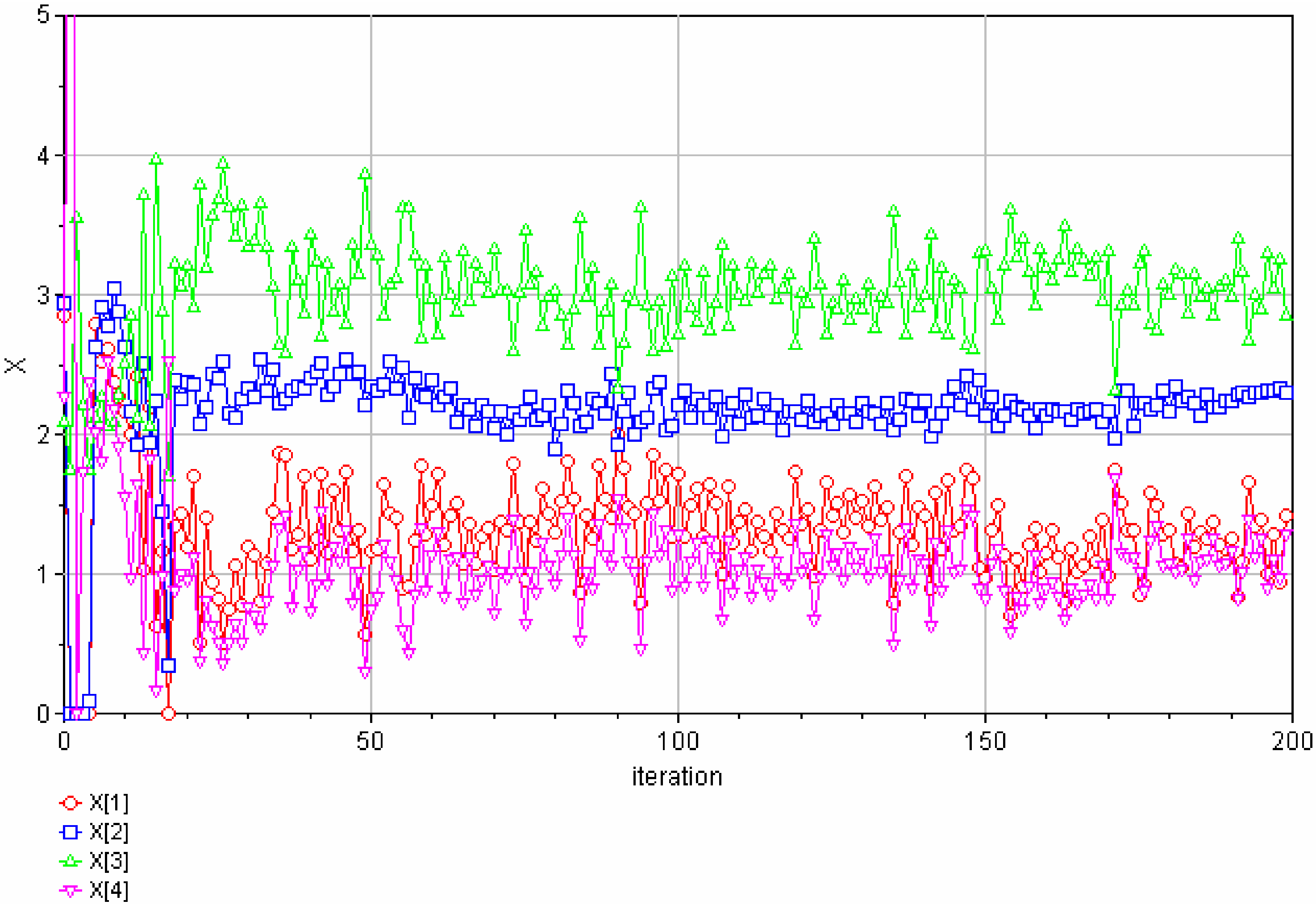
X= 0 0 2.20457 1.73698
First-stage cost: 17.0044
Estimated second-stage cost Q(X) = _¯944.114
Total (estimated) expected value: ¯927.11

Output for 200 iterations

Subproblems were solved approximately, except for most recent x and the incumbent!

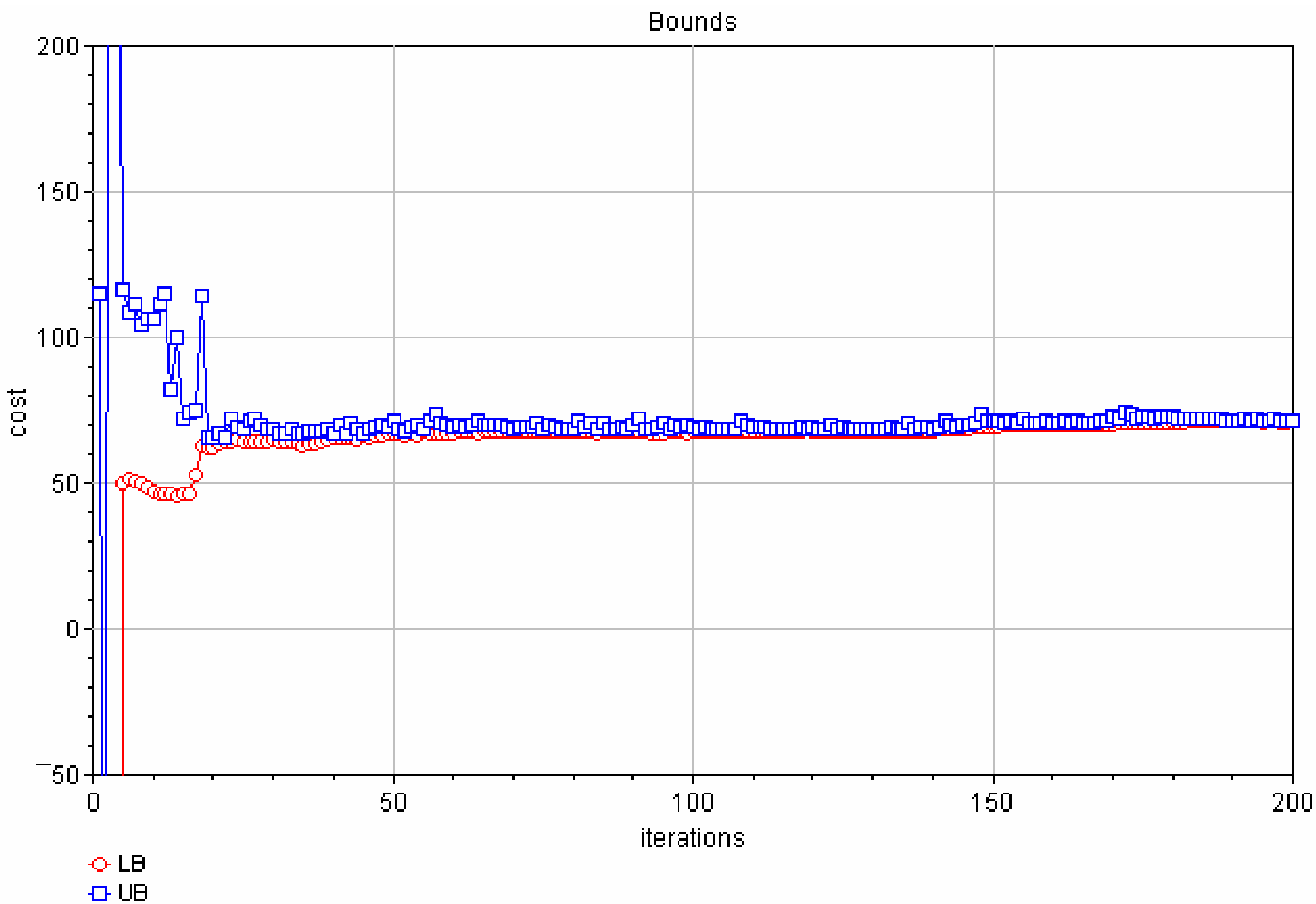
```
Stochastic Decomposition

Randomly-generated SLPwR problem (seed= 17853)
Random number seed used in computation: 17977

Method: Subproblems solved  approximately
Tolerance for distinguishing first-stage solutions X:
   1.0E¯3

# iterations (= # right-hand-sides sampled): 200
# second-stage problems solved: 399

# first-stage solutions generated: 200
Best solution found is #189 with estimated cost 71.3121
12 second-stage problems were solved using this X

# second-stage dual solutions generated: 6
```

Values of first-stage variables (solutions of Master Problem):



"Lower" and "Upper" Bounds

(found by Master & approximate Subproblems):



The Incumbent Solution

Evaluation of trial solution # 189

i	variable	X[i]
1	X[1]	1.21096
2	X[2]	2.18995
3	X[3]	3.05608
4	X[4]	1.06174

Three different methods are used to estimate the expected cost of this solution:

Evaluation by:

- Use cuts
- Use recorded dual solutions *(i.e., solve subproblems with dual variables restricted to the identified dual extreme points)*
- Use recorded Q values *(i.e., use actual optimal subproblem solutions computed with this first-stage solution)*

----

1. Using optimality cuts as approximation of expected second-stage cost.

First stage objective:	31.76
Expected second stage objective:	**39.84**
Total:	71.60

----

2. Using expected second-stage costs approximated by restriction to 6 recorded dual solutions.

First stage objective:	31.76
Expected second stage objective:	**39.65**
Total:	71.41

----

3. Using 12 evaluations of second-stage costs.

First stage objective:	31.76
Expected second stage objective:	**33.47**
Total:	65.23

Suppose that we had expended the extra effort to solve the subproblems optimally for every scenario (rather than only the most recently-generated scenario):

```
Random number seed used in computation: 19138
Method: Subproblems solved  exactly

Tolerance for distinguishing first-stage solutions X:   1.0E¯3

# iterations (= # right-hand-sides sampled): 200
# second-stage problems solved: 20299

# first-stage solutions generated: 200

Best solution found is #111 with estimated cost 66.6435
200 second-stage problems were solved using this X

# second-stage dual solutions generated: 10
```

Compared to 6 dual solutions found previously! But over fifty times the number of subproblems were solved, a substantial increase in effort!