Submitted to *INFORMS Journal on Computing*
manuscript (Please, provide the manuscript number!)

SUPPLEMENTAL MATERIAL for Software for data-based stochastic programming using bootstrap estimation

Xiaotie Chen
Department of Mathematics, UC Davis, xtjchen@ucdavis.edu,

David L Woodruff
Graduate School of Management, UC Davis,dlwoodruff@ucdavis.edu,

---

Appendix A: Experiments

We demonstrate experiments on the algorithms over a few examples as detailed in the following subsections. For simulation purpose, we grant ourselves the access to the distribution of the entire population F_Ω, and approximate the theoretical optimal function value z^* by drawing an extremely large number of samples from the distribution F_Ω, then compute the corresponding optimal function value, which we use as z^* for the simulations.

We replicate the execution of each algorithm a large number (e.g., 500) times with a fixed candidate solution $\hat{x}$. At each replication, a new set D is drawn independently from F_Ω, and the algorithm is executed to return a $(1-\alpha)$ confidence interval for the optimal function value z^*. The coverage rate is then reported as the percentages of the confidence intervals that contains z^* over the replications. We also report the average length of the confidence intervals, denoted as "avg len", over the repeated experiments for each algorithm. The average length of the CIs to some extent represents the sharpness of the intervals computed by the algorithms.

A.1. Small Schultz Examples

A.1.1. Unique Solution This example is from Schultz et al. (1998) used in (Eichhorn and Römisch 2007, p. 129). Results are shown in Table 1.

A.1.2. Nonunique Solution This example is a modified version of the previous problem that used in (Eichhorn and Römisch 2007, p. 131). Results are shown in Table 2. This problem has multiple optima.

method	N	B	k	avg len	coverage
Classical_gaussian	40	100	-	8.04	0.92
Classical_gaussian	80	100	-	5.58	0.90
Classical_gaussian	40	500	-	8.01	0.92
Classical_gaussian	80	500	-	5.73	0.93
Classical_quantile	40	100	-	7.82	0.88
Classical_quantile	80	100	-	5.41	0.87
Classical_quantile	40	500	-	7.97	0.90
Classical_quantile	80	500	-	5.69	0.89
Bagging_with_replacement	40	100	16	9.36	0.93
Bagging_with_replacement	40	100	24	9.38	0.95
Bagging_with_replacement	40	100	32	9.42	0.96
Bagging_with_replacement	80	100	32	7.62	0.97
Bagging_with_replacement	80	100	48	7.39	0.97
Bagging_with_replacement	80	100	64	7.49	0.98
Bagging_with_replacement	40	500	16	8.21	0.90
Bagging_with_replacement	40	500	24	8.15	0.89
Bagging_with_replacement	40	500	32	8.26	0.91
Bagging_with_replacement	80	500	32	6.08	0.93
Bagging_with_replacement	80	500	48	6.05	0.95
Bagging_with_replacement	80	500	64	6.07	0.95
Bagging_without_replacement	40	100	16	9.32	0.95
Bagging_without_replacement	40	100	24	9.59	0.93
Bagging_without_replacement	40	100	32	9.43	0.94
Bagging_without_replacement	80	100	32	7.54	0.98
Bagging_without_replacement	80	100	48	7.83	0.98
Bagging_without_replacement	80	100	64	7.68	0.95
Bagging_without_replacement	40	500	16	8.42	0.91
Bagging_without_replacement	40	500	24	8.58	0.94
Bagging_without_replacement	40	500	32	8.60	0.93
Bagging_without_replacement	80	500	32	6.16	0.93
Bagging_without_replacement	80	500	48	6.29	0.95
Bagging_without_replacement	80	500	64	6.26	0.95
Subsampling	40	100	16	6.15	0.79
Subsampling	40	100	24	5.08	0.73
Subsampling	40	100	32	3.52	0.58
Subsampling	80	100	32	4.36	0.80
Subsampling	80	100	48	3.62	0.72
Subsampling	80	100	64	2.54	0.55
Subsampling	40	500	16	6.30	0.81
Subsampling	40	500	24	5.18	0.74
Subsampling	40	500	32	3.65	0.56
Subsampling	80	500	32	4.45	0.82
Subsampling	80	500	48	3.68	0.74
Subsampling	80	500	64	2.58	0.56
Extended	40	100	-	8.29	0.88
Extended	80	100	-	5.37	0.81
Extended	40	500	-	8.22	0.90
Extended	80	500	-	5.60	0.86

Table 1 Results for unique_schultz ($z^* \approx$ -62.29) with α=0.1 based on 100 replications.

method	N	B	k	avg len	coverage
Classical_gaussian	40	100	-	7.92	0.90
Classical_gaussian	80	100	-	5.46	0.91
Classical_gaussian	40	500	-	7.89	0.89
Classical_gaussian	80	500	-	5.60	0.93
Classical_quantile	40	100	-	7.72	0.89
Classical_quantile	80	100	-	5.30	0.86
Classical_quantile	40	500	-	7.84	0.90
Classical_quantile	80	500	-	5.57	0.92
Bagging_with_replacement	40	100	16	9.23	0.93
Bagging_with_replacement	40	100	24	9.24	0.94
Bagging_with_replacement	40	100	32	9.28	0.96
Bagging_with_replacement	80	100	32	7.45	0.96
Bagging_with_replacement	80	100	48	7.23	0.97
Bagging_with_replacement	80	100	64	7.32	0.97
Bagging_with_replacement	40	500	16	8.10	0.89
Bagging_with_replacement	40	500	24	8.02	0.89
Bagging_with_replacement	40	500	32	8.13	0.91
Bagging_with_replacement	80	500	32	5.95	0.92
Bagging_with_replacement	80	500	48	5.92	0.94
Bagging_with_replacement	80	500	64	5.94	0.94
Bagging_without_replacement	40	100	16	9.19	0.95
Bagging_without_replacement	40	100	24	9.43	0.93
Bagging_without_replacement	40	100	32	9.27	0.92
Bagging_without_replacement	80	100	32	7.38	0.98
Bagging_without_replacement	80	100	48	7.63	0.98
Bagging_without_replacement	80	100	64	7.50	0.96
Bagging_without_replacement	40	500	16	8.29	0.90
Bagging_without_replacement	40	500	24	8.44	0.91
Bagging_without_replacement	40	500	32	8.46	0.93
Bagging_without_replacement	80	500	32	6.03	0.93
Bagging_without_replacement	80	500	48	6.14	0.96
Bagging_without_replacement	80	500	64	6.11	0.94
Subsampling	40	100	16	6.04	0.82
Subsampling	40	100	24	4.99	0.72
Subsampling	40	100	32	3.45	0.61
Subsampling	80	100	32	4.28	0.81
Subsampling	80	100	48	3.52	0.76
Subsampling	80	100	64	2.48	0.57
Subsampling	40	500	16	6.21	0.83
Subsampling	40	500	24	5.10	0.78
Subsampling	40	500	32	3.59	0.60
Subsampling	80	500	32	4.36	0.83
Subsampling	80	500	48	3.60	0.78
Subsampling	80	500	64	2.52	0.56
Extended	40	100	-	8.16	0.86
Extended	80	100	-	5.16	0.81
Extended	40	500	-	8.08	0.88
Extended	80	500	-	5.45	0.89

Table 2 **Results for nonunique_schultz ($z^* \approx$ -61.36) with α=0.1 based on 100 replications.**

A.2. CVaR

A one-stage CVaR problem is used by Lam and Qian (2018):

$$\min_x \left\{ x + \frac{1}{a} E\left[(\xi - x)_+\right] \right\}$$

where $(\cdot)_+$ is defined as $\max \cdot, 0$, $a = 0.1$ and ξ is a drawn from a standard normal distribution. Results are shown in Table 3.

A.3. Scalable Farmer

This example is based on the well-known farmer example from Birge and Louveaux (1997) as modifed for stress-testing various pieces of software such as Knueven et al. (2020, updated 2022). To make it scalable, two instance creation parameters *cropsmult* and *numscens* are added. The original problem has three crops and three scenarios. The scalable instances have *cropsmult* sets of the original three crops with the characteristics as in the original problem and yields that depend on the scenario. Scenarios are in groups of three with a uniformly distributed psuedo-random number added to the yield values of the original three scenarios.

For the results shown in Tables 4 and 5, we used the original three crops and 1000 scenarios to get an assumed value for z^*.

A.4. Discussion of Results

These experiments are intended mainly to illustrate that the software can be used for such experiments. They do illustrate the unsurprising result that if the samples are too small, the confidence intervals will not be very good. They also suggest that the method we call Extended, which is sort of an afterthought in Eichhorn and Römisch (2007), does not seem to work all that well.

The results are mostly reasonable, but mixed and depend on the availability of enough data as well as the choice of method and parameters. Detailed conclusions are beyond the scope of this small study. A preliminary indication is that bagging might be the best thing to try first.

One other thing to note, though, concerns CVaR. Since CVaR considers the tail, getting good confidence intervals requires a larger value of N. Perhaps for similar reasons, quantile-based intervals do not seem to be as good as Gaussian.

References

Birge JR, Louveaux F (1997) *Introduction to Stochastic Programming* (Springer Verlag).

Eichhorn A, Römisch W (2007) Stochastic integer programming: Limit theorems and confidence intervals. *Mathematics of Operations Research* 32(1):118–135.

Knueven B, Mildebrath D, Muir C, Siirola J, Watson JP, Woodruff D (2020, updated 2022) A parallel hub-and-spoke system for large-scale scenario-based optimization under uncertainty. Preprint 2020-11-8088, Optimization Online.

Lam H, Qian H (2018) Assessing solution quality in stochastic optimization via bootstrap aggregating. *2018 Winter Simulation Conference (WSC)*, 2061–2071 (IEEE).

Schultz R, Stougie L, Van Der Vlerk MH (1998) Solving stochastic programs with integer recourse by enumeration: A framework using gröbner basis. *Mathematical Programming* 83(1):229–252.

method	N	B	k	avg len	coverage
Classical_gaussian	300	100	-	0.36	0.90
Classical_gaussian	600	100	-	0.26	0.82
Classical_gaussian	300	1000	-	0.36	0.94
Classical_gaussian	600	1000	-	0.26	0.80
Classical_quantile	300	100	-	0.35	0.50
Classical_quantile	600	100	-	0.25	0.57
Classical_quantile	300	1000	-	0.36	0.69
Classical_quantile	600	1000	-	0.26	0.58
Bagging_with_replacement	300	100	120	0.63	1.00
Bagging_with_replacement	300	100	180	0.62	1.00
Bagging_with_replacement	300	100	240	0.62	1.00
Bagging_with_replacement	600	100	240	0.62	1.00
Bagging_with_replacement	600	100	360	0.63	1.00
Bagging_with_replacement	600	100	480	0.62	1.00
Bagging_with_replacement	300	1000	120	0.20	1.00
Bagging_with_replacement	300	1000	180	0.20	1.00
Bagging_with_replacement	300	1000	240	0.20	1.00
Bagging_with_replacement	600	1000	240	0.20	1.00
Bagging_with_replacement	600	1000	360	0.20	1.00
Bagging_with_replacement	600	1000	480	0.20	1.00
Bagging_without_replacement	300	100	120	0.81	1.00
Bagging_without_replacement	300	100	180	0.99	1.00
Bagging_without_replacement	300	100	240	1.39	1.00
Bagging_without_replacement	600	100	240	0.80	1.00
Bagging_without_replacement	600	100	360	0.99	1.00
Bagging_without_replacement	600	100	480	1.39	1.00
Bagging_without_replacement	300	1000	120	0.26	1.00
Bagging_without_replacement	300	1000	180	0.31	1.00
Bagging_without_replacement	300	1000	240	0.45	1.00
Bagging_without_replacement	600	1000	240	0.26	1.00
Bagging_without_replacement	600	1000	360	0.32	1.00
Bagging_without_replacement	600	1000	480	0.45	1.00
Subsampling	300	100	120	0.36	0.65
Subsampling	300	100	180	0.35	0.61
Subsampling	300	100	240	0.35	0.60
Subsampling	600	100	240	0.26	0.64
Subsampling	600	100	360	0.26	0.60
Subsampling	600	100	480	0.25	0.50
Subsampling	300	1000	120	0.36	0.67
Subsampling	300	1000	180	0.36	0.65
Subsampling	300	1000	240	0.36	0.62
Subsampling	600	1000	240	0.26	0.63
Subsampling	600	1000	360	0.26	0.62
Subsampling	600	1000	480	0.26	0.67
Extended	300	100	-	0.50	0.79
Extended	600	100	-	0.36	0.71
Extended	300	1000	-	0.51	0.85
Extended	600	1000	-	0.36	0.81

Table 3 **Results for cvar ($z^* \approx 1.79$) with α=0.1 based on 100 replications.**

method	N	B	k	avg len	coverage
Classical_gaussian	30	100	-	28259.46	0.887
Classical_gaussian	60	100	-	20181.52	0.892
Classical_gaussian	30	1000	-	28294.08	0.885
Classical_gaussian	60	1000	-	20272.41	0.907
Classical_quantile	30	100	-	27343.29	0.870
Classical_quantile	60	100	-	19645.95	0.877
Classical_quantile	30	1000	-	28236.30	0.880
Classical_quantile	60	1000	-	20235.36	0.905
Bagging_with_replacement	30	100	12	31422.27	0.932
Bagging_with_replacement	30	100	18	31257.34	0.938
Bagging_with_replacement	30	100	24	32061.31	0.932
Bagging_with_replacement	60	100	24	25164.00	0.955
Bagging_with_replacement	60	100	36	25298.68	0.953
Bagging_with_replacement	60	100	48	25285.28	0.950
Bagging_with_replacement	30	1000	12	28323.03	0.895
Bagging_with_replacement	30	1000	18	28489.61	0.902
Bagging_with_replacement	30	1000	24	28554.10	0.900
Bagging_with_replacement	60	1000	24	20614.13	0.907
Bagging_with_replacement	60	1000	36	20688.71	0.907
Bagging_with_replacement	60	1000	48	20751.51	0.905
Bagging_without_replacement	30	100	12	32638.99	0.935
Bagging_without_replacement	30	100	18	32640.54	0.920
Bagging_without_replacement	30	100	24	32558.19	0.930
Bagging_without_replacement	60	100	24	25546.72	0.965
Bagging_without_replacement	60	100	36	25369.97	0.963
Bagging_without_replacement	60	100	48	25711.17	0.970
Bagging_without_replacement	30	1000	12	29438.81	0.902
Bagging_without_replacement	30	1000	18	29629.62	0.905
Bagging_without_replacement	30	1000	24	29664.24	0.895
Bagging_without_replacement	60	1000	24	21069.04	0.915
Bagging_without_replacement	60	1000	36	21175.03	0.917
Bagging_without_replacement	60	1000	48	21287.06	0.922
Subsampling	30	100	12	21560.28	0.777
Subsampling	30	100	18	17703.34	0.680
Subsampling	30	100	24	12593.48	0.502
Subsampling	60	100	24	15389.09	0.785
Subsampling	60	100	36	12457.75	0.680
Subsampling	60	100	48	8943.31	0.500
Subsampling	30	1000	12	22124.33	0.797
Subsampling	30	1000	18	18209.22	0.690
Subsampling	30	1000	24	12877.64	0.510
Subsampling	60	1000	24	15739.58	0.790
Subsampling	60	1000	36	12903.63	0.700
Subsampling	60	1000	48	9166.89	0.525
Extended	30	100	-	26890.50	0.848
Extended	60	100	-	19620.82	0.863
Extended	30	1000	-	28080.65	0.870
Extended	60	1000	-	20435.31	0.875

Table 4 Results for farmer ($z^* \approx$ -132750.32) with α=0.1 based on 400 replications.

method	N	B	k	avg len	coverage
Classical_gaussian	30	100	-	21809.65	0.750
Classical_gaussian	60	100	-	15622.85	0.740
Classical_gaussian	30	1000	-	21955.60	0.770
Classical_gaussian	60	1000	-	15648.47	0.730
Classical_quantile	30	100	-	21348.92	0.740
Classical_quantile	60	100	-	15337.17	0.740
Classical_quantile	30	1000	-	22030.79	0.740
Classical_quantile	60	1000	-	15641.96	0.740
Bagging_with_replacement	30	100	12	24013.20	0.800
Bagging_with_replacement	30	100	18	24468.50	0.790
Bagging_with_replacement	30	100	24	25117.26	0.830
Bagging_with_replacement	60	100	24	19617.59	0.830
Bagging_with_replacement	60	100	36	19544.47	0.820
Bagging_with_replacement	60	100	48	19788.37	0.860
Bagging_with_replacement	30	1000	12	21923.88	0.730
Bagging_with_replacement	30	1000	18	22105.87	0.760
Bagging_with_replacement	30	1000	24	22159.09	0.770
Bagging_with_replacement	60	1000	24	16040.28	0.750
Bagging_with_replacement	60	1000	36	16049.88	0.760
Bagging_with_replacement	60	1000	48	16071.97	0.760
Bagging_without_replacement	30	100	12	25178.82	0.830
Bagging_without_replacement	30	100	18	25316.79	0.780
Bagging_without_replacement	30	100	24	25361.81	0.790
Bagging_without_replacement	60	100	24	19903.04	0.890
Bagging_without_replacement	60	100	36	19887.07	0.860
Bagging_without_replacement	60	100	48	19734.46	0.880
Bagging_without_replacement	30	1000	12	22641.34	0.760
Bagging_without_replacement	30	1000	18	23047.99	0.780
Bagging_without_replacement	30	1000	24	23065.88	0.770
Bagging_without_replacement	60	1000	24	16416.91	0.750
Bagging_without_replacement	60	1000	36	16428.75	0.780
Bagging_without_replacement	60	1000	48	16453.14	0.760
Subsampling	30	100	12	17165.90	0.700
Subsampling	30	100	18	13965.11	0.550
Subsampling	30	100	24	9980.19	0.400
Subsampling	60	100	24	12253.03	0.640
Subsampling	60	100	36	9822.92	0.520
Subsampling	60	100	48	6949.98	0.370
Subsampling	30	1000	12	17245.27	0.670
Subsampling	30	1000	18	14230.53	0.550
Subsampling	30	1000	24	10112.13	0.410
Subsampling	60	1000	24	12336.30	0.630
Subsampling	60	1000	36	10034.51	0.510
Subsampling	60	1000	48	7141.78	0.350
Extended	30	100	-	21392.68	0.760
Extended	60	100	-	16125.62	0.730
Extended	30	1000	-	21603.71	0.770
Extended	60	1000	-	16038.97	0.780

Table 5 Results for farmer ($z^* \approx$ -132750.32) with α=0.2 based on 100 replications.