Proceedings of the IISE Annual Conference & Expo 2022
K. Ellis, W. Ferrell, J. Knapp, eds.

Reliability Assessment of Scenarios for CVaR Minimization in Two-Stage Stochastic Programs

Sarah M. Ryan and Ghazal Shah Abadi

Iowa State University, Ames, IA, USA

Abstract

The mass transportation distance rank histogram (MTDRh) was developed to assess the reliability of any given scenario generation process for a two-stage, risk-neutral stochastic program. Reliability is defined loosely as goodness of fit between the generated scenario sets and corresponding observed values over a collection of historical instances. This graphical tool can diagnose over- or under-dispersion and/or bias in the scenario sets and support hypothesis testing of scenario reliability. If the risk-averse objective is instead to minimize CVaR of cost, the only important, or effective, scenarios are those that produce cost in the upper tail of the distribution at the optimal solution. We describe a procedure to adapt the MTDRh for use in assessing the reliability of scenarios relative to the upper tail of the cost distribution. This adaptation relies on a conditional probability distribution derived in the context of assessing the effectiveness of scenarios. For a risk-averse newsvendor formulation, we conduct simulation studies to systematically explore the ability of the CVaR-adapted MTDRh to diagnose different ways that scenario sets may fail to capture the upper tail of the cost distribution near optimality. We conjecture that, as with the MTDRh and its predecessor minimum spanning tree rank histogram, the nature of the mismatch between scenarios and observations can be observed according to the non-flat shape of the rank histogram. On the other hand, scenario generation methods can be calibrated according to uniform distribution goodness of fit to the distribution of ranks.

Keywords
Risk-averse stochastic programming, scenario reliability assessment, goodness-of-fit

1. Introduction
Stochastic programming is a technique for optimization under uncertainty in which the uncertain parameters of a mathematical program are modeled as random variables. In two-stage stochastic programs, decisions are classified into first-stage decisions, which must be made before the realizations of the random variables are known, and second-stage ones that can provide recourse to the realizations observed. To compute optimal first-stage solutions, it is usually necessary to generate probabilistic scenarios for the uncertain parameter values. This task amounts to forming a discrete approximation of the random variables' joint distribution. If the resulting set of scenarios is too large, a subsequent step of scenario reduction may be used to find a coarser approximation.

The quality of the solution found, which is ideally measured in terms of closeness of its objective value (evaluated according to the true joint distribution) to that of the true optimal solution, depends on how well the probabilistic scenarios represent the true distribution. Under the risk-neutral objective to minimize expected cost, it is well known that an upper bound on the cost error is minimized by employing a set of scenarios with minimal Wasserstein (a.k.a., mass transportation) distance to the true distribution. However, the true distribution is itself a model of the uncertain future outcomes. Moreover, in applications of stochastic programming, researchers have observed that scenario sets formed by Wasserstein-based scenario reduction of large samples are outperformed by using heuristic scenario reduction methods that employ the cost function explicitly. The data-driven optimization philosophy also suggests that observational data on the uncertain parameter values should be exploited to the greatest extent possible.

This work is aimed at applications in which similar instances of a problem are to be solved continually and observational data are available for a collection of past instances. We define the *reliability* of a scenario generation process as the correspondence between the scenario probabilities and the relative frequencies of corresponding observations over these past instances [1]. The mass transportation distance rank histogram (MTDRh) was developed to assess the reliability of a scenario generation process for a two-stage, risk-neutral stochastic program. By comparing the scenarios generated for a collection of historical instances with the corresponding observed values, this tool can diagnose bias or a dispersion mismatch in the scenario sets. If the risk-averse objective is instead to minimize conditional value-at-risk (CVaR) of cost, the only important scenarios are those that produce cost in the upper tail of the distribution at the optimal solution. The goal of the experiments described in this paper is to test the ability of a CVaR-adapted MTDRh to diagnose different ways that scenario sets may fail to capture the upper tail of the cost distribution near optimality.

2. Problem Description

The general problem notation is adapted from [2]. Suppose the problem is to $\min_{x \in X} CVaR_{\alpha}\left[G(x,\xi)\right]$, the expected cost in the upper $(1-\alpha)$-probability tail of the cost distribution. We have a collection of *m* instances, for each of which we have generated *n* scenarios. For $i = 1,\dots,m$, let $\Xi_i = \{\xi_{i1}, \xi_{i2}, \dots, \xi_{in}\}$ be the set of scenarios for instance *i*, where ξ_{ij} has probability P_{ij} and $\sum_{j=1}^{n} P_{ij} = 1$, and let ξ_{i0} be the observed value of ξ in instance *i*.

For simplicity of solution, we focus on a newsvendor model, where the goal is to maximize the CVaR of profit. We reformulate the model of [3], simplified by removing the price-dependence of demand, as a risk-averse cost minimization where *x* is the order quantity and ξ is the uncertain demand for the product. Denoting the parameters as *c* = unit ordering cost, *s* = salvage value, and *p* = selling price ($p > c > s \geq 0$), the cost is given by:

$$G(x,\xi) = cx - p\min\{x,\xi\} - s(x-\xi)^{+} = (p-s)(x-\xi)^{+} - (p-c)x \quad , \tag{1}$$

where $z^{+} = \max\{0,z\}$. The authors of [3] show that $x^{*} = F^{-1}\left((1-\alpha)(p-c)/(p-s)\right)$, where $F(\cdot)$ is the demand distribution function; i.e., the critical fractile of demand is simply multiplied by $1-\alpha$ to minimize CVaR rather than expectation of cost (note that if $\alpha = 0$ the CVaR is equivalent to expectation). We assume the objective function parameters are fixed, while each instance has a different set of scenarios. That is, each day the newsvendor confronts the same economic situation but generates a new set of scenarios to approximate the demand distribution, and records the observed demand after the fact.

3. Background

The MTDRh was introduced in [4, 5] to assess wind power scenario reliability and its relationship to the quality of solutions to a stochastic unit commitment formulation for short term power system planning. The R package [6], coupled with a discrete goodness-of-fit test for uniformity of the ranks, has been employed to tune parameter estimation methods for simulating returns of stock indices to generate scenarios for financial asset allocation [7, 8].

In these and many other applications, risk-neutral objectives have been replaced or combined with CVaR as a coherent and tractable risk measure for stochastic programming (see, for example, [9]). However, some pitfalls of scenario generation for use in CVaR formulations have been noted. Because the tail of the cost distribution is difficult to represent accurately, large numbers of scenarios may be required [10, 11] to avoid "fragile" solutions [12], and standard scenario generation methods may fail to represent risk accurately [13]. A recently developed scenario reduction approach for risk-averse stochastic programming [2] that uses the concept of *effective* scenarios; i.e., those whose removal from the problem lead to changes in the optimal objective value, forms the basis for the reliability assessment method in this paper.

4. Approach

For each instance, *i*, we use Ξ_i and its distribution $P_{i\cdot}$ to solve the problem and obtain an optimal solution, x_i^*. The corresponding set of effective scenarios is defined as:

$$\overline{D}_i^* = \left\{ \xi \in \Xi_i : G\left(x_i^*, \xi\right) \geq VaR_\alpha\left[G\left(x_i^*, \cdot\right)\right]\right\}. \tag{2}$$

The effective scenarios are the subset that that determine the CVaR at optimality. We also partition the scenarios: $D_i^* = \left\{ \xi \in \Xi_i : G\left(x_i^*, \xi\right) < VaR_\alpha\left[G\left(x_i^*, \cdot\right)\right]\right\}$, $E_i^* = \left\{ \xi \in \Xi_i : G\left(x_i^*, \xi\right) = VaR_\alpha\left[G\left(x_i^*, \cdot\right)\right]\right\}$, $F_i^* = \Xi_i \setminus \left(D_i^* \cup E_i^*\right)$. Then we revise the scenario probabilities to $\hat{P}_{ij} = 0, j \in D_i^*$; $\hat{P}_{ij} = P_{ij} \big/ \left(1-\alpha\right), j \in F_i^*$; and $\hat{P}_{ij} = \left(1-\alpha-\sum_{j \in F_i^*} P_{ij}\right) \big/ \left(\left|E_i^*\right|\left(1-\alpha\right)\right), j \in E_i^*$. (Note this last expression is a special case of the third line of (4.10) in [2]). We use the revised probabilities to construct the MTDRh. The scenario set for instance *i* is replaced by the corresponding set of values of $G\left(x_i^*, \xi_{ij}\right), j = 1, \ldots n$, and the observation for instance *i* is replaced by $G\left(x_i^*, \xi_{i0}\right)$.

In our numerical experiment, we simulated $m = 200$ instances, with $n = 200$ scenarios randomly generated for each; all from Normal distributions. We fixed the values of selling price, purchase cost and salvage cost for each instance equal to $p = 15$, $c = 5$, and $s = 1$, respectively. We performed the numerical studies considering values for α of 0.5, 0.75, and 0.9. To test the ability of the MTDrh to identify whether the scenarios represented the tail of the cost distribution accurately, we generated all observations from the same standard normal distribution but systematically varied the parameters of the simulated scenario distribution. The steps of the simulation algorithm are summarized as:

Repeat steps 1 and 2 for $\alpha \in \{0.5,\ 0.75, 0.9\}$, $A \in \{0,\ 0.5, 1\}$, and $B \in \{0.3,\ 0.5, 0.8, 1, 2, 5\}$:

1. For $i = 1, \ldots, m$:
   Simulate *n* scenarios for demand from $N\left(\mu = A, \sigma = B\right)$;
   Simulate observed demand from $N\left(0,1\right)$;
   Solve for x_i^* and modify probabilities as explained in Section 4.
2. Construct the MTDrh, and compute *p*-values for a Cramér-von Mises test of uniformity.

5. Results

Figures 1-3 depict panels of histograms for each value of α. All scenarios and observations are generated from Normal distributions. Observations are generated from a standard Normal distribution throughout, but scenarios are generated from Normal distributions with the 18 combinations of *A* and *B*. The rows represent the three mean values of the scenario distributions, which also quantify their bias, and the columns correspond to the various values of the scenario standard deviation or, equivalently, its ratio to the observation standard deviation.

When the scenario distributions match that of the observations (i.e., $A = 0$, $B = 1$ in the first row, fourth column), the histogram appears to be flat, as expected. For biased scenarios, a mild upward slope is seen in the second and third rows of the fourth column. In the first three columns, the histograms slope downward from left to right, reflecting overpopulation in the low ranks because the scenarios are underdispersed relative to the observations. In the fifth and sixth columns, we expect from the development in [5] that the histograms would slope upward from left to right, because overdispersed scenarios should result in the high ranks being overpopulated. However, we instead observe hill-shaped histograms, which were observed in [5] to occur when time series scenarios differed from observations in terms of their autocorrelations. This effect grows more pronounced with larger values of α; i.e., when the CVaR is defined on a more extreme tail of the cost distribution.

The *p*-values of Cramér-von Mises hypothesis tests of the uniformity of ranks are provided in **Table 1**. As expected, the *p*-values for the instances having scenarios generated from a standard Normal distribution (the same distribution as observations) are large for all values of α. However, we also fail to reject the null hypothesis in some trials where the variance of scenarios is only somewhat smaller than that of the observations.

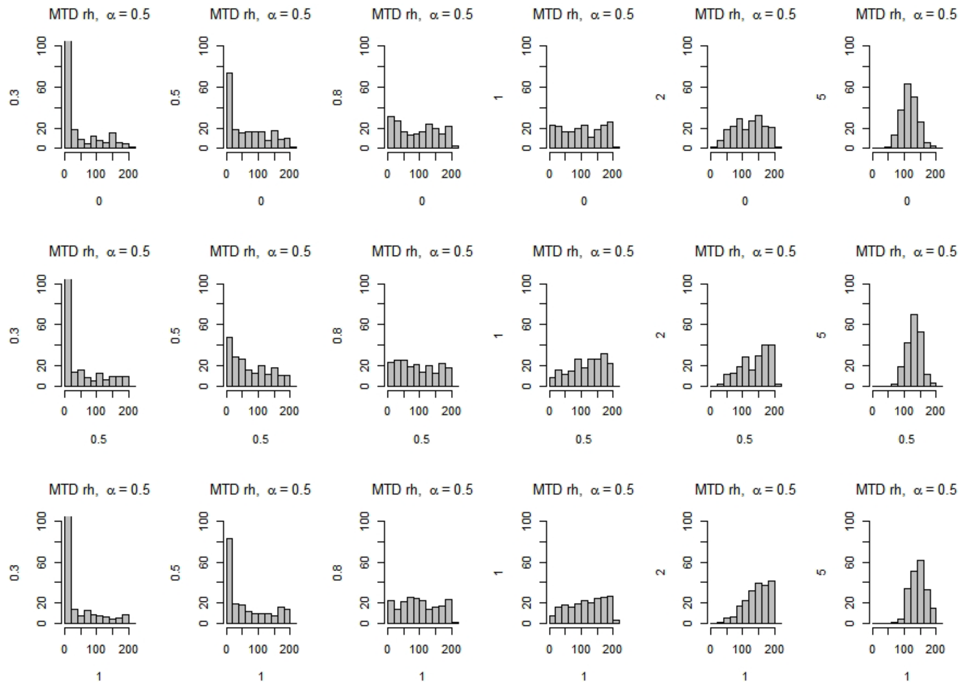

Figure 1: MTDrh simulation results for α = 0.5. Each row has a constant value of A, shown as a horizontal axis label, and each column columns has a constant value of B, labeling the vertical axis.

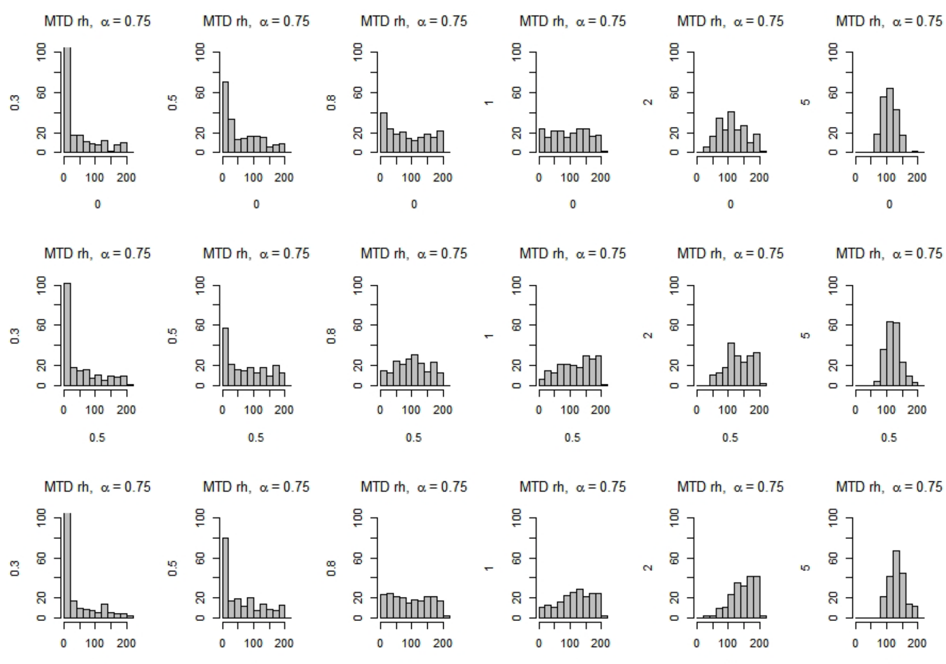

Figure 2: MTDrh simulation results for α = 0.75. Each row has a constant value of A, shown as a horizontal axis label, and each column columns has a constant value of B, labeling the vertical axis.

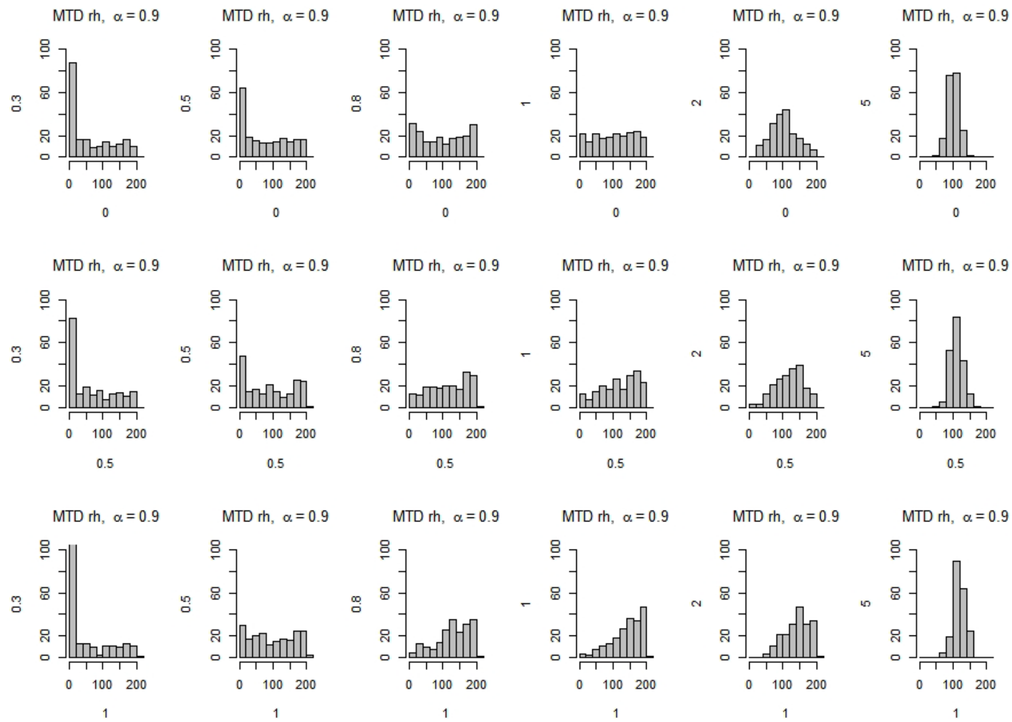

Figure 3: MTDrh simulation results for α = 0.9. Each row has a constant value of *A*, shown as a horizontal axis label, and each column columns has a constant value of *B*, labeling the vertical axis.

Table 1: *p*-values of the hypothesis test for uniformity of ranks.

	μ_{scen} \ $\sigma_{scen}/\sigma_{obs}$	0.3	0.5	0.8	1	2	5
$\alpha = 0.5$	0	≈ 0	≈ 0	0.089	0.537	≈ 0	≈ 0
	0.5	≈ 0	≈ 0	0.071	≈ 0	≈ 0	≈ 0
	1	≈ 0	≈ 0	0.586	≈ 0	≈ 0	≈ 0
$\alpha = 0.75$	0	≈ 0	≈ 0	≈ 0	0.826	≈ 0	≈ 0
	0.5	≈ 0	≈ 0	0.157	≈ 0	≈ 0	≈ 0
	1	≈ 0	≈ 0	0.415	≈ 0	≈ 0	≈ 0
$\alpha = 0.9$	0	≈ 0	≈ 0	0.104	0.629	≈ 0	≈ 0
	0.5	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0	≈ 0
	1	≈ 0	0.156	≈ 0	≈ 0	≈ 0	≈ 0

6. Conclusions and Final Thoughts

This paper reports the results of an initial set of simulation experiments designed to test the capability of the MTDRh to assess the reliability, or its lack, of probabilistic scenarios to represent the upper tail of a cost distribution at optimality. In our simulated newsvendor instances we notice some encouraging results. When scenario distributions match that of the observation, the resulting histogram is almost uniform, as detected visually and with a large *p*-value for the hypothesis test. On the other hand, an upward sloping histogram occurs when scenarios are biased and a downward slope can be seen, with generally smaller or negligible p-values, when scenarios are underdispersed. However, instead of the upward-sloping histograms expected when scenarios are overdispersed relative to the observations, we observe hill shapes, especially when attention is restricted to an extreme tail of the cost distribution.

One possible cause of the unexpected hill shapes is that scenarios with probabilities revised to zero were retained in the sets when computing the mass transportation distances, possibly distorting the ranks. We are experimenting with discarding such scenarios and, further, only including instances in which the observed cost falls in the upper tail. We also plan to test more types of distributions to investigate whether mismatch according to higher-order moments such as skewness and kurtosis can be diagnosed with this tool. In addition, for more complicated problems, solutions that are only approximately optimal can be used in the procedure to revise scenario probabilities.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. EFMA-2132200.

References

[1] D. Sari Ay and S. M. Ryan, "Observational data-based quality assessment of scenario generation for stochastic programs," *Computational Management Science,* vol. 16, pp. 521-540, 2019, doi: 10.1007/s10287-019-00349-1.

[2] S. Arpón, T. Homem-de-Mello, and B. Pagnoncelli, "Scenario reduction for stochastic programs with Conditional Value-at-Risk," *Mathematical Programming,* vol. 170, no. 1, pp. 327-356, 2018, doi: 10.1007/s10107-018-1298-9.

[3] Y. Chen, M. Xu, and Z. G. Zhang, "Technical Note—A Risk-Averse Newsvendor Model Under the CVaR Criterion," *Operations Research,* vol. 57, no. 4, pp. 1040-1044, 2009, doi: 10.1287/opre.1080.0603.

[4] D. Sari, Y. Lee, S. Ryan, and D. Woodruff, "Statistical metrics for assessing the quality of wind power scenarios for stochastic unit commitment," *Wind Energy,* vol. 19, pp. 873–893, 2016, doi: 10.1002/we.1872.

[5] D. Sari and S. M. Ryan, "Statistical Reliability of Wind Power Scenarios and Stochastic Unit Commitment Cost," *Energy Systems,* vol. 9, no. 4, pp. 873-898, 2018, doi: 10.1007/s12667-017-0255-7.

[6] *Package MTDrh*. (2016). The Comprehensive R Archive Network. [Online]. Available: https://cran.r-project.org/package=MTDrh

[7] X. Guo and S. M. Ryan, "Reliability assessment of scenarios generated for stock index returns incorporating momentum," *International Journal of Finance & Economics,* vol. 26, pp. 4013-4031, 2021, doi: 10.1002/ijfe.2002.

[8] X. Guo and S. M. Ryan, "Portfolio rebalancing based on time series momentum and downside risk," *IMA Journal of Management Mathematics,* 2021, doi: 10.1093/imaman/dpab037.

[9] R. Khodabakhsh and S. Sirouspour, "Optimal Control of Energy Storage in a Microgrid by Minimizing Conditional Value-at-Risk," *IEEE Transactions on Sustainable Energy,* vol. 7, no. 3, pp. 1264-1273, 2016, doi: 10.1109/tste.2016.2543024.

[10] P. Krokhmal, J. Palmquist, and S. Uryasev, "Portfolio optimization with conditional value-at-risk objective and constraints," *Journal of Risk,* 2002.

[11] J. Fairbrother, A. Turner, and S. W. Wallace, "Scenario Generation for Single-Period Portfolio Selection Problems with Tail Risk Measures: Coping with High Dimensions and Integer Variables," *INFORMS Journal on Computing,* vol. 30, no. 3, pp. 472-491, 2018, doi: 10.1287/ijoc.2017.0790.

[12] A. E. B. Lim, J. G. Shanthikumar, and G.-Y. Vahn, "Conditional value-at-risk in portfolio optimization: Coherent but fragile," *Operations Research Letters,* vol. 39, no. 3, pp. 163-171, 2011, doi: 10.1016/j.orl.2011.03.004.

[13] J. Fairbrother, A. Turner, and S. W. Wallace, "Problem-driven scenario generation: an analytical approach for stochastic programs with tail risk measure," *Mathematical Programming,* 2019, doi: 10.1007/s10107-019-01451-7.