Scenario estimation and generation

Georg Ch. Pflug

October 10, 2004
The problem

Discrete Approximations
   The one-period case
   Distances of Probability Measures

Tree construction
   Tensor products of trees
   Tree reduction

Examples

Georg Ch. Pflug
Scenario estimation and generation
A decision problem is subject to uncertainty
Uncertainty is represented by probability
To reduce computational complexity, a discrete probability with only few mass points is preferred

The discrete probability distribution representing uncertainty is called the \textit{scenario model}.
The two phases of scenario generation

- Phase 1. Find the correct (appropriate, reasonable) probability distribution - the uncertainty model
- Phase 2. Find a scenario model, such that the information loss going from the uncertainty model to the scenario model is small

Phase 1 is done by
- Statistical Estimation
- Expert opinion
- Regulator’s rules
- Any combination of the above
Multistage decisions and multistage scenarios

\(\xi_1, \ldots, \xi_T\) a multivariate time series process (e.g. future interest rates, future asset prices, etc.)
\(\mathcal{F}_0, \ldots, \mathcal{F}_{T-1}\) a filtration
$x_0, \ldots, x_{T-1}$ decisions

$t = 0$

observation of the r.v. $\xi_1$

decision $x_0$

$t = t_1$

observation of the r.v. $\xi_2$

decision $x_1$

$t = t_2$

observation of the r.v. $\xi_3$

decision $x_2$

$t = t_3$

observation of the r.v. $\xi_3$

decision $x_3$
The success of the strategy is expressed in terms of a cost function

\[ H(x_0, \xi_1, x_1, \ldots, x_{T-1}, \xi_T) \]

The final objective is to minimize a functional \( F \) of the cost success function, such as the expectation, a quantile or some other functional

\[ F[H(x_0, \xi_1, x_1, \ldots, x_{T-1}, \xi_T)] \]

Minimize in \( x_0, x_1(\xi_1), \ldots, x_{T-1}(\xi_1, \ldots, \xi_{T-1}) : F[H(x_0, \xi_1, \ldots, x_{T-1}, \xi_T)] \)
under the information constraints as well as other constraints

Information (measurability) constraints

\[
\begin{align*}
x_0 &= \text{a constant} & \mathcal{F}_0 \text{ measurable} \\
x_1 &= x_1(\xi_1) & \mathcal{F}_1 \text{ measurable} \\
& \vdots & \\
x_{T-1} &= x_{T-1}(\xi_1, \xi_2, \ldots, \xi_{T-1}) & \mathcal{F}_{T-1} \text{ measurable}
\end{align*}
\]
Stochastic processes

- See a stochastic process $\xi_1, \ldots, \xi_T$ as a random element in $\mathbb{R}^T$
- See a stochastic process and a sequence of transition probabilities $P\{\xi_1 \in A_1\}$, $P\{\xi_2 \in A_2|\xi_1\}$, $P\{\xi_3 \in A_3|\xi_1, \xi_2\}$, ...

Only the second view models the information structure along with the uncertainty structure.
Tree Approximations

The transition probability view of a stochastic process leads to tree processes in the discrete world. A stochastic process \( \nu_1, \ldots, \nu_T \) is a tree process, if all conditional distributions \( P\{\nu_1, \ldots, \nu_{t-1}|\nu_t\} \) are degenerated or - equivalently - \( \sigma(\nu_t) = \sigma(\nu_1, \nu_2, \ldots, \nu_t) \). Every stochastic process is a function of a tree process.
The approximating problem

\[ \begin{align*}
\text{Minimize in } & x_0, x_1(\tilde{\xi}_1), \ldots, x_{T-1}(\tilde{\xi}_1, \ldots, \tilde{\xi}_{T-1}) : \mathbb{F}[H(x_0, \tilde{\xi}_1, \ldots, x_{T-1}, \tilde{\xi}_T)] \\
& \text{under the information constraints as well as other constraints}
\end{align*} \]

Information (measurability) constraints

\[
\begin{align*}
x_0 &= \text{a constant} & \tilde{\mathcal{F}}_0 \text{ measurable} \\
x_1 &= x_1(\tilde{\xi}_1) & \tilde{\mathcal{F}}_1 \text{ measurable} \\
x_2 &= x_2(\tilde{\xi}_1, \tilde{\xi}_2) & \tilde{\mathcal{F}}_2 \text{ measurable} \\
\vdots & & \vdots \\
x_{T-1} &= x_{T-1}(\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_{T-1}) & \tilde{\mathcal{F}}_{T-1} \text{ measurable}
\end{align*}
\]

Error = | min(·, ·· ·, ·) − min(˜, ·· ·, ˜)|.
The one-period case

\[ G \text{ original distribution function} \]
\[ \tilde{G} \text{ discrete approximation} \]

\[ \tilde{G} : \begin{array}{c|cccc}
\text{values} & v_1 & v_2 & \cdots & v_m \\
\text{probabilities} & p_1 & p_2 & \cdots & p_m \\
\end{array} \]
Distances of Probability measures

Let $\mathcal{H}$ be a family of functions. Define a semidistance by

$$d_\mathcal{H}(G_1, G_2) = \sup\{| \int h(w) \, dG_1(w) - \int h(w) \, dG_2(w) | : h \in \mathcal{H} \}.$$

Examples.

- If $\mathcal{H}$ is the class of the indicator functions $1_A$ of all measurable sets, the variational distance is obtained

  $$d_V(G_1, G_2) = \sup\{ \int 1_A(u) \, d[G_1 - G_2](u) : A \text{ measurable} \}.$$

- If $\mathcal{H}$ is the class of indicator functions of half-unbounded rectangles in $\mathbb{R}^k$ the Kolmogorov-Smirnov distance is obtained

  $$d_{KS}(G_1, G_2) = \sup\{| G_1(x) - G_2(x) | : x \in \mathbb{R}^k \}.$$

Georg Ch. Pflug

Scenario estimation and generation
If \( \mathcal{H} \) is the class of Lipschitz-continuous functions with Lipschitz-constant 1 the Wasserstein distance is obtained.

\[
d_W(G_1, G_2) = \sup\{ \int h \ dG_1 - \int h \ dG_2 : |h(u) - h(v)| \leq |u - v| \}.
\]

If \( Hc \) is the class of Lipschitz functions of order \( p \), the Fortet-Mourier distance is obtained.

\[
d_{FM_p}(G_1, G_2) = \sup\{ \int h \ dG_1 - \int h \ dG_2 : L_p(h) \leq 1 \}
\]

where

\[
L_p(f) = \inf\{ L : |h(u) - h(v)| \leq L|u - v| \max(1, |u|^{p-1}, |v|^{p-1}) \}.
\]
Suppose that the stochastic optimization problem is to maximize \( \int h(x, u) \, dG(u) \). If \( \mathcal{H} \) is a family of functions containing the functions \( \{h(x, \cdot) : x \in \mathbb{X}\} \). Then defining a semidistance as

\[
d_{\mathcal{H}}(G_1, G_2) = \sup \{ | \int h(w) \, dG_1(w) - \int h(w) \, dG_2(w) | : h \in \mathcal{H} \}.
\]

entails that

\[
\text{Error} \leq \sup_x | \int h(x, w) \, dG_1(w) - \int h(x, w) \, dG_2(w) | \leq d_{\mathcal{H}}(G_1, G_2).
\]

Typical integrands are of the form

\[
u \mapsto \sum_i \alpha_i(x)[c_i(x) \cdot u - d_i(x)]^+
\]

i.e. they are Lipschitz in \( u \).
KS-distance and W-distance: A comparison

Let $\xi_1 \sim G_1$, $\xi_2 \sim G_2$

$$d_{KS}(G_1, G_2) = \sup_x \{ |G_1(u) - G_2(u)| \}$$

$$= \sup \{ |\mathbb{P}(\xi_1 \leq u) - \mathbb{P}(\xi_2 \leq u)| \}. $$

Hlawka-Koksma Inequality:

$$\int h(u) \, dG_1(u) - \int h(u) \, dG_2(u) \leq d_{KS}(G_1, G_2) \cdot V(h).$$

where $V(h)$ is the total variation of $h$

$$V(h) = \sup \left\{ \sum_i |h(z_i) - h(z_{i-1})| : z_1 < z_2 < \cdots < z_n, n \text{ arbitrary} \right\}. $$
If $K$ is a monotonic function, then

$$d_{KS}(K(\xi_1), K(\xi_2)) = d_{KS}(\xi_1, \xi_2).$$

Therefore, using the quantile transform, the problem reduces to approximate the uniform $[0,1]$ distribution. If

$$\begin{array}{cccc}
u_1 & u_2 & \cdots & u_m \\
p_1 & p_2 & \cdots & p_m
\end{array}$$

approximates the uniform $[0,1]$ distribution with $d_{KS}$ distance $\epsilon$, then

$$\begin{array}{cccc}
G^{-1}(u_1) & G^{-1}(u_2) & \cdots & G^{-1}(u_m) \\
p_1 & p_2 & \cdots & p_m
\end{array}$$

approximates $G$ also with $d_{KS}$ distance $\epsilon$. 
The optimal approximation of a continuous distribution $G$ by a distribution sitting on mass points $v_1, \ldots, v_n$ with probabilities $p_1, \ldots, p_n$ w.r.t. the KS distance is given by

$$z_i = G^{-1}\left(\frac{2i - 1}{2n}\right), \quad p_i = 1/n.$$  

The distance is $1/n$. 

\begin{center}
\includegraphics[width=0.5\textwidth]{graph.png}
\end{center}
Other KS approximations

- **The Monte Carlo method** If $U_i$ is a sequence of i.i.d. random variables from a uniform $[0,1]$, then distribution $\tilde{G}_n$ sitting on points $v_i = G^{-1}(U_i)$ with masses $1/n$ satisfies

$$\mathbb{P}\{\sup_v |G(v) - \tilde{G}_n(v)| \geq \epsilon/\sqrt{n}\} \leq 58 \exp(-2\epsilon^2).$$

(Dvoretzky, Kiefer, Wolfowitz inequality)

- **The Quasi Monte Carlo method** A sequence $u_i$ is called a low discrepancy sequence in $[0,1]$ (e.g. the Sobol sequence), if the KS distance between the uniform distribution and the empirical distribution based on $(u_i)$ is $O(\log(n)/n)$. Then with $v_i = G^{-1}(u_i)$ with masses $1/n$ we have

$$d_{KS}(G, \tilde{G}_n) \leq C \frac{\log n}{n} \quad \text{for all } n.$$
The W-distance and the facility location problem

\[ d_W(G_1, G_2) = \sup \left\{ \int h \, dG_1 - \int h \, dG_2 : |h(u) - h(v)| \leq |u - v| \right\}. \]

**Theorem (Kantorovich-Rubinstein).** Dual version of Wasserstein-distance:

\[ d_W(G_1, G_2) = \inf \{ \mathbb{E}(\|X - Y\|) : (X, Y) \text{ is a bivariate r.v. with given marginal distribution functions } G_1 \text{ and } G_2 \}. \]
Remark If both measures sit on a finite number of mass points \( \{v_1, v_2, \ldots \} \), then \( d_W(G_1, G_2) \) is the optimal value of the following linear optimization problem:

\[
\begin{align*}
\text{Minimize} & \quad \sum_{i,j} \kappa_{ij} r_{ij} \\
\text{subject to} & \quad \sum_i \kappa_{ij} = \mu_j \quad \text{for all } j \\
& \quad \sum_j \kappa_{ij} = \nu_i \quad \text{for all } i
\end{align*}
\]

Here \( \mu_i \) resp. \( \nu_i \) is the mass sitting on \( v_i \) and \( r_{i,j} = \| v_i - v_j \| \).

Interpretation as an optimal mass transportation problem:
For an infinite space, it is in general impossible to find the explicit form of the joint distribution of $X$ and $Y$ which minimizes $\mathbb{E}(\|X - Y\|)$. For the case $k = 1$ and for a convex function $\psi$, the explicit solution of the mass transportation problem is known: 

**Lemma.** (Dall’ Aglio (1972))

$$\inf \left\{ \mathbb{E}(\psi(X - Y)) : (X, Y) \text{ is a bivariate r.v. with marginals } G_1 \text{ and } G_2 \right\}$$

$$= \int_0^1 \psi(G_1^{-1}(u) - G_2^{-1}(u)) \, du.$$ 

**Corollary.** The Wasserstein distance on $\mathbb{R}^1$ satisfies

$$d_W(G_1, G_2) = \int |G_1^{-1}(u) - G_2^{-1}(u)| \, du = \int |G_1(u) - G_2(u)| \, du.$$
An example

Suppose we want to approximate the $t$-distribution with 2 degrees of freedom by a probability measure sitting on five points. Using the $KS$- distance one gets the solution $\tilde{G}_1$

<table>
<thead>
<tr>
<th>value</th>
<th>-1.8856</th>
<th>-0.6172</th>
<th>0</th>
<th>0.6172</th>
<th>1.8856</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Using the Wasserstein distance one gets $\tilde{G}_2$

<table>
<thead>
<tr>
<th>value</th>
<th>-4.58</th>
<th>-1.56</th>
<th>0</th>
<th>1.56</th>
<th>4.58</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability</td>
<td>0.0446</td>
<td>0.2601</td>
<td>0.3906</td>
<td>0.2601</td>
<td>0.0446</td>
</tr>
</tbody>
</table>

Scenario estimation and generation
Example (continued)- A newsboy problem

Consider the family of functions

\[ w \mapsto h(w, a) = [w - a]^+ + 0.3[w - a]^- \]

We have plotted

\[ a \mapsto \int h(w, a) \, dG(w) \] (solid)
\[ a \mapsto \int h(w, a) \, d\tilde{G}_1(w) \] (dotted)
\[ a \mapsto \int h(w, a) \, d\tilde{G}_2(w) \] (dashed).

One sees that \( \tilde{G}_2 \) gives a better approximation and the location of the minimum is also closer to the true one.
Finding the WD approximation

Calculating a W-distance minimal $n$-point approximation $\tilde{G}_n$ of a distribution $G$ equals solving a transportation problem. The goal is find the set of borders $B b_i, i = 1, \ldots, n + 1$ which solve the problem

$$\text{minimize } \sum_{i=1}^{n} \int_{b_i}^{b_{i+1}} |g(x) - v_i| \, d(x)$$

The $i$th transportation-center $v_i$ can be calculated as follows:

$$v_i = G^{-1}(\frac{1}{2} \int_{b_i}^{b_{i+1}} g(x) \, dx - \int_{b_1}^{b_i} g(x) \, dx)$$
Global Optimization

Branch and Bound can be applied to calculate the global minimum. Divide search space into $N$ sections and create the B&B tree recursively: Set the first border $b_2$ to all $N - n + 1$ positions, these form subproblems with $n - 1$ transportation centers. This method lowers the dimension of the problem. Upper bounds are evaluated at the leaves of the B&B tree and the lower bounds can be calculated by finding the minimum possible transportation cost within two borders.

Local Optimization

Meta-heuristics may be applied to find local minima much faster. E.g. a combination of iterative local search with simulated annealing:

1. Generate initial set of borders $B$ (e.g. KS-distance optimal borders)
2. Move border $b_x$ which best improves the optimal value of problem.
3. If no border can be moved improving the optimal value, start simulated annealing (Randomly choose a border, if movement improves the optimal value, move it, otherwise draw random number and decide whether to move it or not, lower temperature, ...)

Georg Ch. Pflug
Approximation: Fortet-Mourier distance

Via transformation approximable with Wasserstein distance minimization:

\[ \chi^p(u) = \begin{cases} 
    u & |u| \leq 1 \\
    \text{sign}(u) \cdot |u|^p & |u| > 1 
\end{cases} \]

- Choose \( p \): Transform \( G \) with \( \chi^p \rightarrow G^{1/p} = G \circ \chi^{1/p} \)
- Approximate \( G^{1/p} \) by \( \tilde{G}^{1/p} \) with \( \text{min(Wasserstein)} \)
- Backtransformation: \( \tilde{G} = \tilde{G}^{1/p} \circ \chi^p \)
Approximation methods: Summary

- Full Monte Carlo
- Quasi Monte Carlo
- Moment matching
- KS-approximation
- Wasserstein and Fortet-Mourier Approximation
- some more
Moment matching?

- Advantage: Easy to do.
- Moments do not characterize the distribution. The "moment distance" $d(G, \tilde{G}) = \sup_k | \int u^k \, dG(u) - \int u^k \, d\tilde{G}(u) |$ is not a distance.

![Graph showing moment matching](image)
Phase 1: Estimation and simulation of paths

Parametric approach: Use data, estimate (time-series) model, simulate paths over whole time/all stages.

Dow Jones Industrial: $n = 10$ simulated GARCH(1,1) paths
Phase 2: Generating multi-variate, multi-period trees

**Step 1.** Approximate data in the first stage (e.g. Wasserstein distance)
Generating multi-variate, multi-period trees

**Step 2.** For subsequent stages use paths conditional to predecessor
Generating multi-variate, multi-period trees

**Step 3.** Iterate through the whole tree
Mathematics behind

Let $P_t(A_t | x^{t-1})$ be the conditional probability of $\xi_t \in A_t$ given the past $\xi^{(t-1)} = (\xi_1, \ldots, \xi_{(t-1)}) = x^{t-1} = (x_1, \ldots, x_{t-1})$.

Let $\tilde{P}$ another such probability, again dissected into its chain of conditional probabilities. We define

$$\bar{d}(P, \tilde{P}) = d(P_1, \tilde{P}_1) + \sum_{t=2}^{T} \sup_{x^{(t-1)}} d(P_t(\cdot | x^{(t-1)}), \tilde{P}_t(\cdot | x^{(t-1)})).$$

where $d$ is the Wasserstein distance. For an empirical measure (or any finitely supported measure), the conditional probability does only exist except for finitely many conditions. However we define the distance as zero, if the conditional distribution is not defined. $\bar{d}$ does not fulfill the triangle inequality. However, $\bar{d}$ reflects the information structure. Notice that the usual empirical does not converge w.r.t. $\bar{d}$, only the tree empirical does.
Empricals and tree empricals

Ordinary empirical and tree empirical
Tensor products of tree processes

Suppose that \((\xi^{(1)}_t)\) and \((\xi^{(2)}_t)\) are two tree processes. There is just one way of constructing the joint tree process \((\xi^{(1)}_t, \xi^{(2)}_t)\) such that the two components are independent. However, more often one needs to construct a joint process with given correlation.
Joint distributions with given marginals and given correlation

Constructing a joint distribution $Q$ with given marginals $X = (x_i, p_{1i})$, $Y = (y_i, p_{2i})$ and given correlation $\rho$ can be calculated such that an additional (tree) reduction is obtained by solving the non-linear program

\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{m} \sum_{j=1}^{n} \eta(q_{ij}) \\
\text{subject to} & \quad \sum_{i=1}^{m} q_{ij} = p_{1j} \quad j = 1, \ldots, n \\
& \quad \sum_{j=1}^{n} q_{ij} = p_{2i} \quad i = 1, \ldots, m \\
& \quad \sum_{i=1}^{m} \sum_{j=1}^{n} q_{ij} x_i y_j = \mathbb{E}(X)\mathbb{E}(Y) - \rho \sqrt{\text{Var}(X)\text{Var}(Y)}
\end{align*}
\]

with

\[
\eta(x) = \begin{cases} 
0 & \text{if } x = 0, \\
M & \text{if } x > 0.
\end{cases}
\]
An example

Two asset categories, first stage, 3 nodes each.

<table>
<thead>
<tr>
<th>Stocks</th>
<th>$x_i$</th>
<th>1.12</th>
<th>1.03</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$p_i$</td>
<td>0.21</td>
<td>0.52</td>
<td>0.27</td>
</tr>
<tr>
<td>Bonds</td>
<td>$x_i$</td>
<td>1.054</td>
<td>1.04</td>
<td>1.031</td>
</tr>
<tr>
<td></td>
<td>$p_i$</td>
<td>0.32</td>
<td>0.42</td>
<td>0.26</td>
</tr>
</tbody>
</table>

Joint distribution with $\rho = 0.4$ and reduction:
Tree reduction

- Distance based tree reduction (Dupacova, Groewe, Roemisch; Heitsch)
- EVPI based tree reduction (Dempster, Thompson, Consigli)
Distance based tree reduction

Rule: If two branches of the tree are closer in distance than some threshold, collapse the two.
EVPI based tree reduction

EVPI = Feasible minimum - Clairvoyant’s minimum

Feasible minimum = \( \min \{ \mathbb{F}[H(x_0, \tilde{\xi}_1, \ldots, x_{T-1}, \tilde{\xi}_T)] : \\
\quad x_0, x_1(\tilde{\xi}_1), \ldots, x_{T-1}(\tilde{\xi}_1, \ldots, \tilde{\xi}_{T-1}) \} \)

Clairvoyant’s minimum = \( \min \{ \mathbb{F}[H(x_0, \tilde{\xi}_1, \ldots, x_{T-1}, \tilde{\xi}_T)] : \\
\quad x_0(\tilde{\xi}_1, \ldots, \tilde{\xi}_{T-1}), \ldots, x_{T-1}(\tilde{\xi}_1, \ldots, \tilde{\xi}_{T-1}) \} \)

EVPI may be defined for every subtree (given the optimal decision in earlier stages). This EVPI process is a supermartingale. However, there is no relation between the EVPI’s of full (finer) model and a reduced (coarser) model.
EVPI $\geq$ $\tilde{\text{EVPI}}$.
Empirical convergence of central moments (1)

Data: Weekly data of DJI and 10-year T-Bonds: 01/1993-01/1999
Approximation: Wasserstein, Monte Carlo, Sobol Sequence, Moment matching

First and second central moment approximation ($n = 2 : 35$)
Empirical convergence of central moments (2)

Third and fourth central moment approximation \((n = 2 : 35)\)

Moment matching (Kaut [2003]) started to match at \(n = 15 : 20\).
Approximation errors: Single-stage portfolio selection

Mathematical models \( \mathcal{M} \sim \textbf{Risk} \) functionals \( \mathbb{F} \):

- CVaR\(_\alpha\) : \( \inf \{ a + \frac{1}{1-\alpha} \mathbb{E}[Y - a]^+ : a \in \mathbb{R} \} \)
- Mean Absolute Deviation (MAD): \( \sum_{i=1}^{n} |x_i - \mathbb{E}(X)| \)

**Approximation** methods \( S \):

- Moment matching (Høyland et al. [2003])
- Probability distance minimization (Wasserstein distance)
- Quasi-random sampling (Low discrepancy series: Sobol sequence)
Risk measures: Linear programming formulations

**Mean Absolute Deviation** [Konno/Yamazaki 1991][Feinstein/Thapa 1993]

maximize \( \mathbb{E}(x, \xi) - \sum_{s=1}^{S} p_s z_s \)
subject to \( z_s - \sum_{s=1}^{S} p_s z_s = 1, \ldots, S \)
\( x \in \mathcal{X} \)

**Conditional Value at Risk** [Rockafellar/Uryasev 1999][Uryasev 2000]

maximize \( \gamma - \frac{1}{1-\alpha} \sum_{s=1}^{S} p_s z_s \)
subject to \( z_s - \sum_{s=1}^{S} p_s z_s = 1, \ldots, S \)
\( x^T e \geq \mathbb{E}(x, \xi) \)
\( x \in \mathcal{X} \)

Portfolio constraints \( \mathcal{X} : \sum_{a=1}^{A} x_a \leq 1, \forall a : \max(x_a) = 0.5, \ldots \)
Empirical results

- 13 (somewhat randomly selected) assets: AOL, C, CSCO, DIS, EMC, GE, HPQ, MOT, NT, PFE, SUNW, WMT, XOM
- Rolling horizon: asset allocation, backtracking
- Optimize MAD and CVaR with full data and approximations
Empirical results: MAD

Example (Asset Allocation): Data 01/1993-01/1995, $n = 50$

- Full dataset
- Wasserstein
- Moment matching
- Sobol sequence
Empirical results: $\text{CVaR}_\alpha$

Example (Asset Allocation): Data 01/1999-01/2001, $\alpha = 0.1$, $n = 50$
Scenario Generation and Backtracking (1)

Expected Shortfall (CV@R), $\alpha = 0.4$, $n = 50$
Scenario Generation and Backtracking (2)

Mean Absolute Deviation (MAD), $n = 50$
Scenario Generation and Backtracking (3)

Expected Shortfall (CV@R), $\alpha = 0.4$, $n = 20$ - Moment Matching failed, substituted with Quasi Random Sampling
Scenario Generation and Backtracking (4)

Mean Absolute Deviation (MAD), $n = 20$ - Moment Matching failed, substituted with Quasi Random Sampling
Scenario Generation and Backtracking (5)

Expected Shortfall (CV@R), $\alpha = 0.4$, $n = 150$
Scenario Generation and Backtracking (6)

Mean Absolute Deviation (MAD), \( n = 150 \)