Simulating Risk Measures

Wei Jiang and Steven Kou

National University of Singapore

Outline

Outline

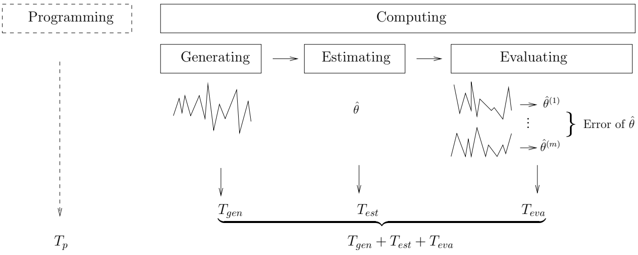
Risk measures

- VaR (Value-at-risk) and ES (Expected shortfall), defined as $v(p) = \inf\{x : F(x) \geq p\}$, $c(p) = E(L|L \leq v(p))$, respectively.
- VaR and ES are widely used risk measures in risk management, e.g. in Basel Accords.
- Among all approaches to estimate VaR and ES, the Monte Carlo simulation approach is widely used,
  - to incorporate various risk models
  - to analyze complex portfolios
- **Efficiency**: simulation cost and accuracy.

Simulation cost

- Total time involved: $T_p + T_{gen} + T_{est} + T_{eva}$.
- Total execution time without the programming time: $T_{gen} + T_{est} + T_{eva}$.

Figure : Standard simulation procedures



Relative Error

- Mean of estimate is 0.1, the SD is 0.3; SD is small but not good.
- Relative error (RE) (Glasserman, 2003):

$$RE(\hat{\theta}_n) = \frac{\sqrt{Var(\hat{\theta}_n)}}{|E\hat{\theta}_n|},$$

where $\hat{\theta}_n$ is an estimator.

Existing methods

- For IID samples: Indirect crude Monte Carlo (Glasserman and Ruiz-Mata, 2006).
- For IID samples: Importance sampling (IS) (indirect: Glasserman et al., 2000; Fuh, 2011; direct: Hong and Sun, 2010).
- For dependent samples (Many financial return series are subject to data dependence, e.g. GARCH, stochastic vol, affined jump diffusion models): Maximum Transformation (Heidelberger and Lewis, 1984)
- Most of the papers did not consider the total execution time, ignoring the cost in evaluating part
- No reported RE

Compare with Heidelberger and Lewis (1984)

This paper	Heidelberger & Lewis (1984)
$X_1, \cdots, X_n$	$X_1, \cdots, X_n$
$\Downarrow$ sort	$\Downarrow$ regroup, find minimal
$X_{(1)}, \cdots, X_{(n)}$	$Y_1, \cdots, Y_m$, $n = mv$
	$\Downarrow$ sort
	$Y_{(1)}, \cdots, Y_{(m)}$
$\Downarrow$	$\Downarrow$
$v_n = X_{([np])}$	$v_n = Y_{([mp^v+1])}$

Note: v_n is the estimator of the quantile.

Compare with Heidelberger and Lewis (1984)

	This paper	Heidelberger and Lewis (1984)
Risk measures	VaR & ES	VaR
Stationary	preserved	not preserved
$\alpha-$mixing	preserved	not preserved
Storage cost	not reduced	reduced
Variance	not inflated	inflated
RE expansion	yes	no

Literature

Table : Some literature on simulating VaR and ES

Panel A: Simulation Methods	
A.1: IID setting	
Glasserman et al. (2000, 2002)	indirect importance sampling (IIS) and indirect importance sampling with stratification (IIS-Q), VaR
Hong and Sun (2010)	direct importance sampling, VaR & ES
Fuh et al. (2011)	indirect importance sampling on portfolio with heavy-tailed risk factors, VaR
A.2: Dependent setting	
Heidelberger and Lewis (1984)	maximum transformation, VaR
This paper	sorted Monte Carlo with reported RE, VaR & ES
Panel B: Relative Error	
Yoshiba and Yamai (2002)	i.i.d, numerical, VaR & ES
Hult and Svensson (2009)	i.i.d, heuristic, VaR & ES
This paper	dependent, theoretical, VaR & ES

Contributions

Our contribution is twofold:

- We propose a general sorting method, allowing dependent samples, to simulate risk measures. The sorting method will preserve stationarity and α mixing properties. Numerical experiments indicate the method is easy to implement and fast, compared to existing methods, in terms of total execution time, even at the level 0.001.
- We rigorously derived approximations for RE's of VaR and ES under dependent setting. With the approximate formulas we are able to compute the necessary sample size needed for VaR and ES.

Outline

1. Introduction
2. An Algorithm: Sorted Monte Carlo with Reported Relative Errors
3. Expansions for Relative Errors
4. Comparison of Relative Errors

Classification: dependent samples

- **Type A:** dependent samples from time series stationary model.
  - Many financial data is subject to stationary time series model (Chen and Tang, 2005; Chen, 2008). Assume that financial losses $\{L_s\}_{s=1}^{T}$ satisfy some stationary time series model.
  - Traditional way: (drop mN (i.e. $m \geq 10000$) samples)

$$\begin{array}{ll} L_1^{(1)}, \cdots, L_m^{(1)}, & L_{m+1}^{(1)} \\ L_1^{(2)}, \cdots, L_m^{(2)}, & L_{m+1}^{(2)} \\ \cdots & \\ L_1^{(N)}, \cdots, L_m^{(N)}, & L_{m+1}^{(N)} \end{array}$$

  - Time series samples: (only drop m (i.e. $m \geq 10000$) samples)

$$L_1, \cdots, L_m, L_{m+1}, \cdots, L_{m+N}$$

Classification: i.i.d samples

- **Type B:** IID samples from time series non-stationary model (Glasserman et al., 2000; Glasserman et al., 2002; Fuh et al., 2011)
  Simulate samples repeatedly.
  Simulate N sample path to get loss samples at time t:

$$\begin{array}{cc} L_1^{(1)}, \cdots, L_{t-1}^{(1)}, & L_t^{(1)} \\ L_1^{(2)}, \cdots, L_{t-1}^{(2)}, & L_t^{(2)} \\ \cdots & \\ L_1^{(N)}, \cdots, L_{t-1}^{(N)}, & L_t^{(N)} \end{array}$$

Estimators

- For n samples of loss: $L_1, \cdots, L_n$, VaR and ES are estimated as follows:

$$v_n(p) = \inf\{x : F_n(x) \geq p\} \quad (1)$$

$$c_n(p) = v_n(p) - \frac{1}{np}\sum_{i=1}^{n}[v_n(p) - L_i]^+ \quad (2)$$

where $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{L_i \leq x\}}$ is the empirical distribution of L.

- Allowing dependent samples, model free sorting method.

Our RE approximation

- Under some regular dependence setting:

$$RE(v_n) \approx -\frac{\sigma_{n,v}}{vf(v)} n^{-1/2}, \quad RE(c_n) \approx -\frac{\sigma_{n,c}}{cp} n^{-1/2} \tag{3}$$

where, $\sigma^2_{n,v} = \{p(1-p) + 2\sum_{k=1}^{(n-1)} \gamma_1(k)\}$,
$\gamma_1(k) = cov\{1_{\{L_1<v\}}, 1_{\{L_{k+1}<v\}}\}$;
$\sigma^2_{n,c} = \{Var[(v-L_1)^+] + 2\sum_{k=1}^{(n-1)} \gamma_2(k)\}$,
$\gamma_2(k) = cov\{(v-L_1)^+, (v-L_{k+1})^+\}$.

  - MA(q) Model: $\sigma^2_{n,v} = p(1-p) + 2\sum_{k=1}^{q} \gamma_1(k)$,
    $\sigma^2_{n,c} = Var[(v-L_1)^+] + 2\sum_{k=1}^{q} \gamma_2(k)$.
  - IID samples: $\sigma^2_{n,v} = p(1-p)$, and $\sigma^2_{n,c} = Var[(v-L_1)^+]$.

- Compute sample size when controlling the RE level to be α by solving Equations (3):

$$n^v_\alpha \approx \frac{\sigma^2_{n,v}}{v^2 f(v)^2 \alpha^2}, \quad n^v_\alpha \approx \frac{\sigma^2_{n,c}}{c^2 p^2 \alpha^2} \tag{4}$$

- In practice, $v_n, c_n, \hat{\sigma}_{n,v}, \hat{\sigma}_{n,c}, \hat{f}(\hat{v})$ should be used in (3) and (4).

Algorithm

Sorted Monte Carlo with reported RE:

- I. Initialize: Set iteration $i = 0$, needed sample size N_0 (e.g. $N_0 = 10000$), $RE_0 = \alpha_0$ ($\alpha_0 \gg \alpha$), sample set $\mathcal{S}_0 = \emptyset$, and total sample size $M_0 = \sharp(\mathcal{S}_0)$.
- II. S-step:
  - Sub-step 1: Generate $\{L_1, \cdots, L_{N_i}\}$, $\mathcal{S}_{i+1} = \mathcal{S}_i \bigcup \{L_1, \cdots, L_{N_i}\}$, $M_{i+1} = \sharp(\mathcal{S}_{i+1}) = M_i + N_i$.
  - Sub-step 2: Sort $\mathcal{S}_{i+1}$ in ascending order, $\hat{\theta}_{i+1} = v_{M_{i+1}}$ for VaR and $\hat{\theta}_{i+1} = c_{M_{i+1}}$ for ES.
  - Sub-step 3: Report $RE_{i+1} = RE(v_{M_{i+1}})$ for VaR and $RE_{i+1} = RE(c_{M_{i+1}})$ for ES by equations (3).
  - Sub-step 4: If $RE_{i+1} < \alpha$, stop with output $\hat{\theta}_{i+1}$ and RE_{i+1}. Otherwise go to the R-step.
- III. R-step: Compute $N_{i+1} = n_\alpha^v$ for VaR and $N_{i+1} = n_\alpha^c$ for ES by equations (4), and return to S-step by replacing i with $i + 1$.

Selection of sorting method

- **Classic quicksort**; dual-pivot quicksort (Yaroslavskiy, 2009); multi-pivot quicksort (Kushagra et al., 2014); .
- The complexity for Quick sorted MC, Bubble sorted MC, and Naive MC are $O(N \log N)$, $O(N^2)$, $O(N^2)$ respectively.

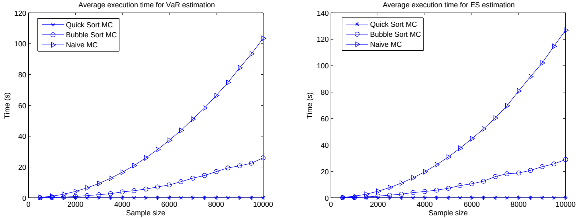


Figure : Left one is for VaR estimation, right one is for ES estimation. Samples are drawing from loss with density function $f(x) = e^x 1_{\{x<0\}}$. p is chosen as 0.05. Average execution time is estimated with 100 repeats.

Easy programming

- Easy programming due to:
  - Samples drawing from original loss model.
  - Based on order statistics, which enables us to use sorting method.
  - Analytic approximate formulae to compute RE and needed sample size.
- Other methods may lack these merits:
  - Estimators are usually more complicated, i.e, importance sampling, nonparametric method.
  - Some may not be easily applicable to dependent samples, i.e. importance sampling.

Fast computing

- **Execution Time:** $\mathbf{T} = \mathbf{T_{gen}} + \mathbf{T_{est}} + \mathbf{T_{eva}}$.
  - T_{est} is improved significantly with sorting method.
    —The computation complexity when choose quick sort method is $O(N \log N)$.
  - T_{eva} is improved by the expansion for RE's.

Time series factor model (10000 stocks; Type A)

Table : VaR and ES of Portfolio return $(1/n\sum_{i=1}^{n} r_{i,t})$ are computed. Portfolio includes $n = 10000$ equally weighted stocks. Each stock return $r_{i,t}, i = 1, \cdots, n$ (normalized by multiply 100) can be explained by three factors $f_{1,t}, f_{2,t}, f_{3,t}$, i.e. $r_{i,t} = \alpha_i + \sum_{k=1}^{3} \beta_k^{(i)} f_{k,t} + \epsilon_{i,t}$, $\epsilon \sim N(0, 1/6)$. Three factors satisfy the AR(2), ARCH(1), and SV (Stochastic Volatility) Model that used in Chen and Tang (2005)respectively. For Panel A, $\alpha_i = -3 + 6i/10000$, $\beta_1^{(i)} = -4 + 8i/10000$, $\beta_2^{(i)} = 8i/10000$, and $\beta_3^{(i)} = -1 + 4i/10000$. ES is only computed with our framework since ES is not considered in Heidelberger and Lewis (1984). Time is in seconds.

Method	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T (= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
Panel A: Portfolio with 10000 equally weighted stocks							
VaR							
$p = 0.01$							
This paper	-23.599(0.1841)	263467	41.392	1.1277	**42.519**	0.0074(0.0013)	0.0078
Heidelberger & Lewis(1984)	-23.409(0.2271)	1750	16.221	1605.9	**1622.1**	N.A.	0.0097
$p = 0.001$							
This paper	-42.053(0.3448)	1435849	260.57	7.0381	**267.61**	0.0080(0.0016)	0.0082
Heidelberger & Lewis(1984)	-42.178(0.4471)	1600	168.79	16711	**16879**	N.A.	0.0106
ES							
$p = 0.01$							
This paper	-30.986(0.2262)	722196	136.23	3.0413	**139.28**	0.0071(0.0020)	0.0073
Heidelberger & Lewis(1984)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
$p = 0.001$							
This paper	-54.671(0.4538)	5603487	1112.7	22.618	**1135.3**	0.0067(0.0021)	0.0083
Heidelberger & Lewis(1984)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

Time series factor model (20000 stocks; Type A)

Table : VaR and ES of Portfolio return $(1/n\sum_{i=1}^{n} r_{i,t})$ are computed. Portfolio includes $n = 20000$ equally weighted stocks. Each stock return $r_{i,t}, i = 1, \cdots, n$ (normalized by multiply 100) can be explained by three factors $f_{1,t}, f_{2,t}, f_{3,t}$, i.e. $r_{i,t} = \alpha_i + \sum_{k=1}^{3} \beta_k^{(i)} f_{k,t} + \epsilon_{i,t}$, $\epsilon \sim N(0, 1/6)$. Three factors satisfy the AR(2), ARCH(1), and SV (Stochastic Volatility) Model that used in Chen and Tang (2005)respectively. For Panel B, $\alpha_i = -8 + 6i/20000$, $\beta_1^{(i)} = -4 + 8i/20000$, $\beta_2^{(i)} = 8i/20000$, and $\beta_3^{(i)} = -1 + 4i/20000$. ES is only computed with our framework since ES is not considered in Heidelberger and Lewis (1984).

Method	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T(= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
Panel B: Portfolio with 20000 equally weighted stocks							
				VaR			
				$p = 0.01$			
This paper	-28.235(0.1836)	285385	78.295	1.2862	**79.581**	0.0074(0.0011)	0.0079
Heidelberger & Lewis(1984)	-28.336(0.2427)	1800	30.137	2983.6	**3013.7**	N.A.	0.0104
				$p = 0.001$			
This paper	-47.344(0.4383)	1524346	482.36	8.2804	**490.64**	0.0073(0.0014)	0.0078
Heidelberger & Lewis(1984)	-47.020(0.4328)	1700	326.58	32331	**32658**	N.A.	0.0103
				ES			
				$p = 0.01$			
This paper	-36.034(0.2099)	867756	248.09	3.7263	**251.81**	0.0063(0.0019)	0.0067
Heidelberger & Lewis(1984)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
				$p = 0.001$			
This paper	-59.597(0.3822)	6282067	2111.3	25.022	**2136.3**	0.0067(0.0018)	0.0070
Heidelberger & Lewis(1984)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

Portfolio of Call and Put Options with Time Series Returns (Type B)

Table : VaR and ES of Portfolio value $(V(t) - V(0))$ are computed. Portfolio includes shorting 10 ATM calls and 5 ATM puts on each of 10 uncorrelated stocks, while all options having a half-year maturity. We investigate losses over 10 days ($t = 10$ days). All stock have an initial value of 100, and they satisfy Johnson NGARCH(1,1) process (model is estimated based on daily data, see Simonato and Sentoft, 2015): $\ln \frac{S_\tau}{S_{\tau-1}} = \alpha + \sigma_\tau \epsilon_\tau$,
$\sigma_\tau^2 = \beta_0 + \beta_1 \sigma_{\tau-1}^2 + \beta_2 \sigma_{\tau-1}^2 (\epsilon - \theta)^2$, $\epsilon_\tau \sim J_{su}(a, b)$, where
$\alpha = 3.3 \times 10^{-4}, \beta_0 = 1.1 \times 10^{-6}, \beta_1 = 0.8664, \beta_2 = 0.0631, \theta = 0.9937, a = 0.3478, b = 2.1610.$

Method	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T(= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
				VaR			
				$p = 0.01$			
This paper	-81.479(0.6599)	27643	5.5602	0.0094	**5.5696**	0.0082(0.0004)	0.0081
Heidelberger & Lewis(1984)	-81.731(0.8664)	400	5.3006	524.76	**530.06**	N.A.	0.0106
				$p = 0.001$			
This paper	-103.83(0.8668)	56479	13.627	0.0242	**13.651**	0.0084(0.0010)	0.0083
Heidelberger & Lewis(1984)	-104.19(1.0836)	90	12.177	1205.5	**1217.7**	N.A.	0.0104
				ES			
				$p = 0.01$			
This paper	-90.946(0.7922)	30080	6.4069	0.0007	**6.4076**	0.0083(0.0005)	0.0087
Heidelberger & Lewis(1984)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
				$p = 0.001$			
This paper	-110.75(0.9414)	95170	21.254	0.0061	**21.261**	0.0080(0.0010)	0.0085
Heidelberger & Lewis(1984)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

Portfolio of Call and Put Options with I.I.D. Gaussian Returns (Type B)

Table : VaR and ES of Portfolio value $(V(t) - V(0))$ are computed. Portfolio (Glasserman et al., 2000): shorting 10 ATM calls and 5 ATM puts (all half year maturity) on each of 10 uncorrelated stocks, where $\Delta S = S_t - S_0$ is normal, $S_0 = 100, t = 0.04$, $\sigma = 0.3$, and $r = 5\%$.

Method	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T(= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
				VaR			
				$p = 0.01$			
This paper	-160.33(1.3391)	23086	0.0731	0.0079	**0.0810**	0.0077(0.0003)	0.0083
Hong and Sun (2010)	-159.38(1.6575)	300	0.0058	0.5743	**0.5801**	N.A.	0.0104
Glasserman et al. (2000)IIS	-158.93(1.5592)	135	0.1119	11.077	**11.189**	N.A.	0.0098
Glasserman et al. (2000)IIS-Q	-159.17(1.6596)	40	0.0995	9.8495	**9.9490**	N.A.	0.0108
				$p = 0.001$			
This paper	-197.86(1.7687)	56316	0.1819	0.0211	**0.2030**	0.0083(0.0010)	0.0088
Hong and Sun (2010)	-198.05(1.9104)	160	0.0031	0.3021	**0.3052**	N.A.	0.0097
Glasserman et al. (2000)IIS	-197.39(1.9937)	38	0.0439	4.3477	**4.3916**	N.A.	0.0101
Glasserman et al. (2000)IIS-Q	-196.96(2.0681)	12	0.0532	5.2666	**5.3198**	N.A.	0.0105
				ES			
				$p = 0.01$			
This paper	-174.07(1.4100)	26551	0.0862	0.0006	**0.0868**	0.0078(0.0004)	0.0081
Hong and Sun (2010)	-175.19(1.7869)	400	0.0018	0.1801	**0.1819**	N.A.	0.0102
Glasserman et al. (2000)IIS	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Glasserman et al. (2000)IIS-Q	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
				$p = 0.001$			
This paper	-213.41(1.7723)	91474	0.2647	0.0083	**0.2730**	0.0081(0.0010)	0.0083
Hong and Sun (2010)	-212.52(2.1677)	50	0.0011	0.1067	**0.1078**	N.A.	0.0102
Glasserman et al. (2000)IIS	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Glasserman et al. (2000)IIS-Q	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

Table : VaR and ES of Portfolio value $(V(t) - V(0))$ are computed. Portfolio (Fuh et al., 2011) consists of options with heavy-tailed underlying stock returns. Parameters satisfy $\nu = 5, b_j = 0.1 + j/100, \lambda_j = 0.05 \times j, j = 1, \cdots, 15$

Method	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T(= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
				VaR			
				$p = 0.01$			
This paper	-59.089(0.4964)	312680	0.9519	0.2176	**1.1695**	0.0080(0.0010)	0.0084
Glasserman et al. (2002)	-59.319(0.5932)	1000	1.3922	137.83	**139.22**	N.A.	0.0100
Fuh et al. (2011)	-59.191(0.6452)	850	13.481	1334.6	**1348.1**	N.A.	0.0109
				$p = 0.001$			
This paper	-159.10(1.3046)	2266233	7.5765	1.1339	**8.7104**	0.0082(0.0011)	0.0084
Glasserman et al. (2002)	-158.73(1.5931)	930	1.2052	119.31	**120.52**	N.A.	0.0101
Fuh et al. (2011)	-159.43(1.6561)	800	9.7891	969.12	**978.91**	N.A.	0.0106
				$p = 0.0005$			
This paper	-212.86(1.7029)	6251108	20.720	3.3899	**24.110**	0.0072(0.0016)	0.0080
Glasserman et al. (2002)	-211.54(2.2423)	850	1.1334	112.21	**113.34**	N.A.	0.0100
Fuh et al. (2011)	-212.13(1.9516)	780	9.1624	907.08	**916.24**	N.A.	0.0092
				ES			
				$p = 0.01$			
This paper	-103.35(0.8061)	1695129	4.9997	0.0718	**5.0715**	0.0078(0.0011)	0.0078
Glasserman et al. (2002)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Fu et al. (2011)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
				$p = 0.001$			
This paper	-268.68(2.2569)	12522977	34.615	0.3624	**34.977**	0.0082(0.0011)	0.0084
Glasserman et al. (2002)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Fuh et al. (2011)	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.

More Application: Intra-horizon risk

- Bakshi and Panayotov (2010, JFE): intra-horizon risk (VaR-I).
- Let $X_t, t \in [0, T]$ with $X_0 = 0$ be a real-valued random process. Denote $X_T^{min} := \min_{0<t<T} X_t$. VaR-I is defined as the value of a quantile of the random variable X_T^{min}.
- VaR-I is related with first passage time, and hence can be computed by solving some PIDE as Bakshi and Panayotov (2010) suggested.
- Our algorithm provides an alternative way to compute VaR-I and VaR without solving PIDE.

Table : VaR-I and VaR multiples over benchmark VaR for a two-week horizon are showed. Reported are values of te multiples VaR-I$/(2.32\hat{\sigma} - \hat{\mu})$ and VaR$/(2.32\hat{\sigma} - \hat{\mu})$, where Benchmark VaR is the quantile of Normal distribution $N(\hat{\mu}, \hat{\sigma})$, where $\hat{\mu} = 0$ and $\hat{\sigma}$ is the standard deviation of the returns time series. Log stock price is modeled by three processes: Merton's jump-diffusion (JD) process, CGMY process, and Stochastic Volatility (SV) process. JD: $dX_t = \mu dt + \sigma dW_t + dJ_t$ with Levy measure $k[x] = \frac{\lambda}{\sigma_J \sqrt{2\pi}} exp(-\frac{(x-\mu_J)^2}{2\sigma_J^2})$, where $\mu = -(\sigma^2/2) - \lambda(\exp(\mu_J + \sigma_J^2/2) - 1)$, $\lambda = 17.8922, \mu_J = -0.0073, \sigma_J = 0.0306, \sigma = 0.1139$ (parameters are calibrated from 1995-2015 weekly S& P 500 index return, also see Bakshi and Panayotov (2010) for the calibration method).

Method	Risk Measure	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T(= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
	Panel A: Merton's jump-diffusion model							
					$p = 0.01$			
This paper	VaR-I	1.3319(0.0115)	26822	76.664	0.0173	**76.681**	0.0085(0.0009)	0.0086
	VaR	1.1958(0.0104)	37171	86.746	0.0271	**86.773**	0.0086(0.0011)	0.0087
Heidelberger & Lewis (1984)	VaR-I	1.3287(0.0154)	350	62.851	6222.2	**6285.1**	N.A.	0.0116
	VaR	1.2026(0.0135)	400	78.878	7808.9	**7887.8**	N.A.	0.0112
					$p = 0.001$			
This paper	VaR-I	1.4152(0.0110)	123011	325.79	0.0852	**325.88**	0.0083(0.0009)	0.0078
	VaR	1.3364(0.0112)	149501	364.95	0.1078	**365.06**	0.0080(0.0008)	0.0084
Heidelberger & Lewis (1984)	VaR-I	1.4213(0.0163)	115	198.46	19648	**19846**	N.A.	0.0115
	VaR	1.3321(0.0143)	130	206.85	20478	**20685**	N.A.	0.0108

Intra-horizon risk: CGMY model

Table : VaR-I and VaR multiples over benchmark VaR for a two-week horizon are showed. Reported are values of te multiples VaR-I$/(2.32\hat{\sigma} - \hat{\mu})$ and VaR$/(2.32\hat{\sigma} - \hat{\mu})$, where Benchmark VaR is the quantile of Normal distribution $N(\hat{\mu}, \hat{\sigma})$, where $\hat{\mu} = 0$ and $\hat{\sigma}$ is the standard deviation of the returns time series. Log stock price is modeled by three processes: Merton's jump-diffusion (JD) process, CGMY process, and Stochastic Volatility (SV) process. CGMY: a pure-jump Levy process with Levy measure $k[x] = \lambda \frac{exp(-\beta_-|x|)}{|x|^{1+\alpha}} 1_{\{x<0\}} + \lambda \frac{exp(-\beta_+ x)}{x^{1+\alpha}} 1_{\{x>0\}}$, where $\lambda = 5.8656, \beta_- = 41.3185, \beta_+ = 60.7789, \alpha = 0.50$ (parameters are calibrated from 1995-2015 weekly S& P 500 index return, also see Bakshi and Panayotov (2010) for the calibration method).

Method	Risk Measure	Estimate (SD)	Sample size	$T_{gen} + T_{est}$	T_{eva}	$T(= T_{gen} + T_{est} + T_{eva})$	Estimated RE (SD)	Resampled RE
	Panel B: Exponentially dampened power law model of CGMY							
					$p = 0.01$			
This paper	VaR-I	1.2166(0.0111)	8523	255.68	0.0049	**255.69**	0.0081(0.0012)	0.0091
	VaR	1.1227(0.0098)	6523	197.58	0.0037	**197.59**	0.0091(0.0006)	0.0087
Heidelberger & Lewis (1984)	VaR-I	1.2113(0.0138)	85	142.68	14125	**14268**	N.A.	0.0114
	VaR	1.1246(0.0136)	75	123.74	12250	**12372**	N.A.	0.0121
					$p = 0.001$			
This paper	VaR-I	1.3899(0.0131)	36793	1155.4	0.0262	**1155.4**	0.0088(0.0014)	0.0094
	VaR	1.3246(0.0122)	25841	687.74	0.0156	**687.76**	0.0087(0.0013)	0.0092
Heidelberger & Lewis (1984)	VaR-I	1.3901(0.0149)	45	830.11	82181	**83011**	N.A.	0.0107
	VaR	1.3197(0.0149)	30	560.70	55509	**56070**	N.A.	0.0113

Outline

Dependence setting

- $\alpha-\mathbf{mixing}$: For a sequence of r.vs $\{L_i\}_{i=1}^n$, let $\mathscr{F}_k^l$ be the $\sigma-$algebra of events generated by $\{L_i, k \leq i \leq l\}$ for $l \geq k$. The $\alpha-$mixing coefficient introduced by Rosenblatt (1956) is

$$\alpha(k) = \sup_{A \in \mathscr{F}_1^i, B \in \mathscr{F}_{i+k}^n} |P(AB) - P(A)P(B)|.$$

  - The series is said to be $\alpha - mixing$ if $\lim_{k\to\infty} \alpha(k) = 0$.
  - The series is said to be **geometric** $\alpha - \mathbf{mixing}$ if $\alpha(k) \leq c\rho^k$ for some constants $c > 0$ and $\rho \in (0,1)$.
- Examples of stationary **geometric** $\alpha - \mathbf{mixing}$ (Chen and Tang, 2005): ARMA(p,q), GARCH(p,q), Diffusion Model (Vasicek Model), Stochastic Volatility Model

A heuristic method in the existing literature

- (Hult and Svensson, 2009) For i.i.d samples, assuming central limit theorem $\sqrt{n}(\hat{\theta} - \theta) \Rightarrow \alpha N(0,1)$, then $SD(\hat{\theta})$ and $E(\hat{\theta})$ can be approximated by

$$SD(\hat{\theta}) \approx \alpha/\sqrt{n} + o_p(1/\sqrt{n}); \quad E(\hat{\theta}) \approx \theta + o_p(1).$$

  Thus

$$RE(\hat{\theta}) = \frac{SD(\hat{\theta})}{|E(\hat{\theta})|} \approx \frac{\alpha/\sqrt{n} + o_p(1/\sqrt{n})}{|\theta| + o_p(1)} \approx \frac{\alpha}{\sqrt{n}|\theta|}.$$

- Extend to dependent samples
  - Asymptotic normality of VaR estimator (Yoshihara, 1995) and ES estimator (Chen, 2008)

$$\sqrt{n} f(v) \sigma_{n,v}^{-1}(v_n - v) \Rightarrow \mathcal{N}(0,1); \quad \sqrt{n} p \sigma_{n,c}^{-1}(c_n - c) \Rightarrow \mathcal{N}(0,1)$$

  - Similarly, we obtain

$$RE(\hat{v}) \approx -\frac{\sigma_{n,v}}{v f(v)} n^{-1/2}, RE(\hat{c}) \approx -\frac{\sigma_{n,c}}{cp} n^{-1/2}$$

An counter example

Let $Z_n = z + T_n + \frac{1}{\sqrt{n}} N(0,1)$, where z is a positive constant and $P(T_n = n) = 1/n, P(T_n = 0) = 1 - 1/n$. Suppose T_n and $N(0,1)$ are independent. Notice that $\sqrt{n}T_n \to 0$ in probability, then $\sqrt{n}(Z_n - z) \Rightarrow N(0,1)$.

- Heuristic method implies:

$$RE(Z_n) = \frac{SD(Z_n)}{|E(Z_n)|} \approx \frac{1/\sqrt{n} + o_p(1/\sqrt{n})}{z + o_p(1)} \approx \frac{1}{z\sqrt{n}}.$$

- However, since $SD(Z_n) = \sqrt{n + 1/n}, EZ_n = z + 1$, we have

$$RE(Z_n) = \frac{SD(Z_n)}{EZ_n} = \frac{\sqrt{(n + 1/n)}}{(z+1)} \to \infty$$

Main technical difficulties

- (Hong and Sun, 2010) develops asymptotic results for v_n and c_n in terms of o_p and $O_{a.s.}$:

$$v_n = v + \frac{1}{f(v)}(p - \frac{1}{n}\sum_{i=1}^{n} 1_{\{L_i \le v\}}) + A_n$$

$$c_n = c + (\frac{1}{n}\sum_{i=1}^{n}[v - \frac{1}{p}(v - L_i)^+] - c) + B_n$$

  where $A_n = o_p(n^{-1/2}), B_n = o_p(n^{-1/2})$, or
  $A_n = O_{a.s.}(n^{3/4}(\log n)^{3/4}), B_n = O_{a.s.}(n^{-1}\log n)$
- However: $X_n = o_p(g(n)) \nRightarrow EX_n^2 = o(g(n)^2)$ or even $EX_n^2 = O(1)$, $X_n = O_{a.s.}(g(n)) \nRightarrow EX_n^2 = O(g(n)^2)$ or even $EX_n^2 = O(1)$.
- Main difficulties: evaluating the moments of errors: $E(v_n - v)^m; E(c_n - c)^m$.

Solution: Careful Analysis of Moments of Errors

- Using Berry-Essen bound (Tikhomirov, 1980) and Bernstein inequality (Merlevede, Peligrad, and Rio, 2009) under dependent setting, and by improving techniques of Liu and Yang (2012, Adv. Appl. Prob.), we obtain

$$E(v_n - v)^m = \frac{\sigma_{n,v}^m}{f(v)^m} E Z^m n^{-m/2} + o(n^{-m/2-1/4}), \quad (5)$$

$$E(c_n - c)^m = \frac{\sigma_{n,c}^m}{p^m} E Z^m n^{-m/2} + o(n^{-m/2-1/4+\epsilon}). \quad (6)$$

- Hall and Martin (1988) provides similar results of (5) for i.i.d samples.
- Chen (2008) heuristically develops a bound for $E(c_n - c)^2$ under dependent setting.

Main Technical Results: Approximation of REs

Theorem 1 (Dependent samples)

For samples that satisfy geometric $\alpha-$mixing and strict stationarity, under some regular conditions, as $n \to \infty$,

$$RE(v_n) = -\frac{\sigma_{n,v}}{vf(v)} n^{-1/2} + o(n^{-3/4}) \quad (7)$$

$$RE(c_n) = -\frac{\sigma_{n,c}}{cp} n^{-1/2} + o(n^{-3/4+\epsilon}) \quad (8)$$

where $\sigma^2_{n,v} = \{p(1-p) + 2\sum_{k=1}^{(n-1)} \gamma_1(k)\}$, $\gamma_1(k) = cov\{1_{\{L_1<v\}}, 1_{\{L_{k+1}<v\}}\}$,
and $\sigma^2_{n,c} = \{Var[(v-L_1)^+] + 2\sum_{k=1}^{(n-1)} \gamma_2(k)\}$, $\gamma_2(k) = cov\{(v-L_1)^+, (v-L_{k+1})^+\}$.

In particular, when the samples are i.i.d, we have $\sigma^2_{n,v} = \{p(1-p)\}$, and $\sigma^2_{n,c} = \{Var[(v-L_1)^+]\}$.

Implication: Estimators for RE

- RE estimators:

$$\hat{RE}(v_n) = -\frac{\hat{\sigma}_{n,v}}{v_n \hat{f}(v_n)} n^{-1/2}, \quad \hat{RE}(c_n) = -\frac{\hat{\sigma}_{n,c}}{c_n p} n^{-1/2}. \tag{9}$$

  where v_n and c_n are from estimates in S-step, $\hat{\sigma}_{n,v}, \hat{\sigma}_{n,c}$ are estimators from spectral method (Heidelberger et al., 1981) and $\hat{f}(\cdot)$ is standard kernel estimator of density function.

- Resampled RE: repeat the estimation m times and obtain $\hat{\theta}^{(1)}, \cdots, \hat{\theta}^{(m)}$, $\bar{\hat{\theta}} = \sum_{i=1}^{m} \hat{\theta}^{(i)}/m$, the resampled RE is computed by

$$\tilde{RE}(\hat{\theta}) = \frac{\sqrt{\sum_{i=1}^{m} (\hat{\theta}^{(i)} - \bar{\hat{\theta}})^2/(m-1)}}{\bar{\hat{\theta}}}. \tag{10}$$

Numerical results of Theorem 1

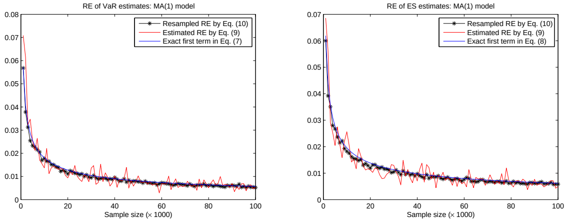


Figure : Estimated REs and resampled REs for VaR and ES. Loss samples satisfy the MA(1) Model: $L_{t+1} = 0.5\epsilon_t + \epsilon_{t+1}$, $\epsilon_t \sim N(0,1)$. $p = 0.05$ and simulated REs are obtained with 200 repeats.

Numerical results of Theorem 1

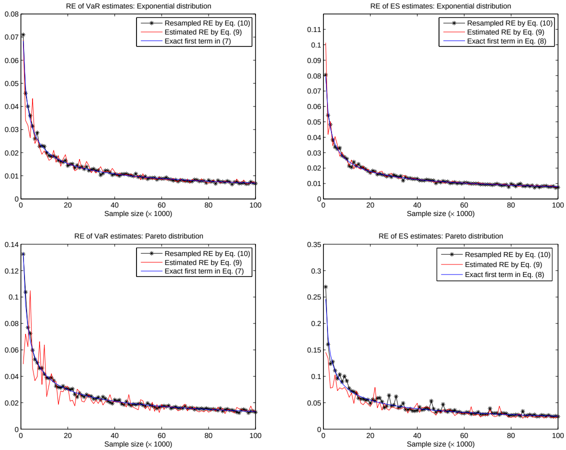


Figure : Estimated REs and resampled REs for VaR and ES. Loss distributions here are i.i.d with density function $f(x) = e^x 1_{\{x<0\}}$, and $f(x) = (1 - x/3)^{-4} 1_{\{x<0\}}$ respectively. $p = 0.05$ and simulated REs are obtained with 200 repeats.

Outline

Comparison of VaR and ES

Theorem 2

(I)(Dependent samples) Under some regular conditions,

$$\lim_{n\to\infty} \frac{RE(v_n(p))}{RE(c_n(p))} = \frac{cp\sigma_{\infty,v}}{vf(v)\sigma_{\infty,c}} < \infty.$$

where $\sigma^2_{\infty,v} = \{p(1-p) + 2\sum_{k=1}^{\infty} \gamma_1(k)\}, \sigma^2_{\infty,c} = \{Var[(v-L_1)^+] + 2\sum_{k=1}^{\infty} \gamma_2(k)\}$.

.

(II)(IID samples) For i.i.d samples under Assumptions A1, B1, C, further assuming the existence of $\lim_{x\to-\infty} xh^{'}(x)$ and $\lim_{x\to-\infty} h(x)$, where $h(x) = \frac{F(x)}{xf(x)}$, then

$$\lim_{p\to 0} \lim_{n\to\infty} \frac{RE(v_n(p))}{RE(c_n(p))} \le \frac{1}{\sqrt{2}}.$$

Remark: 1) Both $\sum_{k=1}^{\infty} \gamma_1(k)$ and $\sum_{k=1}^{\infty} \gamma_2(k)$ converge under geometric $\alpha - mixing$.
2) Most distributions with smooth tail densities satisfy our regular conditions, including Normal, Lognormal, Exponential, Weibull, Pareto distributions.

Comparison of Media Shortfall (MS) and ES

Theorem 2

(I)(Dependent samples) Under some regular conditions,

$$\lim_{n\to\infty} \frac{RE(m_n(p))}{RE(c_n(p))} = \frac{c(p)p\sigma_{\infty,m}}{m(p)f(m(p))\sigma_{\infty,m}} < \infty$$

where $\sigma^2_{\infty,m} = \{1/2p(1-p/2) + 2\sum_{k=1}^{\infty} \gamma_1(k, p/2)\}$,
$\sigma^2_{\infty,m} = \{Var[(v(p/2) - L_1)^+] + 2\sum_{k=1}^{\infty} \gamma_2(k, p/2)\}$.
.
(II)(IID samples) Under the conditions of Assumptions A1, B1, C, denote $\frac{F(x)}{xf(x)} = h(x)$, and assume the existence of $\lim_{x\to-\infty} xh^{'}(x)$ and $\lim_{x\to-\infty} h(x)$. Moreover, we assume $\exists\theta \leq 1$ such that $\frac{v(p)f(v(p))}{p^\theta}$ is a slowly varying function at $p = 0$. Then

$$\lim_{p\to 0}\lim_{n\to\infty} \frac{RE(m_n(p))}{RE(c_n(p))} \leq 1.$$

Remark: According to Kou and Peng (2014), MS at level p (i.e. $m(p)$) is exactly the VaR at level $\frac{p}{2}$ (i.e. $v(\frac{p}{2})$), i.e. $m(p) = v(\frac{p}{2})$.

Numerical results: dependent samples and i.i.d samples

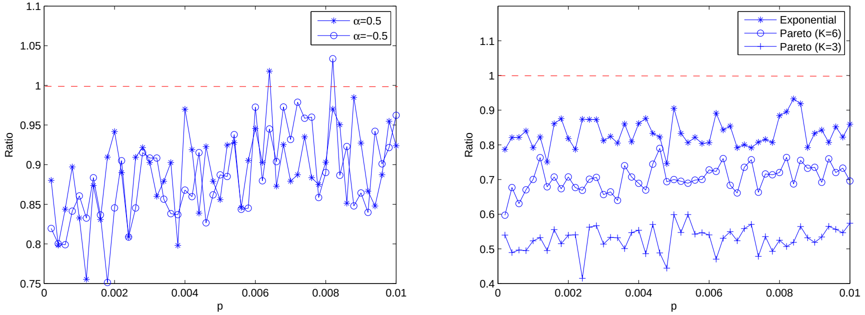


Figure : Ratio of REs for VaR and ES as p varies. Loss samples in left figure satisfy MA(1) Model: $L_{t+1} = \alpha\epsilon_t + \epsilon_{t+1}$, $\epsilon_t \sim N(0,1)$, $\alpha = 0.5$ or -0.5. Loss p.d.f in right satisfies $f(x) = e^x 1_{\{x<0\}}$ and $f(x) = (1 - \frac{x}{K})^{-K-1} 1_{\{x<0\}}$ respectively. Sample size is chosen to be 10^6, and p varies from 0.0002 to 0.01. The estimation of RE is based on 200 times.

Summary

- A sorted Monte Carlo with reported RE is introduced to simulate risk measures for dependent samples. Numerical experiments indicate this method is easy to implement and fast, compared to existing methods, in terms of total execution time, even at the level 0.001.
- Rigorous expansions of RE's of risk measures are given, which enable us to compute needed sample size, and to make comparison between different risk measures.

Thanks!

Appendix A: Regularity Conditions

Assumption A: Sample process $L_1, \cdots, L_n$ is strictly stationary and geometric $\alpha-$mixing. Each L_i is distributed with f and F as its density function and distribution function, respectively.

Assumption A1: Samples $L_1, \cdots, L_n$ are i.i.d. Each L_i is distributed with f and F as its density function and distribution function, respectively

Assumption B: F has bounded second derivative in a neighborhood of v, and $vf(v) < 0$.

Assumption B1: F has bounded second derivative in a neighborhood of v, and $xf(x) < 0$ is satisfied for all $x \leq v$.

Assumption C: L has finite moment for all orders, that is $E|L|^{2+\delta} < \infty$ for all $\delta > 0$.

Assumption C1: L has finite first moment, that is $E|L| < \infty$.

Appendix B

- **Example:** Let random variables X_n defined as:

$$X_n(\omega) = \frac{1}{n} 1_{\{\frac{1}{n^2} \le \omega \le 1\}} + (\frac{1}{n} + n^2) 1_{\{0 \le \omega < \frac{1}{n^2}\}}.$$

  By definition $P(|n^{1-\epsilon} X_n| > \theta) = 1/n^2 \to 0$, hence $X_n = o_p(n^{-1+\epsilon})$.
- However, $EX_n^2 = O(n^2) \neq o(n^{-2+2\epsilon})$.
  Furthermore, $EX_n^2 = O(n^2) \to \infty$ as $n \to \infty$.

Appendix B

- **Example:** Define random variables as $Y_n = 2^n 1_{U_n} + n^{-1} 1_{V_n}$, where U_n, V_n are defined recursively: $U_1 = (0, 1/3) \cup (2/3, 1), V_1 = [1/3, 2/3]$, if $U_k = \cup_{i=1}^{2^k}(a_i, b_i), V_k = \cup_{i=1}^{2^k-1}[c_i, d_i]$, then $U_{k+1} = \cup_{i=1}^{2^k}(a_i, a_i + (b_i - a_i)/3) \cup (a_i + 2(b_i - a_i)/3, b_i), V_{k+1} = V_k \cup_{i=1}^{2^k} [a_i + (b_i - a_i)/3, a_i + 2(b_i - a_i)/3]$.
- Since $\sum_{k=1}^{\infty} P(|Y_n| > 1/n) = \sum_{k=1}^{\infty}(2/3)^k < \infty$, by Borel-Cantelli Lemma, we have $\mathbf{Y_n = O_{a.s.}(1/n)}$.
- However, since $\lim_{n\to\infty} n^2 EY_n^2 = \lim_{n\to\infty} n^2\{(2/3)^n 2^{2n} + (1 - (2/3)^n)1/n^2\} = \infty$, then $\mathbf{EY_n^2 \neq O(1/n^2)}$.
  Furthermore, we have $\mathbf{EY_n^2 = O((8/3)^n)} \to \infty$.