Adaptive stratification

Gersende FORT

CNRS - TELECOM ParisTech
Paris, France

Joint work with B. JOURDAIN (ENPC, France), E. MOULINES (TELECOM ParisTech, France) and P. ETORE (Univ. J. Fourier, France)

Stratification

Let $\phi : \mathbb{R}^d \to \mathbb{R}$ and a random variable $Y \in \mathbb{R}^d$

Given

- a direction of stratification $\mu \in \mathbb{R}^d$
- a partition of $\mathbb{R}$ in I strata A_i

write

$$\mathbb{E}\left[\phi(Y)\right] = \sum_{i=1}^{I} \mathbb{P}(\mu^t Y \in \mathsf{A}_i)\ \mathbb{E}\left[\phi(Y)|\ \mu^t Y \in \mathsf{A}_i\right]$$

Stratification

Let $\phi : \mathbb{R}^d \to \mathbb{R}$ and a random variable $Y \in \mathbb{R}^d$

Given

- a direction of stratification $\mu \in \mathbb{R}^d$
- a partition of $\mathbb{R}$ in I strata A_i

write

$$\mathbb{E}\left[\phi(Y)\right] = \sum_{i=1}^{I} \mathbb{P}(\mu^t Y \in \mathsf{A}_i)\ \mathbb{E}\left[\phi(Y) \middle|\ \mu^t Y \in \mathsf{A}_i\right]$$

and by a Monte Carlo method, approximate

$$\mathbb{E}\left[\phi(Y)\right] \approx \sum_{i=1}^{I} \mathbb{P}(\mu^t Y \in \mathsf{A}_i)\ \frac{1}{n_i} \sum_{k=1}^{n_i} \phi(Y_k^{(i)})$$

where $Y_k^{(i)}, k \le n_i$ i.i.d. under $\mathbb{P}(Y \in \cdot |\ \mu^t Y \in \mathsf{A}_i)$.

Variance reduction

Given

- I strata A_i
- a direction of stratification μ
- an allocation policy : $q_1 + \cdots + q_I = 1$ and $n_i \sim_{n \to +\infty} n q_i$

the variance of the stratified estimator is given by assuming $n_i > 0$

$$n \; \sigma^2_{I,n}\left(\mu, \mathsf{A}_., q_.\right) = \sum_{i=1}^{I} \frac{p_i^2}{q_i} \; \sigma_i^2$$

with

$$p_i := \mathbb{P}(\mu^t Y \in \mathsf{A}_i) \qquad\qquad \sigma_i^2 := \mathbb{V}\left(\Phi(Y) \middle|\, \mu^t Y \in \mathsf{A}_i\right).$$

- Proportional allocation : $q_i = p_i = \mathbb{P}(\mu^t Y \in \mathsf{A}_i)$

$$\sigma^2_{I,n}(\mu,\mathsf{A}.,p.) \leq \sigma^2_{MC}$$

$\hookrightarrow$ stratification can improve on the usual MC

- Proportional allocation : $q_i = p_i = \mathbb{P}(\mu^t Y \in \mathsf{A}_i)$

$$\sigma^2_{I,n}(\mu, \mathsf{A}_., p_.) \leq \sigma^2_{MC}$$

$\hookrightarrow$ stratification can improve on the usual MC

- Optimal allocation : $q_i^\star \propto p_i \ \sigma_i$

$$\sigma^2_{I,n}(\mu, \mathsf{A}_., q_.^\star) \leq \sigma^2_{I,n}(\mu, \mathsf{A}_., q_.)$$

$\hookrightarrow$ σ_i unknown $\Longrightarrow$ it has to be estimated

"Design parameters" for the stratified estimator

Different parameters to be fixed

- the I strata A_i: equivalently, a distribution g and its (generalized) inverse-cdf G^{-1} s.t.

$$\mathsf{A}_i = \left[G^{-1}\left(\frac{i-1}{I}\right); G^{-1}\left(\frac{i}{I}\right) \right[$$

- the allocation policy q_i : equivalently, a distribution χ s.t.

$$q_i = \int_{\mathsf{A}_i} \chi(x) dx.$$

- the direction of stratification $\mu \in \mathbb{R}^d$ $\quad (\mu^t \mu = 1)$

$\hookrightarrow$ how to fix these quantities?

"Design parameters" for the stratified estimator

Different parameters to be fixed

- the I strata A_i: equivalently, a distribution g and its (generalized) inverse-cdf G^{-1} s.t.

$$\mathsf{A}_i = \left[G^{-1}\left(\frac{i-1}{I}\right); G^{-1}\left(\frac{i}{I}\right) \right[$$

- the allocation policy q_i : equivalently, a distribution χ s.t.

$$q_i = \int_{\mathsf{A}_i} \chi(x) dx.$$

- the direction of stratification $\mu \in \mathbb{R}^d$ $\quad (\mu^t \mu = 1)$

$\hookrightarrow$ how to fix these quantities? Answer: s.t. the asymptotic variance $(I, n \to +\infty)$ is minimal.

Results

$\psi_\mu(x) := \mathbb{E}\left[\phi(Y)|\mu^t Y = x\right]$ $\zeta_\mu(x) := \mathbb{E}\left[\phi^2(Y)|\mu^t Y = x\right]$ $f_\mu(x)$ density of $\mu^t Y$

Theorem

Assume

1. $\int \chi^2 g^{-1}\, dx < +\infty, \quad \mathrm{essinf}_g(\chi/g) > 0.$
2. $\int h^2 g^{-1} dx < +\infty$ *for* $h \in \{f_\mu, \zeta_\mu f_\mu, \psi_\mu f_\mu\}$.
3. $\int f_\mu^4 (\zeta_\mu - \psi_\mu^2)^2 / [\chi^2 g] dx < +\infty.$

Then, when $I, n \to +\infty$ *or* $I_n^{-1} + I_n^2/n \to 0$,

$$n\ \sigma^2_{I,n}(\mu, g, \chi) \longrightarrow \int \frac{f_\mu(x)}{\chi(x)} \{\zeta_\mu(x) - \psi_\mu^2(x)\}\ f_\mu(x)\ dx.$$

Results

$\psi_\mu(x) := \mathbb{E}\left[\phi(Y)|\mu^t Y = x\right]$ $\zeta_\mu(x) := \mathbb{E}\left[\phi^2(Y)|\mu^t Y = x\right]$ $f_\mu(x)$ density of $\mu^t Y$

Theorem

Assume

1. $\int \chi^2 g^{-1}\, dx < +\infty, \quad \operatorname{essinf}_g(\chi/g) > 0.$
2. $\int h^2 g^{-1} dx < +\infty$ for $h \in \{f_\mu, \zeta_\mu f_\mu, \psi_\mu f_\mu\}$.
3. $\int f_\mu^4 (\zeta_\mu - \psi_\mu^2)^2 / [\chi^2 g] dx < +\infty.$

Then, when $I, n \to +\infty$ *or* $I_n^{-1} + I_n^2/n \to 0$,

$$n\ \sigma^2_{I,n}(\mu, g, \chi) \longrightarrow \int \frac{f_\mu(x)}{\chi(x)} \{\zeta_\mu(x) - \psi_\mu^2(x)\}\ f_\mu(x)\ dx.$$

- it does not depend on g i.e. on the strata.
- it depends on the allocation policy χ and is minimal for

$$\chi_\mu^\star \propto f_\mu \sqrt{\zeta_\mu(x) - \psi_\mu^2(x)}.$$

$\hookrightarrow \chi_\mu^\star$ not available in practice.

The **minimal** asymptotic variance (when $\chi = \chi^\star_\mu$) is

$$\int \sqrt{\zeta_\mu(x) - \psi^2_\mu(x)}\ f_\mu(x)\ dx = \mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2.$$

Theorem

Assume 2. Then

$$\lim_{I \to +\infty} \sum_{i=1}^{I} \left| q^\star_i - \int_{\mathsf{A}_i} \chi^\star(x)\ dx \right| = 0,$$

and

$$\lim_n n\ \sigma^2_{I,n}\left(\mu, g, q^\star_\cdot\right) = \int \sqrt{\zeta_\mu(x) - \psi^2_\mu(x)}\ f_\mu(x)\ dx.$$

The **minimal** asymptotic variance (when $\chi = \chi^\star_\mu$) is

$$\int \sqrt{\zeta_\mu(x) - \psi^2_\mu(x)}\ f_\mu(x)\ dx = \mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2.$$

Theorem

Assume 2. Then

$$\lim_{I \to +\infty} \sum_{i=1}^{I} \left| q_i^\star - \int_{\mathsf{A}_i} \chi^\star(x)\ dx \right| = 0,$$

and

$$\lim_n n\ \sigma^2_{I,n}\left(\mu, g, q^\star_\cdot\right) = \int \sqrt{\zeta_\mu(x) - \psi^2_\mu(x)}\ f_\mu(x)\ dx.$$

↪ In order to "reach" the minimal variance, apply the stratified estimator with

- the optimal allocation $q^\star_\cdot$ ↪ estimate the quantity !
- th direction μ that minimizes $n\ \sigma^2_{I,n}\left(\mu, g, q^\star_\cdot\right)$; closed form : NOT available: ↪ estimate the quantity !

Comparison with Glasserman et al. (1999)

- We obtained a **minimal** asymptotic variance:

$$\mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2.$$

- In Glasserman et al. the **minimal** asymptotic variance **in the case of proportional allocation** is

$$\mathbb{E}\left[\mathbb{V}\left[\phi(Y)|\mu^t Y\right]\right]$$

Comparison with Glasserman et al. (1999)

- We obtained a **minimal** asymptotic variance:

$$\mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2.$$

- In Glasserman et al. the **minimal** asymptotic variance **in the case of proportional allocation** is

$$\mathbb{E}\left[\mathbb{V}\left[\phi(Y)|\mu^t Y\right]\right]$$

We have

$$\mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2 \leq \mathbb{E}\left[\mathbb{V}\left[\phi(Y)|\mu^t Y\right]\right],$$

but this is due to the allocation : optimal (left) / proportional (right).

Comparison with Glasserman et al. (1999)

- We obtained a **minimal** asymptotic variance:

$$\mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2.$$

- In Glasserman et al. the **minimal** asymptotic variance **in the case of proportional allocation** is

$$\mathbb{E}\left[\mathbb{V}\left[\phi(Y)|\mu^t Y\right]\right]$$

We have

$$\mathbb{E}\left[\sqrt{\mathbb{V}\left[\phi(Y)|\mu^t Y\right]}\right]^2 \leq \mathbb{E}\left[\mathbb{V}\left[\phi(Y)|\mu^t Y\right]\right],$$

but this is due to the allocation : optimal (left) / proportional (right).

$\hookrightarrow$ in both papers, use this expression to derive the : (optimal) direction of stratification μ

Algorithm : adaptive stratification

1. Choose the optimal allocation.
2. Choose μ as the direction that minimizes

$$V(\mu) := n\ \sigma^2_{I,n}(\mu, \mathsf{A}_\cdot, q^\star_\cdot) = \left(\sum_{i=1}^{I} p_i\ \sigma_i\right)^2 .$$

Algorithm : adaptive stratification

1. Choose the optimal allocation.
2. Choose μ as the direction that minimizes

$$V(\mu) := n\ \sigma^2_{I,n}(\mu,\mathsf{A}_\cdot,q^\star_\cdot) = \left(\sum_{i=1}^{I} p_i\ \sigma_i\right)^2 .$$

We have

- $n\ \sigma^2_{I,n}(\mu,\mathsf{A}_\cdot,q^\star_\cdot) = \Xi\left(\nu_i(h,\mu)\right)$ for $h \in \{f,f\phi,f\phi^2\}$

$$\nu_i(h,\mu) := \int_{\{y,\mu^t y \in \mathsf{A}_i\}} h(x)dx.$$

Algorithm : adaptive stratification

1. Choose the optimal allocation.
2. Choose μ as the direction that minimizes

$$V(\mu) := n\ \sigma^2_{I,n}(\mu,\mathsf{A}_\cdot,q^\star_\cdot) = \left(\sum_{i=1}^{I} p_i\ \sigma_i\right)^2 .$$

We have

- $n\ \sigma^2_{I,n}(\mu,\mathsf{A}_\cdot,q^\star_\cdot) = \Xi\left(\nu_i(h,\mu)\right)$ for $h \in \{f, f\phi, f\phi^2\}$

$$\nu_i(h,\mu) := \int_{\{y,\mu^t y \in \mathsf{A}_i\}} h(x)dx.$$

- The gradient $\nabla V(\mu)$ is a function of $\nu_i(h,\mu)$ and $\nabla_\mu \nu_i(h,\mu)$:

$$\nabla_\mu \left[\int_{\{y,\mu^t y \leq z\}} h(y)dy\right] = -|\mu|^{-1}\ \mathbb{E}\left[Y\ h(Y)|\mu^t Y = z\right]$$

Algorithm : adaptive stratification

1. Choose the optimal allocation.
2. Choose μ as the direction that minimizes

$$V(\mu) := n \ \sigma^2_{I,n}(\mu, \mathsf{A}_\cdot, q^\star_\cdot) = \left(\sum_{i=1}^{I} p_i \ \sigma_i \right)^2 .$$

We have

- $n \ \sigma^2_{I,n}(\mu, \mathsf{A}_\cdot, q^\star_\cdot) = \Xi\left(\nu_i(h,\mu)\right)$ for $h \in \{f, f\phi, f\phi^2\}$

$$\nu_i(h,\mu) := \int_{\{y, \mu^t y \in \mathsf{A}_i\}} h(x) dx.$$

- The gradient $\nabla V(\mu)$ is a function of $\nu_i(h,\mu)$ and $\nabla_\mu \nu_i(h,\mu)$:

$$\nabla_\mu \left[\int_{\{y, \mu^t y \leq z\}} h(y) dy \right] = -|\mu|^{-1} \ \mathbb{E}\left[Y \ h(Y) | \mu^t Y = z \right]$$

$\hookrightarrow$ $\nabla V(\mu)$ can be estimated by a (stratified) Monte Carlo sampler

- Repeat, given the current allocation and the current direction $\mu^{(t)}$
  - estimate $\nabla V(\mu^{(t)})$
  - Update the direction of stratification : $\mu^{(t+1)}$.
  - Update the allocation policy.

- Repeat, given the current allocation and the current direction $\mu^{(t)}$
  - estimate $\nabla V(\mu^{(t)})$
  - Update the direction of stratification : $\mu^{(t+1)}$.
  - Update the allocation policy.

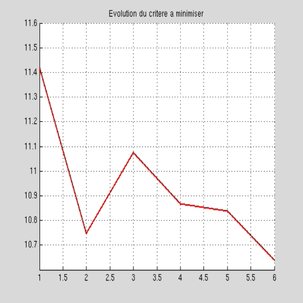
- At each iteration, we have a stratified estimate $\varepsilon^{(t)}$ of $\mathbb{E}[\phi(Y)]$ and an estimate of its variance $\varsigma^{(t)}$. Define the weighted stratified estimator

$$\frac{1}{\sum_{\tau=1}^{t} 1/\varsigma^{(\tau)}} \sum_{\tau=1}^{t} \frac{1}{\varsigma^{(\tau)}} \varepsilon^{(\tau)}.$$

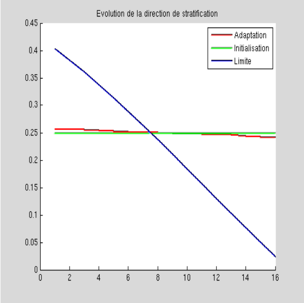
Ex. Pricing of an Asian call (stratif + importance sampling)

$$\mathbb{E}\left[\phi(Y+\nu)\exp(-\nu^t Y - 0.5\nu^t \nu)\right] \qquad Y \sim \mathcal{N}_d(0,\mathbb{I}) \qquad \phi(y) = \exp(-rT)\left(\frac{1}{d}\sum_{k=1}^{d} S_{t_k}(y) - K\right)_+$$

$$T = 1 \qquad S_0 = 50 \qquad r = 0.05 \qquad \upsilon = 0.1 \qquad K = 45 \qquad d = 16.$$



Criterion to minimize



Direction of Stratification

FIG.: [left] Evolution of the variance vs the number of iterations. [right] Successive directions of stratification

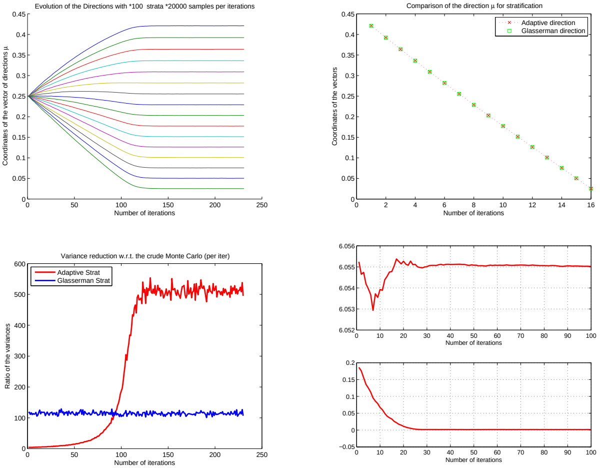


FIG.: As a function of the number of iterations t: [top, left] evolution of the components of the directions $\mu^{(t)}$. [top, right] comparison with the direction of stratification of Glasserman et al. [bottom, left] variance reduction w.r.t. standard MC. [bottom right] weighted estimate and variance $\varsigma^{(t)}$

Confidence interval for the variance of the estimators

Model				Price	Variance			
υ	K	ν	q_i	-	MC	AdaptStr	GHS	μ_{reg}
0.1	45	0	prop	6.05	(8.63922; 8.64018)	-	-	(0.01736; 0.01744)
			opt	6.05	(8.63922; 8.64018)	(0.00348; 0.00353)	-	(0.00472; 0.00477)
		$\nu_\star$	prop	6.05	(0.80272; 0.80311)	-	(0.01375; 0.01376)	(0.01375; 0.01376)
			opt	6.05	(0.80272; 0.80311)	(0.00155; 0.00157)	-	(0.00513; 0.00519)
0.5	45	0	prop	9.00	(158.067; 158.159)	-	-	(2.083; 2.091)
			opt	9.00	(158.067; 158.159)	(0.348; 0.356)	-	(0.358; 0.366)
		$\nu_\star$	prop	9.00	(14.9458; 14.9476)	-	(0.2027; 0.2028)	(0.2248; 0.2249)
			opt	9.00	(14.9458; 14.9476)	(0.1454; 0.1496)	-	(0.1597; 0.1628)
0.5	65	0	prop	2.16	(48.3833; 48.4423)	-	-	(1.8542; 1.8594)
			opt	2.16	(48.3833; 48.4423)	(0.0916; 0.0954)	-	(0.0939; 0.0972)
		$\nu_\star$	prop	2.16	(2.3203; 2.3204)	-	(0.0387; 0.0388)	(0.0457; 0.0458)
			opt	2.16	(2.3203; 2.3204)	(0.0199; 0.0207)	-	(0.0236; 0.0245)
1	45	0	prop	14.01	(851.670; 852.314)	-	-	(52.011; 52.477)
			opt	14.01	(851.670; 852.314	(5.245; 5.540)	-	(5.353; 5.608)
		$\nu_\star$	prop	14.01	(42.759; 42.763)	-	(3.012; 3.015)	(3.183; 3.187)
			opt	14.01	(42.759; 42.763)	(2.219; 2.320)	-	(2.365; 2.446)
1	65	0	prop	7.79	(586.14; 586.94)	-	-	(50.71; 51.05)
			opt	7.79	(586.14; 586.94)	(3.50; 3.98)	-	(2.96; 3.08)
		$\nu_\star$	prop	7.78	(22.331; 22.337)	-	(1.550; 1.551)	(1.744; 1.746)
			opt	7.78	(22.331; 22.337)	(0.963; 1.017)	-	(1.109; 1.1562)

μ_{reg} : regression of $\phi(Y)$ on Y.

Basket options (stratif + importance sampling)

$$\mathbb{E}\left[\phi(Y+\nu)\exp(-\nu^t Y - 0.5\nu^t\nu)\right] \qquad Y \sim \mathcal{N}_d(0,\mathbb{I}) \qquad \phi(y) = \exp(-rT)\left(\sum_{k=1}^{d}\alpha_k S_T^{(k)} - K\right)_+$$

correlated assets

$T = 1 \quad \alpha_k = 1/d \quad r = 0.05 \quad \upsilon = 0.1 \quad K = 45 \quad d = 40.$

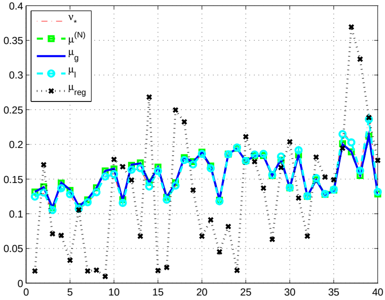


FIG.: Different directions of stratification

Barrier options (stratif + importance sampling)

$$\mathbb{E}\left[\phi(Y+\nu)\exp(-\nu^t Y - 0.5\nu^t\nu)\right] \qquad Y \sim \mathcal{N}_d(0,\mathbb{I}) \qquad \phi(y) = \exp(-rT)\left(d^{-1}\sum_{k=1}^{d} S_{t_k} - K\right)_+ \mathbb{1}_{S_T \leq B}$$

$$S_0 = 50 \quad T = 1 \quad r = 0.05 \quad \upsilon = 0.1 \quad K = 45 \quad B = 60 \quad d = 16.$$

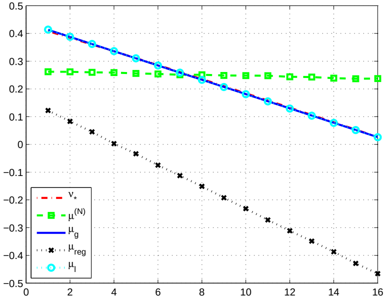


FIG.: Different directions of stratification

Conclusion

- Theoretical results
  - We provided new results on the asymptotic variance of the stratified Monte Carlo estimator.
  - We proved that the choice of the strata does not play a role on this limiting variance.
  - But the variance depends upon (a) the direction of stratification , (b) the allocation.
- Algorithmic discussion
  - we proposed an algorithm to iteratively compute (on the fly) the (a) optimal direction of stratification, (b) the optimal allocation.
- Applications to option pricing
  - Illustration of the convergence of the algorithm + variance reduction along the iterations.
  - Comparison with the results by Glasserman et al. (1999)
  - Comparison with simple guesses for the direction of stratification.

P. Etoré, G. Fort, B. Jourdain and E. Moulines. *On adaptive stratification, 2008. ArXiv math.PR/0809.1135*