THE BLOCKING METHOD

The blocking method is based on the idea of RECURSIVE BINNING

RUN: $X_0, X_1, X_2 \cdots X_{N-1}$ N data

$f_0, f_1, f_2 \cdots f_{N-1}$ $f_i = f(X_i)$

$$\left.\begin{matrix} f_0 \\ f_1 \end{matrix}\right\} F_0^{(1)} \qquad \left.\begin{matrix} f_2 \\ f_3 \end{matrix}\right\} F_1^{(1)} \qquad \left.\begin{matrix} f_4 \\ f_5 \end{matrix}\right\} F_2^{(1)} \qquad \left.\begin{matrix} f_6 \\ f_7 \end{matrix}\right\} F_3^{(1)} \qquad \left.\begin{matrix} f_8 \\ f_9 \end{matrix}\right\} F_4^{(1)}$$

$$\left.\begin{matrix} F_0^{(1)} \\ F_1^{(1)} \end{matrix}\right\} F_0^{(2)} \qquad \left.\begin{matrix} F_2^{(1)} \\ F_3^{(1)} \end{matrix}\right\} F_1^{(2)} \qquad \left.\begin{matrix} F_0^{(2)} \\ F_1^{(2)} \end{matrix}\right\} F_0^{(3)}$$

$$F_0^{(1)} = \frac{1}{2}(f_0 + f_1)$$

$$F_1^{(1)} = \frac{1}{2}(f_2 + f_3)$$

$$\vdots$$

$$F_0^{(2)} = \frac{1}{2}\left(F_0^{(1)} + F_1^{(1)}\right) = \frac{1}{4}(f_0 + f_1 + f_2 + f_3)$$

$$F_1^{(2)} = \frac{1}{2}\left(F_2^{(1)} + F_3^{(1)}\right) = \frac{1}{4}(f_4 + f_5 + f_6 + f_7)$$

$$\vdots$$

$$F_0^{(3)} = \frac{1}{2}\left(F_0^{(2)} + F_1^{(2)}\right) = \frac{1}{8}(f_0 + f_1 \cdots + f_7)$$

Thus from the original data we generate

$F_0^{(1)} \cdots F_{N/2}^{(1)}$ $N/2$ data

$F_0^{(2)} \cdots F_{N/4}^{(2)}$ $N/4$ data

and so on

Now we want to compute the variance of the $F_i^{(k)}$. ②

But before, note that, <u>IN EQUILIBRIUM</u> (the distribution of X_i is always π, independent of time)

$$\langle f(X_i) \rangle_{MC} = \langle f \rangle_\pi$$

$$\langle F_i^{(1)} \rangle_{MC} = \langle \tfrac{1}{2}(f_{2i} + f_{2i+1}) \rangle_{MC} =$$

$$= \left(\langle f_{2i} \rangle_{MC} + \langle f_{2i+1} \rangle_{\pi} \right) \frac{1}{2} = \langle f \rangle_\pi$$

$$\langle F_i^{(2)} \rangle_{MC} = \langle \tfrac{1}{2} (F_{2i}^{(1)} + F_{2i+1}^{(1)}) \rangle_{MC} = \langle f \rangle_\pi$$

All variables have the same average $\langle f \rangle_\pi = F$

Now, let us compute the variance

We begin with $\langle (F_i^{(1)} - F)(F_i^{(1)} - F) \rangle_{MC} = \Delta^{(1)}$

<u>IN EQUILIBRIUM</u> it DOES NOT DEPEND ON TIME.

We set $i=0$

$$\Delta^{(1)} = \langle \left(\tfrac{1}{2} f_0 + \tfrac{1}{2} f_1 - F \right)^2 \rangle_{MC} = \frac{1}{4} \langle (f_0 - F + f_1 - F)^2 \rangle_{MC}$$

$$= \frac{1}{4} \left\{ \langle (f_0 - F)^2 \rangle_{MC} + 2 \langle (f_0 - F)(f_1 - F) \rangle_{MC} + \langle (f_1 - F)^2 \rangle_{MC} \right] =$$

$$= \frac{1}{4} \left[C(0) + 2C(1) + C(0) \right] = \frac{1}{2} \left[C(0) + C(1) \right]$$

$C(n)$ is the autocorrelation function

Now we compute

$$\Delta^{(2)} = \left\langle \left(\frac{1}{4} f_0 + \frac{1}{4} f_1 + \frac{1}{4} f_2 + \frac{1}{4} f_3 - F \right)^2 \right\rangle_{MC}$$

$$= \frac{1}{16} \left[\left\langle \left[(f_0 - F) + (f_1 - F) + (f_2 - F) + (f_3 - F) \right]^2 \right\rangle_{MC} \right]$$

$$= \frac{1}{16} \Big[\langle (f_0 - F)^2 \rangle_{MC} + \langle (f_1 - F)^2 \rangle_{MC} + \langle (f_2 - F)^2 \rangle_{MC} + \langle (f_3 - F)^2 \rangle_{MC}$$

$$+ 2 \langle (f_0 - F)(f_1 - F) \rangle_{MC} + 2 \langle (f_1 - F)(f_2 - F) \rangle_{MC} + 2 \langle (f_2 - F)(f_3 - F) \rangle_{MC}$$

$$+ 2 \langle (f_0 - F)(f_2 - F) \rangle_{MC} + 2 \langle (f_1 - F)(f_3 - F) \rangle_{MC}$$

$$+ 2 \langle (f_0 - F)(f_3 - F) \rangle_{MC} \Big]$$

$$= \frac{1}{16} \left[4\, C(0) + 2 \cdot 3\, C(1) + 2 \cdot 2\, C(2) + 2\, C(3) \right]$$

$$= \frac{1}{4} \left[C(0) + \frac{3}{2} C(1) + \frac{2}{2} C(2) + \frac{1}{2} C(3) \right]$$

We can generalize the result

$$\Delta^{(3)} = \frac{1}{2^3} \Big[C(0) + \frac{7}{4} C(1) + \frac{6}{4} C(2) + \frac{5}{4} C(3) + \frac{4}{4} C(4) + \frac{3}{4} C(5)$$
$$+ \frac{2}{4} C(6) + \frac{1}{4} C(7) \Big]$$

$$\Delta^{(k)} = \frac{1}{2^k} \left[C(0) + \sum_{n=1}^{2^k - 1} \frac{(2^k - n)}{2^{k-1}} C(n) \right]$$

$$= \frac{1}{2^k} \left[C(0) + 2 \sum_{n=1}^{2^{k-1}} C(n) - \frac{1}{2^{k-1}} \sum_{n=1}^{2^{k-1}} n\, C(n) \right]$$

Now let us define the "errors" σ_k^2 that are obtained by assuming that the variables $F^{(k)}$ are independent (uncorrelated)

$$\sigma_1^2 = \frac{1}{\underset{\substack{\uparrow \\ \#\text{ of measures } F^{(1)}}}{N/2}} \operatorname{Var} F^{(1)} = \frac{2}{N}\Delta^{(1)} = \frac{1}{N}\left[C(0)+C(1)\right]$$

$$\sigma_2^2 = \frac{1}{N/4} \operatorname{Var} F^{(2)} = \frac{4}{N}\Delta^{(2)} = \frac{1}{N}\left[C(0)+\frac{3}{2}C(1)+\frac{2}{2}C(2)+\frac{C(3)}{2}\right]$$

$$\sigma_3^2 = \frac{1}{N/8}\Delta^{(3)} = \frac{1}{N}\left[C(0)+\frac{7}{4}C(1)+\frac{6}{4}C(2)+\frac{5}{4}C(3)+\frac{4}{4}C(4) + \frac{3}{4}C(5)+\frac{2}{4}C(6)+\frac{1}{4}C(7)\right]$$

$$\sigma_k^2 = \frac{1}{N/2^k}\Delta^{(k)} = \frac{1}{N}\left[C(0)+2\sum_{n=1}^{2^k-1}C(n) - \frac{1}{2^{k-1}}\sum_{n=1}^{2^k-1} n\,C(n)\right]$$

$$= \frac{1}{N}\Big[C(0)+2\sum_{n=1}^{\infty}C(n) \underbrace{- 2\sum_{n=2^k}^{\infty}C(n) - \frac{1}{2^{k-1}}\sum_{n=1}^{2^k-1} n\,C(n)}_{\text{goes to zero for } k\to\infty}\Big]$$

$$\approx \frac{1}{N}\left[C(0)+2\sum_{n=1}^{\infty}C(n)\right] = \text{correct error for the sample mean of } f:\ \overline{f}.$$

$$\sigma_k^2 - \sigma^2 \sim O\Big(\frac{1}{\underset{\substack{\uparrow \\ \text{length of the block } (m=2^k)}}{m}}\Big) \qquad [\text{Decreases slowly}]$$

CONCLUSION:

The quantity σ_K^2 (the error on the $F^{(k)}$ assuming that these variables are UNCORRELATED) converges to the correct error σ^2 of the variables f_i

THE PRACTICAL PROCEDURE

- Compute $F_i^{(1)}$ $\quad i: 0, \dots \frac{N}{2} - 1$

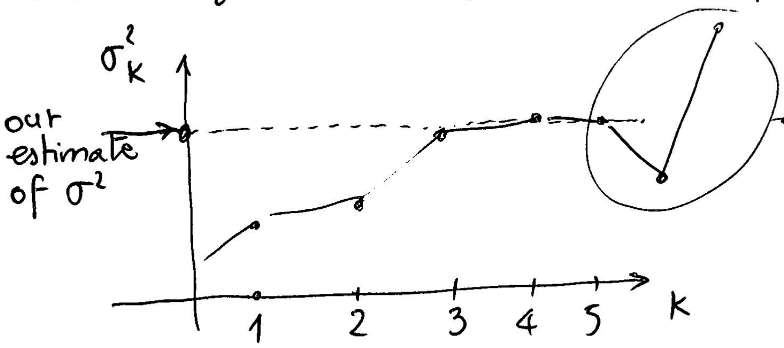
and $$\sigma_1^2 = \frac{1}{N/2} \operatorname{Var} F^{(1)} = \frac{2}{N} \underbrace{\left[\frac{1}{N/2} \sum_{i=0}^{N/2-1} \left(F_i^{(1)}\right)^2 - \left(\frac{1}{N/2} \sum_{i=0}^{N/2-1} F_i^{(1)} \right)^2 \right]}_{\text{This is the variance estimator}}$$

- Compute $F_i^{(2)}$ $\quad i: 0, \dots \frac{N}{4} - 1 = m$

$$\sigma_2^2 = \frac{1}{m} \operatorname{Var} F^{(2)} = \frac{1}{m} \left[\frac{1}{m} \sum_{i=0}^{m-1} \left(F_i^{(2)}\right)^2 - \left(\frac{1}{m} \sum_{i=0}^{m-1} F_i^{(2)} \right)^2 \right]$$

and so on

- Plot σ_i vs. i (the level of blockings)



Fluctuations that cannot be correct (MC fluctuations)

$\underbrace{\hspace{6em}}_{\text{The estimates increase with k}}$

$\boxed{\text{We } \underline{\text{must}} \text{ have } \sigma_{K+1}^2 > \sigma_k^2}$