Statistical properties of time series (continued)

Remark on Millikan (Feynman story about biased data): first diffraction result for e also 1% too low, later data correct within error bars.

Real-world example: DMFT-QMC data + extrapolation

Solution of homework: Is $\sigma^2_{est} = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$ an unbiased estimator for σ^2 in autocorrelated case?

$$\left\langle \sum_{i=1}^{N}(X_i - \bar{X})^2 \right\rangle = \frac{1}{N^2}\sum_i \left\langle \left(N X_i - \sum_{j=1}^{N} X_j\right)^2 \right\rangle$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left\langle \left(N \Delta X_i - \sum_{j=1}^{N}\Delta X_j\right)^2\right\rangle \qquad (\Delta X_i \equiv X_i - \langle x\rangle)$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left\langle N^2 \Delta X_i^2 + \left(\sum_{j=1}^{N}\Delta X_j\right)^2 - 2N\Delta X_i \sum_{j=1}^{N}\Delta X_j\right\rangle$$

$$= \frac{1}{N^2}\sum_{i=1}^{N}\left\langle N^2 \Delta X_i^2 + \sum_{j=1}^{N}\sum_{k=1}^{N}\Delta X_j \Delta X_k - 2N\sum_{j=1}^{N}\Delta X_i \Delta X_j\right\rangle$$

$$= N\sigma^2 - \underbrace{\frac{1}{N}\sum_{i,j=1}^{N}\langle \Delta X_i \Delta X_j\rangle}_{\approx \tau\sigma^2 \text{ (see below)}} \approx (N-\underset{\substack{\uparrow \\ \tau \geq 1}}{\tau})\,\sigma^2$$

$\Rightarrow$ above estimator systematically underestimates σ^2 for autocorrelated data ($\frac{1}{N}$-effect). Better: $\sigma^2_{est,imp} = \frac{\sum_i (X_i - \bar{X})^2}{N-\tau}$

Notes: • no exact unbiased estimator (but bias can be minimized using good approximation for τ)

• σ^2_{est} may be „good enough"

• In general $\langle (X_i - \bar{X})^2 \rangle \neq \langle (X_j - \bar{X})^2 \rangle$ for $i \neq j$ and autocorrelated data (less deviations for inner values)

Step (ii) variance of mean value $\bar{X}$:

example for autocorrelated data: uncorr. Y_i

$\to X_i := \frac{Y_{i-1} + 2Y_i + Y_{i+1}}{4}$

$\Rightarrow \sigma_X^2 = \frac{6}{16}\sigma_Y^2$

$\langle \Delta X_i \Delta X_{i+1} \rangle = \frac{4}{16}\sigma_Y^2$

$$\langle (\bar{X} - \langle x \rangle)^2 \rangle = \langle \left(\frac{1}{N} \sum_{i=1}^{N} x_i - \langle x \rangle \right)^2 \rangle$$

$$= \langle \left[\frac{1}{N} \sum_{i=1}^{N} (x_i - \langle x \rangle) \right]^2 \rangle$$

$$= \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \langle (x_i - \langle x \rangle)(x_j - \langle x \rangle) \rangle$$

(gen. Def: $C_{A,B} = \frac{\langle (A - \langle A \rangle)(B - \langle B \rangle) \rangle}{\sigma_A \; \sigma_B}$)

covariance

$$= \frac{\sigma^2}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i,j}$$

$$\overset{\text{translation inv.}}{=} \frac{\sigma^2}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} C_{i-j}$$

$k = i - j$

$j = i - k$

$$= \frac{\sigma^2}{N^2} \sum_{i=1}^{N} \sum_{k=i-N}^{i-1} C_k$$

$$\underset{C_k \approx 0 \text{ for } |k| > K \ll N}{\approx} \frac{\sigma^2}{N^2} \sum_{i=1}^{N} \sum_{k=-\infty}^{\infty} C_k = \frac{\sigma^2}{N} \tau$$

autocorrelation time $\tau = \sum_{k=-\infty}^{\infty} C_k \overset{C_k = C_{-k}}{=} 1 + 2 \sum_{k=1}^{\infty} C_k$

$$C_k \approx \frac{\frac{1}{N-k-1} \sum_{i=1}^{N-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2}$$

autocorrelation function

Estimate for $S_{ij} = \langle (X_i - \langle X \rangle)(X_j - \langle X \rangle) \rangle$? $\quad (i \neq j)$

How good is $\tilde{S}_{ij} = \langle (X_i - \bar{X})(X_j - \bar{X}) \rangle$?

$$\tilde{S}_{ij} = \langle (X_i - \frac{1}{N}\sum_{k=1}^{N} X_k)(X_j - \frac{1}{N}\sum_{m=1}^{N} X_m) \rangle$$

$$= \langle [(X_i - \langle X \rangle) - \frac{1}{N}\sum_{k=1}^{N}(X_k - \langle X \rangle)][(X_j - \langle X \rangle) - \frac{1}{N}\sum_{m=1}^{N}(X_m - \langle X \rangle)] \rangle$$

$$= S_{ij} - \frac{1}{N}\sum_{m=1}^{N} S_{im} - \frac{1}{N}\sum_{k=1}^{N} S_{kj} + \frac{1}{N^2}\sum_{k,m=1}^{N} S_{km}$$

$$= \left(1 - \frac{1}{N}\right)^2 S_{ij} - \frac{1}{N}(S_{ii} + S_{jj}) + \frac{1}{N^2}\sum_k S_{kk} + \text{offdiag terms}$$

$$= \left(1 - \frac{1}{N}\right)^2 S_{ij} - \frac{\sigma^2}{N} + \text{offdiag terms}$$

$$\xrightarrow[\text{data}]{\text{uncorr}} -\frac{\sigma^2}{N}$$

$$\Rightarrow \boxed{S_{ij} \approx \frac{\tilde{S}_{ij} + \frac{\sigma^2}{N}}{\left(1 - \frac{1}{N}\right)^2}}$$

example: $N = 3$: uncorr.

$$\tilde{S}_{12} = \langle \frac{2\Delta X_1 - \Delta X_2 - \Delta X_3}{3} \, \frac{2\Delta X_2 - \Delta X_1 - \Delta X_3}{3} \rangle$$

$$= \frac{1}{9}[\Delta X_3^2 - 2\Delta X_1^2 - 2\Delta X_2^2] = -\frac{\sigma^2}{3}$$

$\frac{1}{9}(4S_{12} + S_{21} - S_{13} - S_{23})$