Revision/practice questions on statistical inference/range-based estimators for σ and α

1. Let $R_t := \bar{X}_t - \underline{X}_t$ denote the **range** of $X_t = \sigma W_t$ over $[0,t]$, where $\bar{X}_t = \max_{0\le s\le t} X_s$ and $\underline{X}_t = \min_{0\le s\le t} X_s$. Using that $\mathbb{E}(|X_t|) = \sigma\sqrt{\frac{2t}{\pi}}$ and $\mathbb{E}(R_t) = 2\sigma\sqrt{\frac{2t}{\pi}}$, derive the **minimal variance** unbiased estimator for σ of the form $\hat{\sigma} = \lambda_1|X_1| + \lambda_2 R_1$. You may use that

$$\mathrm{Var}(\hat{\sigma}) \quad = \quad \sigma^2[\lambda_1^2(1-\frac{2}{\pi}) + \lambda_2^2(4\log 2 - \frac{8}{\pi}) + 2\lambda_1\lambda_2(\frac{3}{2}-\frac{4}{\pi})].$$

Solution. $\mathbb{E}(\hat{\sigma}) = \mathbb{E}(\lambda_1|X_1| + \lambda_2 R_1) = \sigma(\lambda_1 + 2\lambda_2)\sqrt{\frac{2}{\pi}}$. Hence $\hat{\sigma}$ is unbiased if $(\lambda_1+2\lambda_2)\sqrt{\frac{2}{\pi}} = 1$, so $\lambda_2 = \frac{1}{2}(\sqrt{\frac{\pi}{2}} - \lambda_1)$. Subject to this constraint, $\mathrm{Var}(\hat{\sigma})$ is then just a quadratic in λ_1 only, and we just minimize $\mathrm{Var}(\hat{\sigma})$ over λ_1 to get

$$\lambda_1^* \quad = \quad \frac{\sqrt{\frac{\pi}{2}}\,(4\log 2 - 3)}{\log 16 - 2} \quad \approx \quad -.369$$

for which we find that $\mathrm{Var}(\hat{\sigma}) \approx .0625\sigma^2$.

2. From the joint density of $(W_t, \bar{W}_t)$ in the Reflection Principle chapter of FM02, we can show that

$$\mathbb{P}(\bar{X}_1 \le b | X_1 = x) \quad = \quad 1 - e^{-2b(b-x)/\sigma^2}$$

for $b \ge \max(x,0)$. Use this to compute the distribution of $2\bar{W}_1(\bar{W}_1 - W_1)$, and the variance of $\hat{\sigma}^2 = 2\bar{X}_1(\bar{X}_1 - X_1)$ where $X_t = \sigma W_t$ (recall from FM02 Hwk 4 that $\mathbb{E}(\hat{\sigma}^2) = \sigma^2$, so $\hat{\sigma}^2$ is an unbiased estimator for σ^2).

Solution. Let $h(b) = 2b(b-x)$. Then

$$\mathbb{P}(2\bar{W}_1(\bar{W}_1 - W_1) \le y | W_1 = x) \;=\; \mathbb{P}(h(\bar{W}_1) \le y | W_1 = x) \;=\; \mathbb{P}(\bar{W}_1 \le h^{-1}(y) | W_1 = x) \;=\; 1 - e^{-h(h^{-1}(y))} \;=\; 1 - e^{-y}.$$

for $y \ge 0$.

But the final answer is independent of x, so $2\bar{W}_1(\bar{W}_1 - W_1)$ is independent of W_1 with distribution $\mathrm{Exp}(1)$ (which has variance 1). Then since $2\bar{X}_1(\bar{X}_1 - X_1) = \sigma^2 \cdot 2\bar{W}_1(\bar{W}_1 - W_1)$, we see that the variance of $2\bar{X}_1(\bar{X}_1 - X_1)$ is σ^4.

Remark 0.1 The **antithetic version** of $\hat{\sigma}^2$ defined by $\hat{\sigma}^2_{RS} = \bar{X}_1(\bar{X}_1 - X_1) + \underline{X}_1(\underline{X}_1 - X_1)$ (discussed in Hwk 3 q4 in FM02) attains a lower variance of $0.331\sigma^4$, which is known as the **Rogers-Satchell** (RS) estimator (and remains unbiased with the same variance even if we add a drift to X). Another well known unbiased estimator called the **Garman-Klass** estimator takes the form

$$\sigma^2_{GK} \quad = \quad \frac{1}{2}R_1^2 - (2\log 2 - 1)X_1^2$$

where $R_1 = \bar{X}_1 - \underline{X}_1$, and σ^2_{GK} attains an even lower variance of $.27\sigma^4$ but is biased for non-zero drift.

3. The joint density of the **drawdown** $Y_t = \bar{X}_t - X_t$ and $\bar{X}_t$ is

$$p(y,b;t) \quad = \quad \frac{2(b+y)}{\sigma^3\sqrt{2\pi t^3}}\, e^{-\frac{(b+y)^2}{2\sigma^2 t}}.$$

What is the admissible range for y and b here? Using that

$$\int_0^\infty \lambda e^{-\lambda t} p(y,b;t)\,dt \quad = \quad 2\lambda e^{-\sqrt{2\lambda}\,(b+y)} \tag{1}$$

what can we say about $\bar{X}_T - X_T$ and $\bar{X}_T$, if $T \sim \mathrm{Exp}(\lambda)$ with T independent of W? (see similar question 3d in Mock-SampleQuestions in FM02.

Solution. The admissible range is just $0 \le y < \infty$ and $0 \le b < \infty$. The right hand side of Eq ... is the joint density of Y_T and $\bar{X}_T$, but we note that it can be broken up as

$$\sqrt{2\lambda}\, e^{-\sqrt{2\lambda}\, y} \cdot \sqrt{2\lambda} e^{-\sqrt{2\lambda}\, b}$$

so we see that $Y_T = \bar{X}_T - X_T$ and $\bar{X}_T$ are both i.i.d. $\mathrm{Exp}(\sqrt{2\lambda})$ random variables.

4. Recall from e.g. Hwk 5 q3 in FM02 that a **symmetric α-stable process** X with parameters $\alpha \in (0,2]$, $\sigma > 0$ is a generalization of Brownian motion, which has **independent stationary increments** like Brownian motion but now $\mathbb{E}(e^{iu(X_t-X_s)}|X_s) = e^{-(t-s)\sigma^\alpha|u|^\alpha}$ for $u \in \mathbb{R}$ and $0 \le s \le t$, so X is only a (multiple of) BM if $\alpha = 2$, but for $\alpha < 2$ the increments of X are not normally distributed and the process exhibits positive and negative **jumps** over any time interval.

Compute $\mathbb{E}(\bar{X}_t(\bar{X}_t - X_t))$ for $\sigma = 1$ and $\alpha \in (1,2)$. You may use that $\bar{X}_T$ and $\bar{X}_T - X_T$ are independent if $T \sim \mathrm{Exp}(\lambda)$ is independent of X and that $\mathbb{E}(\bar{X}_t) = \frac{\alpha\Gamma(1-\frac{1}{\alpha})}{\pi} t^{\frac{1}{\alpha}}$ (also seen in FM02), and that

$$\int_0^\infty \lambda e^{-\lambda t} t^q dt \quad = \quad \lambda^{-q}\Gamma(1+q)\,. \tag{2}$$

Solution. Setting $q = \frac{1}{\alpha}$, we see that

$$\mathbb{E}(\bar{X}_T) \quad = \quad \int_0^\infty \lambda e^{-\lambda t}\mathbb{E}(\bar{X}_t)dt \quad = \quad c_\alpha \lambda^{-\frac{1}{\alpha}}$$

where $c_\alpha = \frac{\alpha}{\pi}\Gamma(1-\frac{1}{\alpha})\Gamma(1+\frac{1}{\alpha})$. Then

$$\int_0^\infty \lambda e^{-\lambda t}\mathbb{E}(\bar{X}_t(\bar{X}_t - X_t))dt \quad = \quad \mathbb{E}(\bar{X}_T(\bar{X}_T - X_T)) \quad = \quad \mathbb{E}(\bar{X}_T)\,\mathbb{E}((\bar{X}_T - X_T)) \quad = \quad \mathbb{E}(\bar{X}_T)^2 \quad = \quad c_\alpha^2 \lambda^{-\frac{2}{\alpha}}$$

where we have used that $\bar{X}_T$ and $\bar{X}_T - X_T$ are independent for the second equality, and that $\mathbb{E}(X_T) = \int_0^\infty \lambda e^{-\lambda t}\mathbb{E}(X_t)dt = 0$, since $\mathbb{E}(X_t) = 0$.

Comparing to (2), we see that $\mathbb{E}(\bar{X}_t(\bar{X}_t - X_t)) = \frac{c_\alpha^2}{\Gamma(1+\frac{2}{\alpha})} t^{\frac{2}{\alpha}}$. After some more simplification, one can show this agrees with the tidier formula $\mathbb{E}(\bar{X}_t(\bar{X}_t - X_t)) = \frac{1}{\Gamma(1+\frac{2}{\alpha})\sin(\frac{\pi}{\alpha})^2} t^{\frac{2}{\alpha}}$ used in Hwk 7 in FM02 (proof not required).

Remark 0.2 Note $\mathbb{E}(\bar{X}_t(\bar{X}_t - X_t))$ is finite even though $\mathbb{E}(X_t^2)$ and $\mathbb{E}(\bar{X}_t^2)$ are not since X_t has fat tails.

5. Let $X_t = \sigma W_t$. Let $\bar{X}_t = \max_{0\le s\le t} X_s$ and $\underline{X}_t = \min_{0\le s\le t} X_s$. Then

$$\mathbb{P}(X_t \le x, \bar{X}_t \le b) \quad = \quad \mathbb{P}(W_t \le \frac{x}{\sigma}, \bar{W}_t \le \frac{b}{\sigma})\,.$$

Then differentiating wrt x and b, we see that the density of $(X_t, \bar{X}_t)$ is $\frac{1}{\sigma^2} f(\frac{x}{\sigma}, \frac{b}{\sigma}; t)$ where $f(x,b;t) = \frac{2(2b-x)}{\sqrt{2\pi t^3}}\, e^{-\frac{(2b-x)^2}{2t}}$ is the joint density of $(W_t, \bar{W}_t)$ from the ReflectionPrinciple chapter in FM02, which evaluates to

$$p(x,b;t) \quad = \quad \frac{2(2b-x)}{\sigma^3\sqrt{2\pi t^3}}\, e^{-\frac{(2b-x)^2}{2\sigma^2 t}}$$

on the domain $-\infty < x \le b$, $b \ge 0$ (since clearly $X_t \le \bar{X}_t$ and $\bar{X}_t \ge 0$). Using that

$$\frac{d}{d\sigma}\log p(x,b;t) \quad = \quad \frac{1}{\sigma^3 t}((2b-x)^2 - 3t\sigma^2)$$

derive the Maximum Likelihood Estimate (MLE) $\hat{\sigma}$ for σ based on a single observation of $(X_t, \bar{X}_t)$ when $t = 1$. Compute the mean and variance of $\hat{\sigma}^2$ when $t = 1$, and compare the latter to the RS and GK estimators above.

Solution. Setting the **score**: $\frac{d}{d\sigma}\log p(X_t, \bar{X}_t; t)$ equal to zero, we see that the MLE $\hat{\sigma}$ for σ is

$$\hat{\sigma} \quad = \quad \frac{2\bar{X}_t - X_t}{\sqrt{3t}}$$

and $\hat{\sigma}^2$ is an unbiased estimator for σ^2. Now letting $t = 1$, from the 3a in Mock-SampleQuestions in FM02 we know that $(2\bar{X}_1 - X_1)^2 \sim \chi_3^2$ i.e. the same distribution as the sum of squares of three independent Brownian motions at time 1 (see also q1b in same mock on the Bessel(3) process), so $\mathrm{Var}(\hat{\sigma}^2) = \frac{2\cdot k}{9}\sigma^4 = \frac{2}{3}\sigma^4$.

Note this is much larger than the variance of the aforementioned RS and GK estimators, since $\hat{\sigma}^2$ is not antithetic, but we make this antithetic by replacing $\hat{\sigma}$ with $\frac{1}{2}\frac{2\bar{X}_t - X_t}{\sqrt{3t}} + \frac{1}{2}\frac{-2\underline{X}_t + X_t}{\sqrt{3t}}$, since $(X_t, \bar{X}_t)$ has the same joint distribution as $(-X_t, -\underline{X}_t)$.

See `https://colab.research.google.com/drive/1ksEEZHm4x7VW129YDTOOpVaEdCFVaTn-?usp=sharing`. for numerical examples to empirically compute the variance of these range-based estimators for Brownian motion.

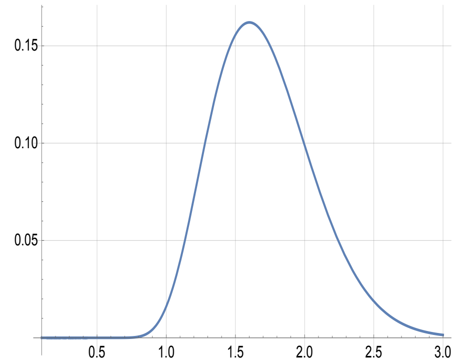

Figure 1: Likelihood function of σ in q6 for $n=1$ observation with $X_t=0$, $\bar{X}_t=1$ and $\underline{X}_t=-1$.

6. Let $X_t=\sigma W_t$, and recall from FM02 that

$$\mathbb{P}(W_t\in dy, M_t<b, m_t>a) \quad = \quad \frac{2}{b-a}\sum_{n=1}^{\infty} e^{-\lambda_n t}\sin(\frac{n\pi(y-a)}{b-a})\sin(\frac{n\pi(0-a)}{b-a})dx \quad = \quad q(y,a,b)dy$$

for $y\in(a,b)$, where $\lambda_n=\frac{n^2\pi^2}{2(b-a)^2}$. Compute the MLE for σ from a single observation of $(X_t,\bar{X}_t,\underline{X}_t)$, and explain how we would use this in practice.

Remark 0.3 This series is absolutely convergent, which implies that it converges, i.e. $\sum|a_n|<\infty\Rightarrow\sum a_n<\infty$, where a_n is the summand in the infinite sum.

Solution. Using a similar argument to the previous question, the joint density of $(X_t,\bar{X}_t,\underline{X}_t)$ is $-q_{ab}(\frac{x}{\sigma},\frac{a}{\sigma},\frac{b}{\sigma})\frac{1}{\sigma^3}$., where q_{ab} means $\frac{\partial^2 q}{\partial a\partial b}q$. The MLE for a single observation is the σ-value which maximizes the **joint likelihood function** $-q_{ab}(\frac{X_t}{\sigma},\frac{\underline{X}_t}{\sigma},\frac{\bar{X}_t}{\sigma})\frac{1}{\sigma^3}$.

Given daily observations r_i of the increments of X, and the (relative) daily highs H_i and lows of X for n days, the MLE is σ-value $\hat{\sigma}_n$ which maximizes the joint likelihood function:

$$L(\sigma) \quad = \quad -\frac{1}{\sigma^{3n}}\prod_{i=1}^{n} q_{ab}(\frac{r_i}{\sigma},\frac{L_i}{\sigma},\frac{H_i}{\sigma})$$

using the same notation as FM02. One can use e.g. the Garman-Klass estimator above as a smart initial guess for minimization scheme here.

Remark 0.4 The MLE is **asymptotically efficient** which means that $\sqrt{n}(\hat{\sigma}_n-\sigma)\sim N(0,\frac{1}{I(\theta)})$ as $n\to\infty$, where $I(\theta)=\mathbb{E}((\frac{\partial}{\partial\sigma}\log L(\sigma))^2)$ is the **Fisher information** ($\frac{1}{nI(\theta)}$ is the **Cramér-Rao** lower bound for the variance of any unbiased estimator).

7. Let $T_1,...,T_n$ denote an i.i.d. sequence of Exp(λ) random variables. Then from the Strong Law of Large Numbers (SLLN) from FM02, we know that the sample mean $\frac{1}{n}\sum_{i=1}^n T_i$ tends to $\mathbb{E}(T_i)=\frac{1}{\lambda}$, and note that $\frac{1}{n}T_i\sim\text{Exp}(n\lambda)$. Now set $\lambda=\frac{1}{t}$ for some **fixed** maturity time t of interest.

For the symmetric α-stable process X above, we have that

$$\mathbb{E}(e^{-\beta\bar{X}_{e_q}}) \quad = \quad e^{-\frac{1}{\pi}\int_0^\infty \frac{\beta}{u^2+\beta^2}\log(1+\frac{u^\alpha}{q})\,du} \tag{3}$$

for $\beta\geq 0$, where $e_q\sim\text{Exp}(q)$ is independent of X. Using the **Gil-Pelaez** formula from Hwk7 in FM02 by setting $\beta=-iz$ with $z\in\mathbb{R}$, we can use this to compute the cdf $F(.)$ of $\bar{X}_{e_q}$. Using the first part of the question, explain how we can approximately jointly sample $\bar{X}_t$ and $\bar{X}_t-X_t$.

Solution. Based on the first paragraph, we set $q=n\lambda=\frac{n}{t}$, and draw n i.i.d. samples of $\bar{X}_{e_q}$ and $\bar{X}_{e_q}-X_{e_q}$, using $F^{-1}(U)$ where $U\sim U[0,1]$ (see AppliedProbabilityRevision chapter in FM02).

From q4, we know $\bar{X}_{e_q}$ and $\bar{X}_{e_q}-X_{e_q}$ are independent, and (by a general result) it is known that $\bar{X}_{e_q}-X_{e_q}\sim -\underline{X}_{e_q}$ and (by symmetry) $-\underline{X}_{e_q}\sim\bar{X}_{e_q}$.

We now proceed as follows: let $(S_i,D_i)_{i\geq 1}$ be i.i.d. pairs with

$$S_i\sim\bar{X}_{e_q},\quad D_i\sim\bar{X}_{e_q}-X_{e_q},\quad \Delta X_i=S_i-D_i$$

with S_i and D_i are independent (and we know $S_i \sim D_i$ from above). Set $X_0 = 0$, $M_0 = 0$, and and for $i \geq 1$ we let

$$X_i = X_{i-1} + \Delta X_i, \qquad M_i = \max(M_{i-1}, X_{i-1} + S_i).$$

Then M_n approximates $\bar{X}_t$ and $M_n - X_n$ approximates $\bar{X}_t - X_t$ when $q = n/t$ and n is large. This procedure is known as **Canadization** or **Erlangization**. There is also another method for sampling X_t and $\bar{X}_t$ called the stick breaking algorithm which does not require the Gil-Pelaez Fourier inversion described above.