Time of the maximum

- Let $\bar{W}_t = \max_{0\leq s\leq t} W_s$ and let $\theta_t = \min\{s : W_s = \bar{W}_t\}$, i.e. the **first time** s that X achieves its **final maximum** at time t.
- You may assume that the s-value for which $W_s = \bar{W}_t$ is **unique** almost surely without proof, so we can simplify the discussion and just refer to $\theta_t \leq t$ as the **time of the maximum** at time t.
- When $\gamma = 0$, one can show that the joint density of $(W_t, \bar{W}_t, \theta_t)$ is

$$q(x,b,s) \quad = \quad 2 f_{\tau_b}(s)\, f_{\tau_{b-x}}(t-s) \tag{1}$$

  where $f_{\tau_b}(t) = \frac{b}{\sqrt{2\pi t^3}} e^{-b^2/2t}$ is the usual hitting time density to $b > 0$ at time t for standard Brownian motion starting at zero which we have previously computed
- Similar to last week, if we multiply the X process by σ, then the joint density of $(X_t, \bar{X}_t, \theta_t^X)$ becomes

$$q(x,b,s) \quad = \quad \frac{2}{\sigma^2} f_{\tau_{\frac{b}{\sigma}}}(s)\, f_{\tau_{\frac{b-x}{\sigma}}}(t-s)\,.$$

  This known as Williams's **path decomposition** at the maximum.

Example 1. Using that

$$\frac{\partial}{\partial\sigma}\log q(x,b,s) \quad = \quad \frac{b^2 t - 2bsx + s(x^2 + 3(s-t)\sigma^2)}{s(t-s)\sigma^3}$$

at $\sigma^2 = \frac{b^2 t - 2bsx + sx^2}{4s(t-s)} \geq 0$, derive the Maximum Likelihood Estimator (MLE) for $\hat{\sigma}^2$ from a single observation of $(X_t, \bar{X}_t, \theta_t)$.

Solution. Setting $x = X_t$, $b = \bar{X}_t$ and $s = \theta$, the "score" is

$$\frac{\partial}{\partial\sigma}\log q(X_t,\bar{X}_t,\theta_t) \quad = \quad \frac{\bar{X}_t^2 t - 2\bar{X}_t\theta_t X_t + \theta_t(X_t^2 + 3(\theta_t - t)\sigma^2)}{s(t-s)\sigma^3}\,. \tag{2}$$

Then setting the numerator to zero, we find that

$$\hat{\sigma}^2 \quad = \quad \frac{\bar{X}_t^2 t - 2\bar{X}_t\theta_t X_t + \theta_t X_t^2}{4\theta_t(t-\theta_t)}\,.$$

Example 2. One can show that the conditional cdf of θ_t given $W_t = x$ and $\bar{W}_t = b$ is

$$F_2(s;x,b) \quad = \quad \frac{1}{2}\mathrm{erfc}\Big(\frac{bt-(2b-x)s}{\sqrt{2ts(t-s)}}\Big) \; - \; \frac{x}{2b-x} e^{\frac{2b(b-x)}{t}} \frac{1}{2}\mathrm{erfc}\Big(\frac{bt-xs}{\sqrt{2ts(t-s)}}\Big)$$

where $\mathrm{erfc}(x) = 2(1-\Phi(x\sqrt{2}))$. Explain how we can use this to jointly sample $(W_1, \bar{W}_1, \theta_1)$.

Solution. Let U_1, U_2, U_3 be i.i.d. $U[0,1]$ random variables. Then set

$$\begin{aligned} W_1 &= \Phi^{-1}(U_1) \\ \bar{W}_1 &= F^{-1}(U_2; W_1) \\ \theta_1 &= F_2^{-1}(U_3; W_1, \bar{W}_1) \end{aligned}$$

where $F(b;x) = \mathbb{P}(\bar{W}_1 \leq b \,|\, W_1 = x) = 1 - e^{-2b(b-x)}$, see q2 last week.

Remark 0.1 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).

Example 3. If we have n i.i.d. samples of $(X_t, \bar{X}_t, \theta_t)$ (e.g. daily observations of a stock price), what can we say about $\sqrt{n}(\hat{\sigma}_n - \sigma)$ as $n \to \infty$?.

Solution. As we saw last week, the MLE $\hat{\sigma}_n$ is **asymptotically efficient**, which means that $\sqrt{n}(\hat{\sigma}_n - \sigma) \sim N(0, \frac{1}{I(\theta)})$ as $n \to \infty$, where

$$I(\theta) \quad = \quad \mathbb{E}\Big(\big(\frac{\partial}{\partial\sigma}\log L(\sigma)\big)^2\Big) \tag{3}$$

is the **Fisher information**, where the log-likelihood function here is $\log L(\sigma) := \log q(X_t, \bar{X}_t, \theta_t)$, so $\frac{\partial}{\partial\sigma}\log L(\sigma)$ is given (2). **Look at Python from last week**

Quadratic variation of a jump process

Let (Y_i) be a sequence of i.i.d. random variables with density $\nu(x)$ and $(N_t)_{t\geq 0}$ be a **Poisson process** for which $\mathbb{E}(e^{pN_t}) = e^{\lambda t(e^p-1)}$ (and recall that $N_t \in \mathbb{N}$), and assume N_t is independent of (Y_i). Compute the mgf of $X_t = \sum_{i=1}^{N_t} g(Y_i)$ for some general function g.

Solution. From the tower property of conditional expectation

$$\mathbb{E}(e^{p\sum_{i=1}^{N_t} g(Y_i)}) = \mathbb{E}(\mathbb{E}(e^{p\sum_{i=1}^{N_t} g(Y_i)}|N_t)) = \mathbb{E}(\mathbb{E}(\prod_{i=1}^{N_t} e^{pg(Y_i)})|N_t) = \mathbb{E}(\mathbb{E}(e^{pg(Y_1)})^{N_t}) = \mathbb{E}(e^{V(p)N_t}) = e^{\lambda t(\mathbb{E}(e^{pg(Y_1)})-1)}$$

where the penultimate equality follows since the Y_i's are i.i.d. We can further re-write the final line as

$$\mathbb{E}(e^{pX_t}) = e^{\lambda t(\int_{-\infty}^{\infty} e^{pg(x)}\nu(x)dx-1)} = e^{\lambda t\int_{-\infty}^{\infty}(e^{pg(x)}-1)\nu(x)dx}.$$

A particular case of interest is when $g(x) = x^2$, in which case $X_t = \sum_{s\leq t}\Delta X_s^2$, which is known as the quadratic variation of X at time t.