Applying the result above gives an exponential approximation. (It is, by the way, summable, so we automatically recover the SLLN).

- ***Theorem (Azuma's Inequality)***: Let $\{Z_n\}$ be a zero-mean martingale with bounded MG differences (ie: $-\alpha \le Z_i - Z_{i-1} \le \beta$ for $\alpha, \beta \ge 0$). Then

$$\mathbb{P}\left(\bigcup_{n=m}^{\infty} \left|Z_n\right| > n\varepsilon\right) \le 2\exp\left(-\frac{2m\varepsilon^2}{(\alpha+\beta)^2}\right)$$

This bound is not as tight as the CLT's, but it requires less.

- ***Example***: $S_n =$ number of heads in n flips, where $\mathbb{P}(\text{Heads}) = p$. $Z_n = S_n - np$ is a martingale with $-p \le Z_i - Z_{i-1} \le 1 - p$. As such, we can use Azuma's inequality and obtain $\mathbb{P}\left(\bigcup_{n=m}^{\infty} \left|\frac{S_n}{n} - p\right| > \varepsilon\right) \le 2\exp\left(-2m\varepsilon^2\right)$. □

- ***Definition (Doob Martingale)***: Let X be a random variable in L^1 and $\mathcal{F}_n$ be a set of filtrations. Then $X_n = \mathbb{E}\left(X \mid \mathcal{F}_n\right)$ is a martingale.

Proof: $\mathbb{E}\left(X_{n+1} \mid \mathcal{F}_n\right) = \mathbb{E}\left[\mathbb{E}\left(X \mid \mathcal{F}_{n+1}\right) \mid \mathcal{F}_n\right] = \mathbb{E}\left(X \mid \mathcal{F}_n\right) = X_n$. ■

- Let $\boldsymbol{X} = (X_1, \cdots, X_n)$, where the X_i are independent and with CDF F_i. Define $\mathcal{F}_i = \sigma(X_1, \cdots, X_i)$. Finally, let $h : \mathbb{R}^n \to \mathbb{R}$ such that, if $\boldsymbol{x}$ differs from $\boldsymbol{y}$ in only one component, $\left|h(\boldsymbol{x}) - h(\boldsymbol{y})\right| \le \ell$, for some $\ell \ge 0$. Then $S_i = \mathbb{E}\left[h(\boldsymbol{X}) \mid \mathcal{F}_i\right]$ is a Doob martingale. Provided we can prove $\left|S_i - S_{i-1}\right| \le \ell$, we can apply Azuma's Inequality with $\alpha + \beta = \ell$ to $S_n = h(\boldsymbol{X})$

$$\mathbb{P}\left(\left|h(\boldsymbol{X}) - \mathbb{E}\left[h(\boldsymbol{X})\right]\right| > n\varepsilon\right) \le 2\exp\left(-\frac{n\varepsilon^2}{2\ell^2}\right)$$

Proof: To prove $\left|S_i - S_{i-1}\right| \le \ell$, note that

$$S_i = \mathbb{E}\left[h(\boldsymbol{X}) \mid \mathcal{F}_i\right] = \int_{x_{i+1}} \cdots \int_{x_n} h\left(X_1, \cdots, X_i, x_{i+1}, \cdots, x_n\right) \mathrm{d}F_n(x_n) \cdots \mathrm{d}F_{i+1}(x_{i+1})$$

$$S_{i-1} = \int_{x_i} \cdots \int_{x_n} h\left(X_1, \cdots, x_i, x_{i+1}, \cdots, x_n\right) \mathrm{d}F_n(x_n) \cdots \mathrm{d}F_{i+1}(x_{i+1}) \mathrm{d}F_i(x_i)$$

As such, remembering that densities integrate to 1

$$\begin{aligned} \left|S_i - S_{i-1}\right| &\le \int \cdots \int \left| \quad \right| \, \mathrm{d}\cdots \\ &\le \ell \int \cdots \int 1 \, \mathrm{d}\cdots \\ &= \ell \end{aligned}$$

As required. ■

This second argument is *incorrect*, because in going from the first to the second line, we condition B_t on the past while ignoring that we are at the same time condition B_t on the future (because we are conditioning on $B_1 = 0$). $\square$

- ***Martingales Associated with Brownian Motion***
  - Recall that in continuous time, the martingale property reads, for all $t > s$, $\mathbb{E}\left(X(t) \mid \mathcal{F}_s\right) = X(s)$. Defining a stopping time is more tricky; if T satisfies $\left\{T \leq t\right\} \in \mathcal{F}_t$, then since $\mathcal{F}$ is an increasing family, $\left\{T < t\right\} = \bigcup_{n=1}^{\infty} \left\{T \leq t - \frac{1}{n}\right\} \in \mathcal{F}_t$. However, if T only satisfies $\left\{T < t\right\} \in \mathcal{F}_t$, then $\left\{T \leq t\right\} = \bigcap_n \left\{T < t + \frac{1}{n}\right\} \notin \mathcal{F}_t$ (again, since the family is increasing). To conclude this, we need to assume right-continuity of the filtration.
  - ***Theorem***: The following three processes are martingales:
    - $\left\{B_t\right\}$ (the "mean martingale")
    - $\left\{B_t^2 - t\right\}$ (the "variance martingale")
    - $\left\{\exp\left(\theta B_t - \frac{1}{2}\theta^2 t\right)\right\}$ where θ is a deterministic parameter (the "exponential martingale").

    Proof: The first two parts are trivial. For the last, recall that $\mathbb{E}\left(e^{\theta N(0,1)}\right) = e^{\theta^2/2}$ and $\mathbb{E}\left(e^{\theta B_t - \frac{1}{2}\theta^2 t} \mid \mathcal{F}_s\right) = \left(e^{\theta B_s - \frac{1}{2}\theta^2 t}\right)\mathbb{E}\left(e^{\theta(B_t - B_s)} \mid \mathcal{F}_s\right) = \left(e^{\theta B_s - \frac{1}{2}\theta^2 t}\right)\mathbb{E}\left(e^{\theta[N(0,1)]\sqrt{t-s}}\right)$. $\blacksquare$
  - The exponential martingale can be used to generate many other martingales. Let

$$f(\theta; t, x) = e^{\theta x - \frac{1}{2}\theta^2 t} = \sum_{n=0}^{\infty} \frac{\theta^n}{n!} H_n(t, x)$$

    Where H_n is the n^{th} Hermite polynomial, $H_n(t, x) = f^{(n)}(0; t, x)$. Feeding this into the martingale property and exchanging summation and expectation, we can use the fact that this holds for *any* θ including $\theta = 0$, and conclude that for each n, $\left\{H_n(t, B_t)\right\}$ is also a martingale.
  - We can apply the Optional Stopping Theorem to these martingales to get some interesting results.
  - ***Example***: Define $T = \inf\left\{t : B_t = -a \text{ or } b\right\}$. Using the mean martingale, we can find $p_b = a / (a + b)$. Using the variance martingale, we can find $\mathbb{E}(T) = ab$. $\square$

- ***Example***: Let $X_t = \mu t - \sigma B_t$, with $\mu, \sigma > 0$; this could be seen as the "net demand up to time *t*", where σB_t is a production process. We are interested in $T = \inf\left\{t : X_t = b > 0\right\}$ (the first time stock depletes). $\mathbb{E}(T) = \frac{b}{\mu}$ is easily found using the mean martingale on B_T. Use the variance martingale for $\mathbb{V}\text{ar}(T) = \frac{\sigma^2 b}{\mu^3}$ □
- ***Example***: $X_t = \mu t + \sigma B_t$, $T = \inf\left\{t : X_t \notin (-a, b)\right\}$. Use $e^{\theta B_T - \frac{1}{2}\theta^2 T} = e^{\theta \frac{X_T - \mu T}{\sigma} - \frac{1}{2}\theta^2 T}$. Choose $\theta = -2\mu / \sigma$, and use the OST (stopped martingale is bounded); $\mathbb{E}(e^{\frac{\theta}{\sigma} X_T}) = 1$. Directly gives p_b and p_a. Use OST on $B_T = (X_T - \mu T) / \sigma$ to find $\mathbb{E}(T)$. Can let $\mu < 0$ and $a \to -\infty$ and find $\mathbb{P}\left(\sup_t X_t \geq b\right) = e^{-2b\mu/\sigma^2}$. □
- ***Example***: Following from the above example and letting $T = \inf\left\{t : X_t = b\right\}$, suppose we want $\mathbb{E}(e^{-\gamma T})$. Write $-\gamma T = \theta B_T - \frac{1}{2}\theta^2 T - \beta b$.
  - Use the OST on $\exp(\theta B_{T\wedge t} - \frac{1}{2}\theta^2[T \wedge t]) \leq \exp(\theta B_{T\wedge t})$ [bounded.]
  - Substitute $b = \mu T + \sigma B_T$ and equate coefficients of T and B_T.

Ito's Formula

- For a deterministic x_t, $\mathrm{d}x_t = \dot{x}\ \mathrm{d}t$ and $\mathrm{d}f(x_t) = f'(x_t)\mathrm{d}x_t + \frac{1}{2}f''(x_t)(\mathrm{d}x_t)^2 + \cdots$, but $(\mathrm{d}x_t)^2 = \dot{x}^2(\mathrm{d}t)^2$ which vanishes. In BM, $(\mathrm{d}B_t)^2 \sim \mathrm{d}t$, and so

$$\mathrm{d}f(B_t) = f'(B_t)\ \mathrm{d}B_t + \tfrac{1}{2}f''(B_t)\ \mathrm{d}t$$
$$f(B_t) = f(0) + \int_0^t f'(B_s)\ \mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\ \mathrm{d}s$$

This is Ito's Formula. The first integral above is called Ito's Integral and can be approximated as

$$\int_0^t X_s\ \mathrm{d}B_s \approx \sum_{i=1}^n X_{t_{i-1}}\left[B_{t_i} - B_{t_{i-1}}\right]$$

This is a martingale transformation, and is therefore a martingale under boundedness and predictability ($\Leftarrow$ left-continuity) of X_t. Furthermore, as a martingale, the mean of the integral is 0. To find its variance, consider

$$\mathbb{E}\left(\int_0^t X_s \mathrm{d}B_s\right)^2 \approx \mathbb{E}\left(\sum_{i=1}^n X_{t_{i-1}}\left[B_{t_i} - B_{t_{i-1}}\right]\right)^2 = \sum_{i=1}^n \mathbb{E}(X_{t_{i-1}}^2)(t_i - t_{i-1}) \approx \mathbb{E}\left(\int_0^t X_s^2 \mathrm{d}s\right)$$

Where we have used orthogonality of martingale differences, and conditioned on $\mathcal{F}_{t_{i-1}}$. This is knows as *Ito's Isometry*. More generally, for a bivariate $f(t, x)$

$$
\begin{aligned}
p &= \mathbb{P}\left((X_0, X_1, \cdots) \in A\right) \\
&=_{A \text{ shift invariant}} \mathbb{P}\left((X_0, X_1, \cdots) \in A, (X_n, X_{n+1}, \cdots) \in A\right) \\
&\to_{\text{mixing}} \mathbb{P}\left((X_0, X_1, \cdots) \in A\right) \cdot \mathbb{P}\left((X_0, X_1, \cdots) \in A\right) \\
&= p^2
\end{aligned}
$$

And so p must be equal to 0 or 1. ■

- ***Example***: Let $X_0 = Y$ w.p. p and Z w.p $1-p$, where $Y, Z \in L_1$. Consider two cases:
  - If $X_n \equiv X_0$, $\bar{X} = Y$ w.p p and Z w.p $1-p$
  - If $X_n =_d X_0$, $\bar{X} = \mathbb{E}(Y)$ w.p p and $\mathbb{E}(Z)$ w.p $1-p$

  Neither case is ergodic.

Part IV – Stochastic Orders

Introduction

- The Hazard Rate of a random variable X with CDF F is

$$\lambda_X(t) = \lim_{dt \to 0} \mathbb{P}\left[X \in (t, t+\mathrm{d}t) \mid X > t\right] = \frac{f(t)}{\bar{F}(t)} \qquad t \geq 0$$

  As such

$$\lambda_X(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[-\ln \bar{F}(t)\right] \qquad \bar{F}(t) = \exp\left(-\int_0^t \lambda_X(s)\, \mathrm{d}s\right)$$

  X follows the exponential distribution $\bar{F}(t) = e^{-t/\mu}$ if and only if $\lambda_X(t) = \frac{1}{\mu}$.

  Example: If $W = \max(X, Y)$, then $\bar{F}_W = \bar{F}_X \bar{F}_Y$, and $\lambda_W = \lambda_X + \lambda_Y$. ■

- Assume X and Y have densities F and G. We define the following variable orderings
  - ***Likelihood ratio ordering***: $X \geq_{\text{LR}} Y$ if $\frac{f(y)}{f(x)} \leq \frac{g(y)}{g(x)}$ for all $x \geq y$.
  - ***Hazard rate ordering***: $X \geq_{\text{HR}} Y$ if $\lambda_X(x) \leq \lambda_Y(x)$ for all x.
  - ***Stochastic ordering***: $X \geq_{\text{ST}} Y$ if $\bar{F}(x) \geq \bar{G}(y)$ for all x.
  - ***Increasing convex ordering***: $X \geq_{\text{ICX}} Y$ if $\mathbb{E}\left[\phi(X)\right] \geq \mathbb{E}\left[\phi(Y)\right]$ for all increasing and convex functions $\phi(x)$,
- ***Theorem***: $X \geq_{\text{LR}} Y \Rightarrow X \geq_{\text{HR}} Y \Rightarrow X \geq_{\text{ST}} Y \Rightarrow X \geq_{\text{ICX}} Y$