Statistics 212: Lecture 23 (April 28, 2025)

Girsanov's theorem

Instructor: Mark Sellke

Scribe: Diego Luca Gonzalez Gauss, Enrico Yao-Bate

0.1 Final Exam

The final exam will be held at Harvard Hall 101 on May 9th, from 9 AM to 12 PM. It will be similar to the midterm in format and length. Material covered will span the whole course, with an emphasis on what has been covered after the midterm. We can bring 2 double-sided pages of notes.

1 Some new motivating questions

Let $\mathbb{P}, \mathbb{Q}$ be probability measures on the same space with $\mathbb{P} \ll \mathbb{Q}$, $\mathbb{Q} \ll \mathbb{P}$ and let

$$Z = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}.$$

Given filtration $(\mathscr{F}_t)$, we can also define

$$Z_t = \left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right)_{\mathscr{F}_t}$$

Proposition 1. Z_t is a martingale with respect to $\mathbb{P}$.

Proof. Let $s < t$. It suffices to show that

$$\mathbb{E}^{\mathbb{P}}\left[Z_t \mid \mathscr{F}_s\right] = Z_s.$$

Let $A \subseteq \mathscr{F}_s$. Since A is $\mathscr{F}_t$-measurable,

$$\mathbb{E}^{\mathbb{P}}[Z_t \mathbf{1}_A] = \mathbb{E}^{\mathbb{P}}[\mathbf{1}_A (\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}})_{\mathscr{F}_t}] = \mathbb{E}^{\mathbb{P}}[Z\mathbf{1}_A].$$

And because A is $\mathscr{F}_s$-measurable,

$$\mathbb{E}^{\mathbb{P}}[Z\mathbf{1}_A] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}[\mathbf{1}_A (\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}})_{\mathscr{F}_t} \mid \mathscr{F}_s]\right] = \mathbb{E}^{\mathbb{P}}[\mathbf{1}_A (\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}})_{\mathscr{F}_s}] = \mathbb{E}^{\mathbb{P}}[Z_s\mathbf{1}_A],$$

so $\mathbb{E}^{\mathbb{P}}[Z_t\mathbf{1}_A] = \mathbb{E}^{\mathbb{P}}[Z_s\mathbf{1}_A]$ holds for every $A \subseteq \mathscr{F}_s$, which implies $\mathbb{E}^{\mathbb{P}}\left[Z_t \mid \mathscr{F}_s\right] = Z_s$ by the uniqueness of conditional expectation, so we are done. □

We can remember our question from last time, that is: if $\mathbb{P}_\mu$ is a Brownian Motion with drift μ, $\mathbb{P} = \mathbb{P}_0$, what is $Z_t = (d\mathbb{P}_\mu / d\mathbb{P})_{\mathscr{F}_t}$. We found that $Z_t(\mu) = \exp(\mu B_t - \frac{\mu^2 t}{2})$, and similarly

$$\frac{Z_t(\mu)}{Z_t(\theta)} = \left(\frac{d\mathbb{P}_\mu}{d\mathbb{P}_\theta}\right)_{\mathscr{F}_t} = \exp((\mu-\theta)B_t - \frac{(\mu^2-\theta^2)t}{2}$$

now, we carry onto a new set of questions:

Proposition 2. **Some motivating questions.** Let $\mu, L > 0$, what is

1. *the probability that a Brownian Motion with drift* μ *hits* $-L$*?*
2. *the conditional behavior of the Brownian Motion, given that we hit* $-L$*?*

Let us first answer 1. To do this, we can identify two methods. In our first, we can take our Brownian Motion with drift B_t to be $B_t = \tilde{B}_t + \mu t$ for $\tilde{B}_t$ our genuine Brownian Motion. Letting

$$M_t = e^{-2\mu B_t},$$

we see that:

$$dM_t = -2\mu M_t d\tilde{B}_t - 2\mu^2 M_t dt + 2\mu^2 M_t dt.$$

Further, $M_0 = 1$ and $M_\tau = e^{2\mu L}$ for τ a hitting time of L, which means for $\tau \to \infty$ that $M_t \to 0$, and by Optional Stopping Theorem we get we altogether equal $e^{-2\mu L}$.

There's another way to think about these questions that leverages our identity from last class. Here we can let $\theta = -\mu$, and get

$$\left(\frac{d\mathbb{P}_\mu}{d\mathbb{P}_{-\mu}}\right)_{\mathscr{F}_t} = \exp((\mu - (-\mu))B_t - \frac{(\mu^2 - (-\mu)^2)t}{2} = \exp(2\mu B_t)$$

by routine Optional Stopping Theorem tricks we can extend this for any stopped filtration, which gives

$$\left(\frac{d\mathbb{P}_\mu}{d\mathbb{P}_{-\mu}}\right)_{\mathscr{F}_\tau} = \exp(2\mu B_\tau)$$

for τ our hitting time for $-L$. For us, $B_\tau = -L$; this tells us that the event $B_\tau = -L$ is $e^{-2\mu L}$ more likely to appear for BM with drift μ than BM with drift $-\mu$.

Then by the LLN, we also have $\mathbb{P}^{-\mu}[\tau < \infty] = 1$ and

$$\mathbb{P}^\mu[\tau < \infty] = \int \mathbf{1}_{\tau<\infty} d\mathbb{P}^\mu = \int \mathbf{1}_{\tau<\infty} e^{-2\mu L} d\mathbb{P} = e^{-2\mu L} \times \mathbb{P}^{-\mu}[\tau < \infty],$$

which by Optional Stopping Theorem again equals $e^{-2\mu L}$.

So, our answer for (1) is $e^{-2\mu L}$, while our answer for (2) is that the conditional behavior of BM, given that we hit $-L$, is as if we have a drift $-\mu$ until our hitting time τ. This is because in the course of the proof above, we actually showed that the Radon-Nikodym derivative $\left(\frac{d\mathbb{P}_\mu}{d\mathbb{P}_{-\mu}}\right)_{\mathscr{F}_\tau}$ is constant on the event that $\tau < \infty$ (and this event holds almost surely under $\mathbb{P}_{-\mu}$).

To build more intuition about the this behavior, we can consider the discrete case. We can let $P > 1/2$ and do SRW with $\mathbb{P}[X_{k+1} - X_k = 1 \mid \mathscr{F}_k] = p$, else $X_{k+1} - X_k = -1$, and again consider the conditional behavior for our BM given that 'X_k hits $-L$'. Each stopped path takes U up-steps and $U + L$ downsteps, and so the probability of this path is $\mathbb{P}^P(\text{path}) = p^U(1-p)^{U+L}$. Then letting $\mathbb{P}^{-p}$ have opposite drift, we get $\mathbb{P}^{-p}[\text{path}] = p^{u+c}(1-p)^u = \mathbb{P}^p[\text{path}] + \left(\frac{p}{1-p}\right)^L$, where the second term is constant since we've crystallized our L. Then for every path ending at the first hitting time of $-L$ we have

$$\frac{d\mathbb{P}^p}{d\mathbb{P}^{-p}}(\text{path}) = \left(\frac{1-p}{p}\right)^L < 1$$

and under $\mathbb{P}^{-p}$, our hitting time is a.s. finite, so

$$\sum_{\text{paths ending at } \tau_t} \mathbb{P}^{-p}[\text{path}] = 1 \Longrightarrow \sum_{\text{paths ending at } \tau_t} \mathbb{P}^p[\text{path}] = \left(\frac{1-p}{p}\right)^L$$

Prof. Sellke is then asked whether a similar mechanic will work for more complicated drifts a la $\mu^2 t + \mu t^2$, and whether negation is sufficient here, to which Prof. Sellke mentions that this is a good lead-in to Girsanov's theorem. To introduce this, we look at a little more machinery.

1.1 Doob's h-transform

Namely, suppose we condition on some event A in the future. We are working here with a discrete-time Markov chain $(X_0,\dots,X_T)$. What is

$$\mathrm{Law}(X_{t+1} \mid X_0,\dots,X_t, A)\Big|_{I(A)}?$$

By Bayes' rule, this is proportional to the ordinary condition law

$$\mathrm{Law}(X_{t+1} \mid X_0,\dots,X_t)\cdot \mathbb{P}(A \mid X_{t+1})$$

for $\mathbb{P}[A|X_0,\dots,X_{t+1}] = h(X_{t+1})$, where hX_{t+1} is the probability that your BM hits -L with drift in the continuous case, by tilting. This h is known as Doob's h-transform.

Girsanov's theorem will let us carry this result to the continuum.

2 Girsanov's theorem

To introduce Girsanov's, we'll start with the Cameorn-Martin theorem, which states

Theorem 2.1 (Cameron-Martin)**.** *For $\mu : [0,T] \to \mathbb{R}$, $\int_0^T \mu(t)^2 dt < \infty$, and $\mathbb{P}_\mu$ the law of*

$$B_t + \int_0^t \mu(S) ds$$

(i.e. a Brownian Motion with drift μ_t) then

$$\left(\frac{d\mathbb{P}_\mu}{d\mathbb{P}_0}\right)_{\mathscr{F}_t} = \exp(\int_0^t \mu_s dB_s - \int_0^t \frac{\mu_s^2}{2} ds)$$

(note that this is a local martingale by Ito's formula). Roughly, this is saying the same as what we see earlier: if we have a BM with drift, then we end up just tilting a BM by some exponential factor, but our machinery is now a bit broader.

Theorem 2.2 (Girsanov's)**.** *Girsanov's theorem is the same conclusion with μ_S progressively measurable. We need some boundedness conditions to ensure we actually have a martingale, but for now we'll ignore these technical hoops to demonstrate the bigger picture (e.g. it suffices that $\mathbb{E}[\exp(\int_0^T \frac{\mu_s^2}{2} dT)] < \infty$ a.k.a. Novikov's criterion, Revuz-Yor chapter 8).*

Proof. To prove Girsanov, we'll (handwaving some technical criteria) simply consider the law of B_t, to get

$$\begin{aligned} \mathbb{E}^{\mathbb{P}_\mu}\left[\exp\left(\alpha(B_t - \int_0^t \mu_s ds)\right)\right] = \mathbb{E}^{\mathbb{P}_0}\left[\exp\left(\alpha(B_t \int_0^t \mu_s ds) + \int_0^t \mu_s dB_s - \int_0^t \mu_s^2/2 ds\right)\right] = \\ \mathbb{E}\left[\exp\left(\int_0^t (\alpha+\mu_S) dB_s - \int_0^t \frac{2\alpha\mu_s + \mu_s^2}{2} ds\right)\right] = \\ = e^{\alpha^2 t/2} \times \mathbb{E}\left[\exp\left(\int_0^t (\alpha t \mu_s) dB_s - \frac{1}{2}\int_0^t (\alpha+\mu_s)^2 ds\right)\right] = e^{\alpha^2 t/2} \end{aligned}$$

where it can be validated that $\exp(\int_0^t \mu_s dB_s - \int_0^t \frac{\mu_s^2}{2} ds)$ is a local martingale as mentioned earlier, so $\mathbb{E}[\exp(\int_0^t (dt\mu_s) dB_s - \frac{1}{2}\int_0^t (\alpha+\mu_s)^2 ds)] = 1$ (we're implicitly using that this martingale has expectation 1). $\square$

Let's look at some examples working through conditional BM behavior. Say we start a $B_0 = 1$ conditioned on reaching 10 before 0, described until some τ. Our intuition tells us that conditioning should just pull us upwards; to check, let's show what this looks like on the discrete case. We have a SRW conditioned again to reach 10 before 0. If $X_t = k$ we get by Bayes' theorem

$$\frac{\mathbb{P}\left[X_{t+1} = k+1 \mid \mathscr{F}_k, A\right]}{\mathbb{P}\left[X_{t+1} = k-1 \mid \mathscr{F}_k, A\right]} = \frac{\mathbb{P}\left(X_{t+1} = k+1 \mid \mathscr{F}_t\right)}{\mathbb{P}\left[X_{t+1} = k-1 \mid \mathscr{F}_t\right]} \times \frac{\mathbb{P}[A \mid X_{t+1} = k+1]}{\mathbb{P}[A \mid X_{t+1} = k-1]} = \frac{k+1}{k-1}.$$

In other words,

$$\frac{\mathbb{P}\left[X_{t+1} = k+1 \mid \mathscr{F}_k, A\right]}{\mathbb{P}\left[X_{t+1} = k-1 \mid \mathscr{F}_k, A\right]} = \frac{k+1}{2k}, \tag{1}$$

which might be summarize by saying the conditional law of X_{t+1} has upward drift $1/k$. Interestingly, this doesn't depend on the value of 10. Thus, if we condition X_t to reach some very large N before reaching 0, we get the same dynamics (until N is reached). We can "take N to infinity" and define a Markov chain with the dynamics of (1) for all time. This is often referred to as "simple random walk conditioned to escape to infinity" (even though one cannot actually condition on this event as it has probability 0).

Now we'll extend to the continuous case. We'll let τ be a hitting time for $\{10, 0.01\}$, and by our earlier work we know conditioning on behavior until τ will tilt the Wiener measure by B_τ up to constant factor. Ito's formula gives us that

$$\log(B_t) = \int_0^t \frac{dB_s}{B_s} - \int_0^t \frac{ds}{2B_s^2}$$

and taking $\mu_t = \frac{1}{B_t}$ we get that

$$\exp\left(\int_0^T \mu_t dB_t - \int_0^t \frac{\mu_t^2}{2} dt\right) = B_t$$

Ultimately this means the tilting is equivalent to having a Brownian Motion with drift $1/B_t$. (We can again send $10 \to \infty$ and $0.1 \to 0$. The reason for using 0.1 is to avoid justifying the above formula for $\log(B_t)$ when B_t reaches zero.)