36-771 Fall 2018

Martingales 1 : Concentration inequalities

The one and only homework

Lecturer : Aaditya Ramdas

There will be 8 questions in total, each worth 5 points.

Question 1 (An exponential-polynomial martingale that looks like a mixture) *Let S_n be the sum of n iid Exp(1) random variables. Show that*

$$M_n = \frac{n!}{(1+S_n)^{n+1}} \exp(S_n)$$

is a martingale with respect to the natural filtration.

Question 2 (Throwing away balls yields a martingale) *Consider an urn with m yellow and m pink balls (so $2m$ balls in total). At each step, we draw a random ball from the urn, note its color, and throw it away. For $n = 0, 1, \ldots, 2m-1$, let Y_n be the number of yellow balls left in the urn after n steps of this process (so $Y_0 = m$), and let*

$$M_n = \frac{Y_n}{2m-n}$$

be the fraction of yellow balls left in the urn. Let $(\mathcal{F}_n)$ be the natural filtration generated by the process Y_n. Prove that (M_n) is a martingale wrt $(\mathcal{F}_n)$.

Question 3 (Adding balls yields a martingale) *Consider an urn with 1 yellow and 1 pink balls (so 2 balls in total). At each step, we draw a random ball from the urn, note its color, and add another ball of the same color back into the urn. For $n = 1, \ldots,$ let Y_n be the number of yellow balls left in the urn after n steps of this process (so $Y_0 = 1$), and let*

$$M_n = \frac{Y_n}{n+2}$$

be the fraction of yellow balls in the urn. Let $(\mathcal{G}_n)$ be the natural filtration generated by the process Y_n. Prove that (M_n) is a martingale wrt $(\mathcal{G}_n)$. Are $(\mathcal{G}_n)$ and $(\mathcal{F}_n)$ the same? Why or why not?

Question 4 (Counterexamples) *Design 2×2 matrix examples to show that (a) the matrix exponential is not operator monotone, (b) the matrix exponential is not operator convex, (c) the matrix square is not operator monotone, (d) in general $\exp(A+B) \neq \exp(A)\exp(B)$. Keep it simple so that it is easy to verify the calculations.*

Question 5 (Will you ever stop?) *Imagine you are tossing independent fair coins taking values ± 1, and let S_n be the running sum. We stop tossing coins when $S_n = 1$, let τ be the achieved stopping time. Prove that $\mathbb{E}\tau = \infty$. However, suppose we stop whenever the S_n reaches either $+a$ or $-b$, for positive a, b, prove that $\mathbb{E}\tau < \infty$.*

Question 6 (Wald's Sequential Probability Ratio Test (SPRT)) *Let's say we are testing $H_0 : X \sim P$ against $H_1 : X \sim Q$, where P and Q have densities f, g by observing an infinite sequence $X_1, X_2, \dots$. Prove that the likelihood ratio sequence (L_n) of Q to P is a non-negative martingale under the null hypothesis wrt the natural filtration. The SPRT stops and rejects the null at the first time such that $L_n > 1/\alpha$. Use Ville's inequality to justify that this test controls the type-1 error at level α (that is, when the null is true, the probability that the test will* ever *stop and reject the null is at most α). (If you're curious: what is the connection of this test to the Neyman-Pearson lemma?)*

Question 7 (What's $L_t(\lambda)$?) *Identify the supermartingales that are implicitly associated with the (a) matrix Hoeffding's inequality for bounded-random variables, (b) matrix Bennett's inequality for upper-bounded random variables, (c) matrix Bernstein's inequality, and (d) matrix self-normalized inequality for symmetric random variables. Which of these inequalities is the tightest, if one is cares about the concentration properties of a sum of independent symmetric random variables bounded between $[-1, 1]$, and why?*

Question 8 (Simulating student tea) *Using Theorem 1(b), write down* ***different*** *explicit, closed-form uniform concentration inequalities that would apply to a sum of independent random variables with a (a) t-distribution with 2 degrees of freedom, (b) a t-distribution with 3 degrees of freedom (but would not apply to (a)), (c) a t-distribution with 4 degrees of freedom (but would not apply to (a,b)). Feel free use wikipedia. In all 3 cases, for any illustrative choice of m, x, plot many long sample paths of S_t against V_t, and report what fraction of them crossed the chosen line and compare it to the Theorem 1(b) prediction. (the words "illustrative", "many" and "long" mean that an estimate of the fraction can be reasonably accurate)*