Solution to homework 11

Statistics 205B: Spring 2008

1. Use Kolmogorov $0-1$ law in order to prove that if B is a BM then $\inf\{t>0 : B_t > 0\} = 0$ a.s.

   Solution: Enough to how that $\tau = \inf\{2^{-n} : B_{2^{-n}} > 0\} = 0$. First note that $\mathbb{P}(\tau \le 2^{-n}) \ge \mathbb{P}(B_{2^{-n}} > 0) = 1/2$ for all n. Hence taking limit as $n \to \infty$ we have $\mathbb{P}(\tau = 0) \ge 1/2$. Now define $X_n = 2^{(n+1)/2}(B_{2^{-n}} - B_{2^{-n-1}})$ for $n \ge 1$. Clearly X_n's are i.i.d $\mathrm{N}(0,1)$ and $B_{2^{-n}} = \sum_{i \ge n} 2^{-(i+1)/2} X_i \in \sigma\{X_n, X_{n+1}, \ldots\}$ (the scaling is only for understanding, it's not needed for the proof). Hence $\tau \in \cap_n \sigma\{X_n, X_{n+1}, \ldots\}$, the tail sigma field and $\mathbb{P}(\tau = 0) = 1$ by Kolmogorov's $0-1$ law.

2. Use Kolmogorov $0-1$ law in order to prove that if B is a BM then $\limsup_{t\to\infty} B_t/\sqrt{t} = \infty$.

   Solution: Clearly we have
   $$\limsup_{t\to\infty} \frac{B_t}{\sqrt{t}} \ge \limsup_{n\to\infty} \frac{B_n}{\sqrt{n}}.$$
   Hence it is enough to show that
   $$\limsup_{n\to\infty} \frac{B_n}{\sqrt{n}} \ge K \text{ a.s.}$$
   for any positive constant K. Fix K. Now define $X_i = B_i - B_{i-1}$ for $i = 1, 2, \ldots$. Then clearly X_i's are i.i.d $\mathrm{N}(0,1)$ and the for the n-th partial sum $S_n := \sum_{i=1}^n X_i$ we have $S_n = B_n$. Hence the event
   $$\left(\limsup_{n\to\infty} \frac{B_n}{\sqrt{n}} \ge K\right) = \limsup_{n\to\infty} \left(\frac{B_n}{\sqrt{n}} \ge K\right) = \limsup_{n\to\infty} \left(\frac{S_n}{\sqrt{n}} \ge K\right)$$
   is in the tail sigma-field of X_i's, Hence if we show that the probability of the above event is positive, by Kolmogorov's $0-1$ law we have
   $$\mathbb{P}\left(\limsup_{n\to\infty} \frac{B_n}{\sqrt{n}} \ge K\right) = 1.$$
   Now the above follows from the following
   $$\mathbb{P}\left(\limsup_{n\to\infty} \left(\frac{S_n}{\sqrt{n}} \ge K\right)\right) \ge \limsup_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \ge K\right) = \mathbb{P}(N(0,1) \ge K) > 0.$$

3. Given a finite metric space (Ω, d) define a metric on probability measures on Ω by letting:
   $$d_{\mathcal{K}}(\mu, \nu) = \min_{X\sim\mu, Y\sim\nu} \mathbb{E}\, d(X, Y).$$

(The inf is taken over all couplings of μ and ν where μ and ν are probability measures on Ω.) Show that $d_{\mathcal{K}}$ is indeed a metric.

Hint: Given the coupling $p(x, y)$ of μ_1 and μ_2 and another $p(y, z)$ of μ_2 and μ_3 consider the coupling $p(x, y)p(y, z)/p(y)$ of the three spaces.

Solution: $d_{\mathcal{K}}$ is a metric because

(a) (non-negativity) Clearly, $d_{\mathcal{K}}(\mu, \nu) \geq 0$.

(b) $d_{\mathcal{K}}(\mu, \mu) = 0$ and $d_{\mathcal{K}}(\mu, \nu) = 0$ implies $\min_{X\sim\mu,Y\sim\nu} \mathbb{E}\ d(X, Y) = 0$. If we prove that, "For any two probability measures μ, ν on a finite space Ω, there is a coupling such that $\mathbb{E}\ d(X, Y) = d_{\mathcal{K}}(\mu, \nu)$" $(*)$

then we are done. Then there is a coupling (X, Y) such that $X \sim \mu, Y \sim \nu$ and $\mathbb{E}\ d(X, Y) = 0$, hence $X \equiv Y$ and $\mu = \nu$.

Now to prove $(*)$, w.l.o.g. assume $\Omega = \{1, 2, \dots, n\}$. Consider the set of all probability measures $\mathcal{P}$ on $\Omega \times \Omega$ considered as a subset of $[0, 1]^{n^2}$. Fix any two probability measures μ, ν on Ω. Then the set $\mathcal{P}(\mu, \nu) = \{\pi \in \mathcal{P} : (X, Y) \sim \pi \implies X \sim \mu, Y \sim \nu\}$ is a compact subset of $[0, 1]^{n^2}$ and the function f on $\mathcal{P}$ defined as $f(\pi) = \sum_{x,y\in\Omega} d(x, y)\pi(x, y)$ is a continuous function on $\mathcal{P}$. Hence f achieves its minima on $\mathcal{P}(\mu, \nu)$.

For this problem the existence of the optimal coupling is not needed, by the following observation that $d_* = \min_{x\neq y} d(x, y) > 0$ and $\mathbb{E}d(X, Y) \geq d_*\mathbb{P}(X \neq Y)$. Hence $d_{\mathcal{K}}(\mu, \nu) \geq d_*\mathrm{TV}(\mu, \nu) = d_* \sup_A |\mu(A) - \nu(A)|$.

(c) (symmetry) Clearly $d_{\mathcal{K}}(\mu, \nu) = d_{\mathcal{K}}(\nu, \mu)$.

(d) (triangle inequality) Given the coupling $(X_1, Y_1) \sim p(x, y)$ of μ_1 and μ_2 and $(Y_2, Z_2) \sim p(y, z)$ of μ_2 and μ_3 consider the coupling $(X, Y, Z) \sim p(x, y)p(y, z)/\mu_2(y)$ of the three probability measures. Then marginally we have $X \sim \mu_1, Y \sim \mu_2, Z \sim \mu_3$ and we have

$$d_{\mathcal{K}}(\mu_1, \mu_3) \leq \mathbb{E}\ d(X, Z) \leq \mathbb{E}\ d(X, Y) + \mathbb{E}\ d(Y, Z) = \mathbb{E}\ d(X_1, Y_1) + \mathbb{E}\ d(Y_2, Z_2).$$

Since this is true for any coupling (X_1, Y_1) and (Y_2, Z_2), minimizing over all couplings we have the result.

4. Let $G = (V, E)$ be a graph, $\Omega = V$ and d be the path metric on Ω - in other words, for $u, v \in \Omega$, $d(u, v)$ is the length of shortest path connecting u and v in the graph G. Let P be a Markov chain defined on Ω and suppose that for any edge $e = (u, v) \in E$ it holds that $d_{\mathcal{K}}(\delta_u P, \delta_v P) \leq a < 1$. Show that for all measures μ and ν on Ω it holds that $d_{\mathcal{K}}(\mu P^t, \nu P^t) \leq a^t d_{\mathcal{K}}(\mu, \nu)$.

   Hint: First show the claim for $t = 1$ and μ and ν delta measures; For the general case and $t = 1$, use a coupling of μ and ν that achieves the distance $d_{\mathcal{K}}$ between them; for higher values of t use induction.

   Solution: It is enough to prove for $t = 1$ then using induction we have $d_{\mathcal{K}}(\mu P^t, \nu P^t) = d_{\mathcal{K}}(\mu P^{t-1}P, \nu P^{t-1}P) \leq a d_{\mathcal{K}}(\mu P^{t-1}, \nu P^{t-1}) \leq \cdots \leq a^t d_{\mathcal{K}}(\mu, \nu)$. Assume that $t = 1$.

Assume $\mu = \delta_u, \nu = \delta_v$ for some $u, v \in V$ (not necessarily adjacent). Then note that $d_{\mathcal{K}}(\mu, \nu) = d(u, v) =: k$ (say). Let $u = u_0, u_1, \ldots, u_k = v$ be a path from u to v of length k where each (u_{i-1}, u_i) is an edge in G. Then we have

$$d_{\mathcal{K}}(\mu P, \nu P) = d_{\mathcal{K}}(\delta_{u_0} P, \delta_{u_k} P) \leq \sum_{i=1}^{k} d_{\mathcal{K}}(\delta_{u_{i-1}} P, \delta_{u_i} P) \leq ak = ad(\mu, \nu)$$

where the first inequality follows by triangle inequality and the second one by hypothesis.

Now for general probability distributions μ_i on Ω use the result that

$$d_{\mathcal{K}}(\sum_i \alpha_i \mu_i) \leq \sum_i \alpha_i d_{\mathcal{K}}(\mu_i)$$

where $\alpha_i > 0$ and $\sum_i \alpha_i = 1$. This follows from a simple coupling argument (EXERCISE!)

Now given two p.m. μ and ν consider the optimal coupling $\pi(x, y)$ attaining the $d_{\mathcal{K}}(\mu, \nu)$ bound. Then $\mu P = \sum_{x,y} \pi(x, y) \delta_x P$ and $\nu P = \sum_{x,y} \pi(x, y) \delta_y P$. Hence

$$d_{\mathcal{K}}(\mu P, \nu P) \leq \sum_{x,y} \pi(x, y) d_{\mathcal{K}}(\delta_x P, \delta_y P) \leq a \sum_{x,y} \pi(x, y) d(x, y) = ad_{\mathcal{K}}(\mu, \nu).$$

5. Show that under the setup of the previous exercise for any two measure μ and ν it holds that $\mathrm{TV}(\mu P^t, \nu P^t) \leq a^t \mathrm{Diam}(G)$ where $\mathrm{Diam}(G)$ is maximal distance between any two vertices in G.

   Solution: Note that $d(u, v)$ is integer valued. Hence $\mathbb{P}(X \neq Y) = \mathbb{P}(d(X, Y) > 0) = \mathbb{P}(d(X, Y) \geq 1) \leq \mathbb{E}d(X, Y)$ and this implies that $\mathrm{TV}(\mu, \nu) \leq d_{\mathcal{K}}(\mu, \nu)$. Also we have $d(u, v) \leq \mathrm{Diam}(G)$ which implies that $d_{\mathcal{K}}(\mu, \nu) \leq \mathrm{Diam}(G)$.