[Solution of Homework 2]

Problem 1 (Optimal Coupling)

Let Ω be a finite state space and μ, ν be two distributions over Ω. Prove that there exists a coupling ω of μ and ν such that

$$\mathbf{Pr}_{(X,Y)\sim\omega}\left[X \neq Y\right] = D_{\mathrm{TV}}(\mu, \nu).$$

You need to explicitly describe how ω is constructed.

Proof.

Let $P(x, y)$ denote $\mathbf{Pr}_{(X,Y)\sim\omega}\left[X = x, Y = y\right]$ and use ρ to denote $D_{\mathrm{TV}}(\mu, \nu)$ for shorthand. The coupling ω can be constructed as follows: First, we set $P(a, a) = \min\{\mu(a), \nu(a)\}, \forall a \in \Omega$. Clearly, if $D_{\mathrm{TV}}(\mu, \nu) = 0$, the forementioned setting is indeed the optimal coupling. Otherwise, for any $a \in \Omega$, let

$$\begin{aligned} R_X(a) &= \mu(a) - P(a, a) \\ R_Y(a) &= \nu(a) - P(a, a). \end{aligned}$$

For any $a, b \in \Omega$, let

$$P(a, b) = \frac{R_X(a)R_Y(b)}{\rho}.$$

It's clear that $\sum_a R_X(a) = \sum_b R_Y(b) = \rho$ by the fact that $D_{\mathrm{TV}}(\mu, \nu) = \max_{A\in\Omega}|\mu(A) - \nu(A)|$. Moreover, $R_X(a)R_Y(a) = 0, \forall a \in \Omega$.

Now we need check it's indeed a coupling. For a fixed $a \in \Omega$,

$$\begin{aligned} \sum_b P(a, b) &= P(a, a) + \sum_{b:b\neq a} P(a, b) \\ &= P(a, a) + \sum_{b:b\neq a} \frac{R_X(a)R_Y(b)}{\rho} \\ &= P(a, a) + \frac{R_X(a)}{\rho}(\rho - R_Y(a)) \\ &= P(a, a) + R_X(a) = \mu(a) \end{aligned}$$

Similarly, you can check that $\sum_b P(a,b) = \nu(b), \forall b \in \Omega$. Hence it's a feasible coupling. As for the optimiality,

$$\begin{aligned}\mathbf{Pr}_{(X,Y)\sim\omega}[X \neq Y] &= 1 - \sum_{a\in\Omega} P(a,a) \\ &= \sum_{a\in\Omega}\mu(a) - \sum_{a\in\Omega} P(a,a) = \rho.\end{aligned}$$

Problem 2 (Stochastic Dominance)

Let $\Omega \subseteq \mathbb{Z}$ be a finite set of integers. Let μ and ν be two distributions over Ω. We say μ is *stochastic dominance* over ν if for $X \sim \mu$, $Y \sim \nu$ and any $a \in \Omega$,

$$\mathbf{Pr}[X \geq a] \geq \mathbf{Pr}[Y \geq a].$$

We write $\mu \succeq \nu$.

1. Consider the binomial distirbution $\texttt{Binom}(n,p)$ where $X \sim \texttt{Binom}(n,p)$ satisfies for any $a = 0,1,\dots,n$, $\mathbf{Pr}[X = a] = \binom{n}{a}\cdot p^a \cdot (1-p)^{n-a}$. Prove that for any $p, q \in [0,1]$, $\mathrm{Binom}(n,p) \succeq \mathrm{Binom}(n,q)$ if and only if $p \geq q$.
Proof.
We consturct the following coupling with respect to $\mathrm{Binom}(n,p)$ and $\mathrm{Binom}(n,q)$:
1. Sample U_i uniformly at random from $[0,1]$ for any $i \in [n]$ i.i.d;
2. $X_i = 1$ iff $U_i \leq p$ and $Y_i = 1$ iff $U_i \leq q$ for any $i \in [n]$;
3. Let $X = \sum_{i=1}^n X_i$ and $Y = \sum_{i=1}^n Y_i$.
It is obvious that $X \sim \mathrm{Binom}(n,p)$ and $Y \sim \mathrm{Binom}(n,q)$ which justifies the above process is indeed a coupling. With this coupling, we know that $\{Y \geq a\} \subseteq \{X \geq a\}$ for any $a = 0,1,\dots,n$ iff $p \geq q$. Therefore, if $p \geq q$, we have $\mathrm{Binom}(n,p) \succeq \mathrm{Binom}(n,q)$.

   On the other hand, let $X \sim \mathrm{Binom}(n,p)$ and $Y \sim \mathrm{Binom}(n,q)$. If $\mathrm{Binom}(n,p) \succeq \mathrm{Binom}(n,q)$, we have $\mathbf{Pr}[X = n] \geq \mathbf{Pr}[Y = n]$ which implies that $p \geq q$.

2. A coupling ω of μ and ν is *monotone* if $\mathbf{Pr}_{(X,Y)\sim\omega}[X \geq Y] = 1$. Prove that $\mu \succeq \nu$ if and only if a monotone coupling of μ and ν exists.
Proof.
Proof of "$\Leftarrow$".Suppose ω is a monotone coupling of μ and ν, which means $\mathbf{Pr}_{(X,Y)\sim\omega}[X \geq Y] = 1$. Then

$$\begin{aligned}\mathbf{Pr}_{Y\sim\nu}[Y \geq a] &= \mathbf{Pr}_{(X,Y)\sim\omega}[Y \geq a] \\ &= \mathbf{Pr}_{(X,Y)\sim\omega}[X \geq Y \wedge Y \geq a] + \mathbf{Pr}_{(X,Y)\sim\omega}[X < Y \wedge Y \geq a] \\ &= \mathbf{Pr}_{(X,Y)\sim\omega}[X \geq Y \geq a] \\ &\leq \mathbf{Pr}_{(X,Y)\sim\omega}[X \geq a] = \mathbf{Pr}_{X\sim\mu}[X \geq a].\end{aligned}$$

Proof of "$\Rightarrow$". Define the cumulative distribution function $F_\mu(x) = \mu((-\infty, x])$ and $F_\nu(x) = \nu((-\infty, x])$. Then use these functions to construct two random variables:

$$X = F_\mu^{-1}(U), Y = F_\nu^{-1}(U),$$

where U is sampled uniform at random from $[0, 1]$ and
$F_\mu^{-1}(u) \triangleq \inf\{x \in \mathbb{R} : F_\mu(x) \geq u\}$ is a generalized inverse(similar for F_ν^{-1}).
Now we claim that $\omega = (X, Y)$ is a monotone coupling of μ and ν.
First we need to check that ω is a coupling of μ and ν which means X and Y follows μ and ν respectively.

$$\mathbf{Pr}\left[X \leq x\right] = \mathbf{Pr}\left[F_\mu^{-1}(U) \leq x\right] = \mathbf{Pr}\left[U \leq F_\mu(x)\right] = F_\mu(x) \quad \text{(similar for Y)}$$

Then we will check $\mathcal{C}$ is a monotone coupling. According to $\mu \succeq \nu$, we have $F_\mu(x) \leq F_\nu(x)$ for $x \in \mathbb{R}$. Thus

$$\mathbf{Pr}\left[X \geq Y\right] = \mathbf{Pr}\left[F_\mu^{-1}(U) \geq F_\nu^{-1}(U)\right] = 1.$$

3. Consider the Erdős–Rényi (https://en.wikipedia.org/wiki/Erd%C5%91s%E2%80%93R%C3%A9nyi_model) model $\mathcal{G}(n, p)$ for random graph. In this model, each $G \sim \mathcal{G}(n, p)$ is a simple undirected random graph with n vertices where each $\{i, j\} \in \binom{[n]}{2}$ is present with probability p independently. Prove that for any $p, q \in [0, 1]$ satisfying $p \geq q$, it holds that $\mathbf{Pr}_{G \sim \mathcal{G}(n,p)}\left[G \text{ is connected}\right] \geq \mathbf{Pr}_{H \sim \mathcal{G}(n,q)}\left[H \text{ is connected}\right]$.
   Proof
   The following coupling justifies the statement:
   1. Sample U_e uniformly at random from $[0, 1]$ for any $e \in \binom{[n]}{2}$ i.i.d;
   2. e occurs in G iff $U_e \leq p$ and e occurs in H iff $U_e \leq q$ for any $e \in \binom{[n]}{2}$.

Problem 3 (Total Variation Distance is Non-Increasing)

Let P be the transition matrix of an irreducible and aperiodic Markov chain with state space Ω. Let π be its stationary distribution. Let μ_0 be an arbitrary distribution on Ω and $\mu_t^{\mathtt{T}} = \mu_0^{\mathtt{T}} P^t$ for every $t \geq 0$. For every $t \geq 0$, let $\Delta(t) = D_{\mathtt{TV}}(\mu_t, \pi)$ be the total variation distance between μ_t and π. Prove that $\Delta(t+1) \leq \Delta(t)$ for every $t \geq 0$.
Proof
Let $X_t \sim \mu_t$ and $Y_t \sim \pi$. By coupling lemma, there exist a coupling ω of X_t and Y_t such that $\mathbf{Pr}\left[X_t \neq Y_t\right] = \Delta(t)$. Equipped with ω, we construct the coupling ω' of X_{t+1} and Y_{t+1} as follows:
1. We first sample (X_t, Y_t) from ω;
2. Next we run the Markov chain according to the transition matrix P on X_t and Y_t as follows:

- If $X_t = Y_t$, the two chains evolve synchronously;
- If $X_t \neq Y_t$, the two chains evolve independently.

Under the coupling ω',

$$\Delta(t+1) \leq \mathbf{Pr}\left[X_{t+1} \neq Y_{t+1}\right] \leq \mathbf{Pr}\left[X_t \neq Y_t\right] = \Delta(t).$$