STOCHASTIC PARTICLE SYSTEMS - EXERCISE 1: SOLUTIONS

1. For $i \in \{1, ..., N\}$, let X_i be a Poisson distributed random variable with parameter λ_i independent of X_j for $j \neq i$. Show that $\sum_{i=1}^N X_i$ is Poisson distributed with parameter $\sum_{i=1}^N \lambda_i$.

Solution: By induction, it's enough to consider the $N = 2$ case. To characterise the distribution of $X_1 + X_2$, it is enough to calculate the probabilities $P(X_1 + X_2 = k)$ for each non-negative integer k. Using independence and the fact that X_i is Poisson with parameter λ_i, we have

$$\begin{aligned}
P(X_1 + X_2 = k) &= \sum_{n=0}^{k} P(X_1 + X_2 = k, X_1 = n) \\
&= \sum_{n=0}^{k} P(X_2 = k - n, X_1 = n) \\
&= \sum_{n=0}^{k} P(X_1 = n) P(X_2 = k - n) \\
&= \sum_{n=0}^{k} \frac{\lambda_1^n e^{-\lambda_1}}{n!} \frac{\lambda_2^{k-n} e^{-\lambda_2}}{(k-n)!} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{n=0}^{k} \frac{\lambda_1^n \lambda_2^{k-n}}{n!(k-n)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{n=0}^{k} \frac{k!}{n!(k-n)!} \lambda_1^n \lambda_2^{k-n} \\
&= \frac{e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2)^k}{k!},
\end{aligned}$$

where in the last step we used the binomial theorem. Thus $X_1 + X_2$ is Poisson with parameter $\lambda_1 + \lambda_2$. $\square$

2. *The setup:* Consider the set $\mathbb{T}_N^d = (\mathbb{Z}/N\mathbb{Z})^d$, i.e. the discrete d-dimensional torus and let Ω be the set of occupation configurations on $\mathbb{T}_N^d$, i.e. $\eta \in \Omega$ is a map $\eta : \mathbb{T}_N^d \to \mathbb{N} = \{0, 1, 2, ...\}$. The interpretation is that for any $x \in \mathbb{T}_N^d$, $\eta(x)$ tells you how many particles are located at x. Consider now the probability measure μ_α on Ω defined by

$$\mu_\alpha(\eta) = \prod_{x \in \mathbb{T}_N^d} \frac{\alpha^{\eta(x)} e^{-\alpha}}{\eta(x)!}.$$

What this means is that the number of particles at each x is independent of the number of particles at the other points and it is Poisson distributed with parameter α (which is constant in x). The total number of particles is a random variable under this distribution (it is Poisson distributed according to problem 1). To relate this measure to the random walk scenario, we condition on the event that the total number of particles is K: let $S_K = \{\eta \in \Omega : \sum_x \eta(x) = K\}$ and let us define the measure $\tilde{\mu}$ on Ω by

$$\tilde{\mu}(\eta) = \frac{\mu_\alpha(\eta; \eta \in S_K)}{\mu_\alpha(S_K)}.$$

The actual problems:

a) Show that $\tilde{\mu}$ is independent of α.

b) Show that if we take K independent simple random walks $\{X^i : i = 1, ..., K\}$ on $\mathbb{T}_N^d$ and let each of their initial distributions be uniform, then $\tilde{\mu}$ is the distribution of $x \mapsto \sum_{i=1}^K \mathbf{1}_{\{X^i = x\}}$

Solutions:

a) Let us begin by calculating $\mu_\alpha(S_K)$. As noted in the setup, problem 1 implies that as a sum of Poisson random variables, $\sum_x \eta(x)$ is a poisson random variable with parameter $|\mathbb{T}_N^d|\alpha = N^d\alpha$. Thus

$$\mu_\alpha(S_K) = \frac{(\alpha N^d)^K e^{-\alpha N^d}}{K!}.$$

The probablity of a configuration $\eta_K \in S_K$ (other configurations are given zero probability under $\tilde{\mu}$) is then

$$\mu_\alpha(\eta_K) = \prod_{x \in \mathbb{T}_N^d} \frac{\alpha^{\eta_K(x)} e^{-\alpha}}{\eta_K(x)!} = \frac{\alpha^{\sum_x \eta_K(x)} e^{-N^d\alpha}}{\prod_x \eta_K(x)!} = \frac{\alpha^K e^{-\alpha N^d}}{\prod_x \eta_K(x)!}.$$

So for $\eta \in \Omega$,

$$\tilde{\mu}(\eta) = \frac{K!}{N^{Kd} \prod_x \eta(x)!} \mathbf{1}(\eta \in S_K),$$

where $\mathbf{1}(\eta \in S_K) = 1$ is $\eta \in S_K$ and 0 if $\eta \notin S_K$. This is independent of α so $\tilde{\mu}$ is independent of α.

b) Let ν be the uniform distribution on $(\mathbb{T}_N^d)^K$. We wish to show that for any $\eta \in \Omega$ and for all n,

$$P^\nu \left(\sum_{i=1}^K \mathbf{1}_{\{X_n^i = x\}} = \eta(x) \text{ for all } x \in \mathbb{T}_N^d \right) = \tilde{\mu}(\eta).$$

Here P^ν is the distribution of the collection of random walks with initial distribution ν. Let us begin by checking that the uniform measure on the starting points is indeed an invariant measure for the K independent random walks. Let us write $\mathbf{X} = (X^i)_{i=1}^K$. By independence, X is a Markov chain and it's transition probability is just given by the product of the transition probabilites of the random walks. The simple random walk is symmetric, i.e. it's transition probability p satisfies $p(x, y) = p(y, x)$. Thus $\mathbf{X}$ has this property as well. Symmetry and ν being uniform imply that

$$\sum_{\mathbf{x}_0 \in (\mathbb{T}_N^d)^K} \nu(\mathbf{x}_0) P(\mathbf{X}_1 = \mathbf{y} | \mathbf{X}_0 = \mathbf{x}_0) = \sum_{\mathbf{x}_0 \in (\mathbb{T}_N^d)^K} \nu(\mathbf{y}) P(\mathbf{X}_1 = \mathbf{y} | \mathbf{X}_0 = \mathbf{x}_0)$$
$$= \nu(\mathbf{y}) \sum_{\mathbf{x}_0 \in (\mathbb{T}_N^d)^K} P(\mathbf{X}_1 = \mathbf{x}_0 | \mathbf{X}_0 = \mathbf{y})$$
$$= \nu(\mathbf{y}).$$

So ν is invariant. A simple induction shows that the distribution of $\mathbf{X}_n$ is ν for every n so it is enough to consider the case $n = 0$. The question of the distribution of $\sum_{i=1}^K \mathbf{1}_{\{X_n^i = x\}}$ now becomes a problem in basic probability: the fact that ν is uniform implies that our problem is equivalent to the following scenario. We have K balls and at each point in $\mathbb{T}_N^d$ we have a slot. We place each ball into a slot uniformly at random and ask how many balls there are in each slot. The distribution of this problem is just the multinomial distribution with K independent trials, N^d possible outcomes and the probability of each outcome being N^{-d}. Explicitly, if $\eta \in S_K$, then plugging the data into the multinomial distribution gives

$$P^\nu \left(\sum_{i=1}^K \mathbf{1}_{\{X_n^i = x\}} = \eta(x) \text{ for all } x \in \mathbb{T}_N^d \right) = \frac{K!}{\prod_x \eta(x)!} \prod_x (N^{-d})^{\eta(x)}$$
$$= \frac{K!}{N^{Kd} \prod_x \eta(x)!}.$$

As the probablity is zero for $\eta \notin S_K$, we see that this indeed agrees with $\tilde{\mu}$.$\square$

3. Let (Y, ρ) be a metric space, $\mathcal{B}(Y)$ be the σ-algebra of the Borel sets of Y and $\mathcal{M}_1(Y)$ the set of probability measures on $(Y, \mathcal{B}(Y))$. For $A \subset Y$ and $\epsilon > 0$, let

$$A^{(\epsilon)} = \{x \in Y : \rho(x, y) < \epsilon \text{ for some } y \in Y\}.$$

Define

$$r(\mu, \nu) = \inf\{\epsilon > 0 : \nu(F) \leq \mu(F^{(\epsilon)}) + \epsilon \text{ for every closed set } F \subset Y\}.$$

Show that r satisfies the triangle inequality, i.e. that for each $\lambda, \mu, \nu \in \mathcal{M}_1(Y)$, $r(\lambda, \nu) \leq r(\lambda, \mu) + r(\mu, \nu)$.

Solution: Let $\lambda, \mu, \nu \in \mathcal{M}_1(Y)$ and $\epsilon, \epsilon' > 0$ so that $r(\lambda, \mu) < \epsilon$ and $r(\mu, \nu) < \epsilon'$. Thus for any closed set F,

$$\nu(F) \leq \mu(F^{(\epsilon)}) + \epsilon.$$

and

$$\mu(F) \leq \lambda(F^{(\epsilon')}) + \epsilon'.$$

We note that certainly $\mu(F^{(\epsilon)}) \leq \mu(\overline{F^{(\epsilon)}})$ ($\overline{A}$ is the closure of A) so composing the two relations we have

$$\nu(F) \le \lambda\left(\overline{F^{(\epsilon)}}^{(\epsilon')}\right) + \epsilon + \epsilon'.$$

Let $x \in \overline{F^{(\epsilon)}}^{(\epsilon')}$. Then there is some $y \in \overline{F^{(\epsilon)}}$ such that $\rho(x, y) < \epsilon'$. As $y \in \overline{F^{(\epsilon)}}$, we may pick some sequence of points $y_n \in F^{(\epsilon)}$ so that $y_n \to y$ so we can find a point $y' \in F^{(\epsilon)}$ so that $\rho(x, y') < \epsilon'$. Since $y' \in F^{(\epsilon)}$, we can find a point $z \in F$ so that $\rho(y', z) < \epsilon$. Thus we have found a point $z \in F$ so that $\rho(x, z) \le \rho(x, y') + \rho(y', z) < \epsilon' + \epsilon$. We conclude that $x \in F^{(\epsilon+\epsilon')}$ so that $\overline{F^{(\epsilon)}}^{(\epsilon')} \subset F^{(\epsilon+\epsilon')}$ and

$$\nu(F) \le \lambda(F^{(\epsilon+\epsilon')}) + \epsilon + \epsilon'.$$

This implies that $r(\lambda, \nu) \le \epsilon + \epsilon'$. Letting $\epsilon \to r(\lambda, \mu)$ and $\epsilon' \to r(\mu, \nu)$ we get the triangle inequality.

Remarks: As noted in the lectures, r is indeed a metric. To see that r is symmetric, pick two probability measures μ and ν and let $\alpha, \beta > 0$ be such that $\mu(F) \le \nu(F^{(\alpha)}) + \beta$ for all closed sets F. Now if we consider the closed set $\tilde{F} = Y \setminus F^{(\alpha)}$, this relation implies that $\mu(F^{(\alpha)}) = 1 - \mu(\tilde{F}) \ge 1 - \nu(\tilde{F}^{(\alpha)}) - \beta = \nu(Y \setminus \tilde{F}^{(\alpha)}) - \beta \ge \nu(F) - \beta$, where the last step requires checking that $F \subset Y \setminus \tilde{F}^{(\alpha)}$. We conclude that for any $\epsilon > 0$, $\mu(F) \le \nu(F^{(\epsilon)}) + \epsilon$ implies that $\nu(F) \le \mu(F^{(\epsilon)}) + \epsilon$. It follows from this that $r(\mu, \nu) = r(\nu, \mu)$. For positive definiteness, we note that $r(\mu, \nu) = 0$ and some elementary measure theory imply that $\mu(F) = \nu(F)$ for every closed set. One can show (see for example Billingsley's Convergence of probability measures) that probability measures on metric spaces with Borel σ-algebras are regular, i.e. the measure of any Borel set can be approximated by the measure of closed sets. From this it follows that $\mu(A) = \nu(A)$ also for any Borel set and $\mu = \nu$.

The point of this metric is that we can define weak convergence of measures in terms of it instead of integrals of continuous functions. Recall that the sequence of measures μ_n is said to converge weakly to a measure μ if

$$\int f d\mu_n \to \int f d\mu$$

for each bounded continuous function $f : Y \to \mathbb{R}$. This is equivalent to $r(\mu_n, \mu) \to 0$. This is a result of the Portmanteau theorem. This theorem gives a few characterisations for weak convergence. One of the characterisations is that the sequence (μ_n) converges to μ weakly if $\limsup_n \mu_n(F) \le \mu(F)$ for all closed sets F. The proof is not especially hard, but requires some measure theory. For a proof, see Billingsley's Convergence of probability measures or Kallenberg's Foundations of modern probability. The benefit of being able to metrize weak convergence is that one can then use the machinery of metric spaces and use concepts such as completeness.

4. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra and $X : \Omega \to \mathbb{R}$ some random variable. Recall that the conditional expectation of X given $\mathcal{G}$ defined to be any $\mathcal{G}$ measurable random variable $\tilde{X}$ satisfying

$$\int_G \tilde{X}(\omega) dP(\omega) = \int_G X(\omega) dP(\omega)$$

for all $G \in \mathcal{G}$. Moreover, the conditional expectation is almost surely defined uniquely by this condition and one usually writes $\tilde{X} = E(X|\mathcal{G})$. When $\mathcal{G}$ is the smallest σ-algebra that makes a mapping $Y : \Omega \to \mathbb{R}$ measurable (i.e. $\mathcal{G} = \sigma(Y)$ is the σ-algebra generated by Y), one often writes $E(X|\mathcal{G}) = E(X|Y)$.

Consider now the special case that Y takes only finitely many values. Show that there is some function $f : \mathbb{R} \to \mathbb{R}$ so that $E(X|Y) = f(Y)$, i.e. that for every $\omega \in \Omega$, $E(X|Y)(\omega) = f(Y(\omega))$. Note that f can depend on the range of Y and naturally it will depend on X, but it does not depend on the randomness ω.

Solution: Let y be one of the values Y takes, i.e. $y \in Y(\Omega)$. Then intuitively one would expect that if such a function f exists, it should satisfy $f(y) = E(X|Y = y)$, where $E(X|Y = y)$ can be defined in terms of elementary probability since we are dealing with a random variable taking only finitely many values:

$$E(X|Y = y) = \frac{E(X1_{\{Y=y\}})}{P(Y = y)}.$$

This should hold for all y, so a natural candidate for $E(X|Y)$ would be

$$Z = \sum_{y \in Y(\Omega)} \frac{E(X1_{\{Y=y\}})}{P(Y = y)} \mathbf{1}_{\{Y=y\}}.$$

To check that indeed $Z = E(X|Y)$, we must check that for any set $A \in \sigma(Y)$, $E(X\mathbf{1}_A) = E(Z\mathbf{1}_A)$. One can then check that $\sigma(Y)$ is in our case a finite collection of sets which are unions of sets of the form $\{\omega : Y(\omega) = y\}$ for some $y \in Y(\Omega)$. Thus it is enough to check the property for sets of this form. So let $y \in Y(\Omega)$. Since $\mathbf{1}_{Y=y}\mathbf{1}_{Y=x} = 0$ if $x \neq y$, we have

$$\begin{aligned} E(Z\mathbf{1}_{\{Y=y\}}) &= E\left(\frac{E(X1_{\{Y=y\}})}{P(Y = y)}\mathbf{1}_{Y=y}\right) \\ &= E(X1_{\{Y=y\}}), \end{aligned}$$

so $E(X|Y) = Z = f(Y)$ for the function $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \sum_{y \in Y(\Omega)} \frac{E(X\mathbf{1}_{Y=y})}{P(Y = y)} \delta_{x,y}.$$

here $\delta_{i,j}$ is the Kroenecker δ: it is one when $i = j$ and zero otherwise. □

5. Let S_n be a simple random walk on $\mathbb{Z}$ with $S_0 = 0$. Show that $X_n = \max\{S_m : 0 \leq m \leq n\}$ is not a Markov chain.

Solution: To be a Markov chain, one must have $P(X_{n+1} = x_{n+1}|X_n = x_n, ..., X_0 = x_0) = P(X_{n+1} = x_{n+1}|X_n = x_n)$ for all n and for all $\{x_n\}$ for which $P(X_n = x_n, ..., X_0 = x_0) > 0$. So let us consider

$$P(X_4 = 2|X_3 = 1, X_2 = 1, X_1 = 1, X_0 = 0) \text{ and } P(X_4 = 2|X_3 = 1, X_2 = 0, X_1 = 0, X_0 = 0).$$

Since we are dealing with a simple random walk, the case $(X_0, X_1, X_2, X_3) = (0, 0, 0, 1)$ can correspond only to a single path: $(S_0, S_1, S_2, S_3) = (0, -1, 0, 1)$. Whereas the case $(X_0, X_1, X_2, X_3) = (0, 1, 1, 1)$ can correspond to two paths: $(S_0, S_1, S_2, S_3) = (0, 1, 0, \pm 1)$. Since each path of length n has a probability 2^{-n} for the simple random walk, we see that

$$P^0(X_3 = 1, X_2 = 1, X_1 = 1, X_0 = 0) = 2^{-2}$$

(this case corresponded to two paths) and

$$P^0(X_3 = 1, X_2 = 0, X_1 = 0, X_0 = 0) = 2^{-3}$$

(this corresponded to only one path). On the other hand there is only one path for the event $(X_0, X_1, X_2, X_3, X_4) = (0, 1, 1, 1, 2)$: $(S_0, S_1, S_2, S_3, S_4) = (0, 1, 0, 1, 2)$. Similarly for the case $(X_0, X_1, X_2, X_3, X_4) = (0, 0, 0, 1, 2)$ there is only the path $(S_0, S_1, S_2, S_3, S_4) = (0, -1, 0, 1, 2)$. Hence each of the events has probability 2^{-4} and we conclude that

$$P(X_4 = 2 | X_3 = 1, X_2 = 1, X_1 = 1, X_0 = 0) = \frac{1}{4}$$

and

$$P(X_4 = 2 | X_3 = 1, X_2 = 0, X_1 = 0, X_0 = 0) = \frac{1}{2}.$$

Since these two probabilities don't agree, they can't both be equal to $P(X_4 = 2 | X_3 = 1)$ and (X_n) can't be a Markov chain. $\square$