2.2.2. *Esimerkkejä korttipakan sekoittamisesta*

Korttipakkaa voidaan sekoittaa monilla tavoilla. Ideana on toistaa jotakin satunnaista operaatiota, joka vaikuttaa sekoittavasti korttien järjestykseen. Jos pakassa on n korttia, korttipakan järjestystä kuvaa permutaatio $\sigma \in S_n$ — tällöin j:s kortti päältä lukien on $\sigma(j)$. Satunnainen sekoitusoperaatio määrittelee siirtymätodennäköisyysmatriisin P, se kertoo millä todennäköisyydellä alkuperäisestä korttien järjestyksestä saadaan operaation jälkeen jokin toinen järjestys. Sekoittamista kuvaa siis Markov-ketju $X = (X(t))_{t\in\mathbb{Z}_{\geq 0}}$ tila-avaruudella $\mathcal{S} = S_n$, siten, että pakan järjestys aluksi on $X(0)$, ja järjestys t:n sekoitusoperaation jälkeen on $X(t)$.

Jotta korttipakan toistuva sekoitus tuottaisi hyvin sekoitetun korttipakan, tulee tilojen $X(t)$ supeta kohti tasaista jakaumaa joukolla S_n, kun $t \to \infty$. Kuten edellä nähtiin, välttämätön ehto tälle on, että tasainen jakauma $\underline{\pi}$, $\pi_\sigma = \frac{1}{n!}$, on stationaarinen jakauma sekoitusoperaation siirtymätodennäköisyyksille P.

Esimerkki V.17 (Korttipakan ylhäältä keskelle -sekoitus)**.** Alkeellinen tapa sekoittaa korttipakkaa olisi ottaa aina päällimmäinen kortti, ja työntää se sitten umpimähkäiseen paikkaan pakassa. Jos järjestys ennen sekoittamista on $\sigma \in S_n$, niin järjestys τ yhden sekoitusoperaation jälkeen on muotoa $\tau = \sigma \circ \lambda_k$, missä λ_k on k-sykli

$$\lambda_k(j) = \begin{cases} j+1 & \text{jos } j < k \\ 1 & \text{jos } j = k \\ j & \text{jos } j > k \end{cases}.$$

Tällaista sekoittamista vastaavan Markov-ketjun siirtymätodennäköisyydet ovat

$$P_{\sigma,\tau} = \begin{cases} \frac{1}{n} & \text{jos } \tau = \sigma \circ \lambda_k \text{ jollakin } k \\ 0 & \text{muutoin} \end{cases}.$$

Tasaisen jakauman $\underline{\pi}$ sekoittaminen tuottaa

$$(\underline{\pi} P)_\tau = \sum_{\sigma \in S_n} \pi_\sigma P_{\sigma,\tau} = \sum_{\substack{\sigma \in S_n \\ \exists k: \tau = \sigma \circ \lambda_k}} \frac{1}{n!} \frac{1}{n}$$

Summassa on nyt n yhtäsuurta termiä, mahdollisista k:n arvoista $1, 2, 3, \ldots, n$, joten tosiaankin $(\underline{\pi} P)_\tau = \frac{1}{n!} = \pi_\tau$ eli tasainen jakauma $\underline{\pi}$ on stationaarinen.

Esimerkki V.18 (Korttipakan limittäissekoitus)**.** Tavallisesti korttipakkaa sekoitetaan siten, että pakka puolitetaan (ainakin suunnilleen) kahteen osaan, ja osat sitten yhdistetään toisiinsa limittäin (yleensä antamalla korttien tipahdella eri osista vuoroillaan yhteiseen pinoon, tai vähemmän sorminäppäryyttä vaativasti vaikkapa työntämällä osia toisiaan vasten). Yksilöllisissäkin sekoitustyyleissä on toki eroja, mutta esitämme tässä mallin, joka tavoittaa kohtuullisen hyvin useita relevantteja korttipakan tavallisen sekoittamisen ominaisuuksia.

Limittäissekoitusoperaatiota voidaan mallintaa seuraavasti. Valitaan binomijakautunut $B \sim \mathrm{Bin}(n, \frac{1}{2})$. Päällimmäiset B korttia ja alimmaiset $n - B$ korttia muodostavat pakan puolituksen kahteen osaan. Sekoitettu järjestys τ konstruoidaan valitsemalla toistuvasti järjestyksen τ seuraavaksi kortiksi päällimmäinen jomman kumman osan jäljellä olevista korteista todennäköisyyksillä, jotka ovat suhteessa osien jäljellä olevien korttien lukumääriin. Yhtäpitävästi sekoitettu järjestys voitaisiin toteuttaa puolituksen jälkeen valitsemalla τ tasaisesti sellaisten järjestysten joukosta, joille $\tau^{-1}(\sigma(j)) < \tau^{-1}(\sigma(j'))$ aina kun $j < j' \leq B$ tai $B < j < j'$ — tämä vaatimus sanoo, että puolikkaat asetetaan keskenään limittäin. Pienen sievennyksen jälkeen nähdään, että tämän sekoitusoperaation siirtymätodennäköisyysmatriisi on

$$P_{\sigma,\tau} = \begin{cases} 2^{-n} & \text{jos jollakin } b \text{ pätee } \tau^{-1}(\sigma(j)) < \tau^{-1}(\sigma(j')) \text{ kun } j < j' \leq b \text{ tai } b < j < j' \\ 0 & \text{muutoin.} \end{cases}$$

2.2.3. *Detailed balance*

Seuraava riittävä ehto stationaarisuudelle on, silloin kun se on voimassa, helpompi tarkistaa kuin stationaarisuus suoraan.

Lemma V.19. *Jos $P \in \mathbb{R}^{\mathcal{S}\times\mathcal{S}}$ on stokastinen matriisi, ja $\underline{\pi} \in \mathbb{R}^{\mathcal{S}}$ on rivivektori, jolle pätee "detailed balance"-ehto*

$$\pi_x\, P_{x,y} = \pi_y\, P_{y,x} \qquad \forall x, y \in \mathcal{S}, \tag{V.12}$$

niin silloin $\underline{\pi} P = \underline{\pi}$.

Todistus. Summaamalla yhtälöä (V.12) indeksin $x \in \mathcal{S}$ yli, vasen puoli on $(\underline{\pi} P)_y$, ja oikea puoli taas $\sum_{x\in\mathcal{S}} \pi_y\, P_{y,x}$. Oikea puoli yksinkertaistuu muotoon π_y käyttäen ominaisuutta (V.10). □

Huomautus V.20. Esimerkkien V.14 ja V.15 jakaumien stationaarisuus on tosiaan suoraviivaisinta tarkistaa käyttäen Lemman V.19 ehtoa.

Esimerkin V.17 stationaarinen jakauma ei toteuta ehtoa (V.12). Erityisesti "detailed balance" ehto ei ole välttämätön statinaarisuudelle.

2.3. Markov-ketjun redusoitumattomuus ja jaksottomuus

Tavoitteemme on antaa käyttökelpoiset riittävät ehdot sille, että Markov-ketju konvergoi, mistä tahansa alkutilasta lähtien, kohti yksikäsitteistä stationaarista jakaumaa. Kriteerit koskevat tilojen välisiä siirtymätodennäköisyyksiä.

Oletamme siis annetuksi äärellisen tila-avaruuden $\mathcal{S}$ sekä stokastisen matriisin $P \in \mathbb{R}^{\mathcal{S}\times\mathcal{S}}$ (vaikka alkujakaumaa ei spesifioitaisi, käytämme tästä termiä Markov-ketju). Luokittelemme seuraavaksi tiloja $x \in \mathcal{S}$ sekä osajoukkoja $S \subset \mathcal{S}$ niitä koskevien siirtymätodennäköisyyksien kvalitatiivisten ominaisuuksien perusteella.

Määritelmä V.21. Sanomme, että:

- Tilasta $x \in \mathcal{S}$ on *pääsy* tilaan $y \in \mathcal{S}$, jos jollakin $t \geq 0$ pätee $(P^t)_{x,y} > 0$. Merkitsemme $x \rightsquigarrow y$.
- Markov-ketju on *redusoitumaton*, jos sen kaikista tiloista on pääsy toisiinsa, eli kaikilla $x, y \in C$ pätee $x \rightsquigarrow y$.
- Tilan $x \in \mathcal{S}$ *jakso* (tai *periodi*) on joukon $\mathcal{T}(x) = \big\{ t \in \mathbb{Z}_{>0} \,\big|\, (P^t)_{x,x} > 0 \big\}$ suurin yhteinen tekijä $\gcd(\mathcal{T}(x))$.
- Markov-ketju on *jaksoton*, jos sen kaikkien tilojen jakso on 1.

Periodi on sama kaikille keskenään kommunikoiville tiloille.

Lemma V.22. *Jos $x \leftrightsquigarrow y$, niin* $\gcd(\mathcal{T}(x)) = \gcd(\mathcal{T}(y))$.

Todistus. Olkoot $x, y \in \mathcal{S}$ kaksi eri tilaa, $x \neq y$, jotka kommunikoivat keskenään, $x \leftrightsquigarrow y$. Silloin on olemassa positiiviset kokonaisluvut $r, s \in \mathbb{Z}_{>0}$ siten, että $(P^r)_{x,y} > 0$ ja $(P^s)_{y,x} > 0$. Asetetaam $m = r + s$. Silloin $(P^m)_{x,x} > 0$ ja $(P^m)_{y,y} > 0$, joten $m \in \mathcal{T}(x) \cap \mathcal{T}(y)$. Lisäksi selvästi $\mathcal{T}(x) \subset \mathcal{T}(y) - m$, joten jokainen joukon $\mathcal{T}(y)$ alkioiden tekijä on myös joukon $\mathcal{T}(x)$ alkioiden tekijä, ja siis $\gcd(\mathcal{T}(x)) \geq \gcd(\mathcal{T}(y))$. Symmetrisesti päätellään $\gcd(\mathcal{T}(x)) \leq \gcd(\mathcal{T}(y))$. □

Seuraus V.23. *Redusoitumattoman Markov-ketjun kaikkien tilojen jakso on sama.*

Huomautus V.24. Seurauksen V.23 perusteella redusoitumaton Markov-ketju on jaksoton, jos on olemassa ainakin yksi tila, jonka jakso on 1.

Propositio V.25. *Jos stokastinen matriisi P on redusoitumaton ja jaksoton, niin on olemassa r siten, että $(P^r)_{x,y} > 0$ kaikilla $x, y \in \mathcal{S}$.*

Todistus. Käytämme lukuteoreettista tulosta, jonka mukaan luonnollisten lukujen joukon sellainen osajoukko, joka on yhteenlaskun suhteen suljettu ja jonka suurin yhteinen tekijä on 1, sisältää äärellisen montaa poikkeusta lukuunottamatta kaikki luonnolliset luvut.

Oletetaan, että P on redusoitumaton ja jaksoton. Olkoon $x \in \mathcal{S}$, ja muistutetaan määritelmä $\mathcal{T}(x) = \{t \in \mathbb{Z}_{>0} \,|\, (P^t)_{x,x} > 0\}$. Jos $t, s \in \mathcal{T}(x)$, niin selvästi myös $t + s \in \mathcal{T}(x)$, eli $\mathcal{T}(x)$ on yhteenlaskun suhteen suljettu. Jaksottomuuden perusteella $\gcd(\mathcal{T}(x)) = 1$. Päättelemme, että on olemassa luku R_x siten, että kaikilla $r \geq R_x$ pätee $r \in \mathcal{T}(x)$. Jos $y \in \mathcal{S}$ on jokin toinen tila, niin redusoitumattomuuden perusteella $(P^t)_{x,y} > 0$ jollakin t. Silloin myös $(P^{r+t})_{x,y} > 0$ kaikilla $r \in \mathcal{T}(x)$. Siis on olemassa luku $R_{x,y}$ (esimerkiksi $t+R_x$) siten, että kaikilla $r \geq R_x$ pätee $(P^r)_{x,y} > 0$. Nyt kun

$$r \geq \max\{R_{x,y} \,|\, x, y \in \mathcal{S}\}$$

pätee $(P^r)_{x,y} > 0$ kaikilla $x, y \in \mathcal{S}$. □

Esimerkki V.26 (Laiska Markov-ketju)**.** Jos P on stokastinen matriisi, ja $\varepsilon \in (0, 1)$, niin myös $P' = (1-\varepsilon)P + \varepsilon I$ on stokastinen matriisi. Stokastisia matriiseja P ja P' vastaavat Markov-ketjut käyttäytyvät joitakin tarkoituksia varten riittävän samalla tavalla — esimerkiksi jokainen stationaarinen jakauma P':lle on stationaarinen myös P:lle. Matriisin P' määrittelemää Markov-ketjua sanotaan *laiskaksi* (verrattuna P:n määrittelemään): sen määrittelemä prosessi saadaan seuraamalla P:n määrittelemän prosessin tilojen jonoa siten, että jokaisella ajanhetkellä on lisäksi todennäköisyys ε olla etenemättä (kts. Tehtävä V.2).

Laiska Markov-ketju on aina jaksoton, koska $P_{x,x} \geq \varepsilon > 0$ kaikille tiloille x.

Tehtävä V.2. Olkoon $X = (X(t))_{t \in \mathbb{Z}_{\geq 0}}$ Markov-ketju, jonka siirtymätodennäköisyysmatriisi on P, ja olkoon $(\xi_j)_{j \in \mathbb{Z}_{\geq 0}}$ jono riippumattomia Bernoulli$(1-\varepsilon)$-jakautuneita satunnaismuuttujia, jotka ovat riippumattomia prosessista X. Asetetaan $S(t) = \sum_{j=1}^{t} \xi_j$, ja

$$X'(t) = X(S(t)).$$

Osoitettava, että $X' = (X'(t))_{t \in \mathbb{Z}_{\geq 0}}$ on Markov-prosessi siirtymätodennäköisyysmatriisilla $P' = (1-\varepsilon)P + \varepsilon I$.

2.4. Stationaarisen jakauman olemassaolo

Kaikilla äärellistilaisilla Markov-ketjuilla on stationaarisia jakaumia. Tämä voidaan periaatteessa todistaa suoralla lineaarialgebralla. Luonnostelemme myös probabilistisen todistuksen, mutta jätämme sen yksityiskohdat harjoitustehtäväksi.

Propositio V.27. *Olkoon $P \in \mathbb{R}^{\mathcal{S} \times \mathcal{S}}$ stokastinen matriisi äärellisellä tila-avaruudella $\mathcal{S}$. Silloin on olemassa stationaarinen jakauma $\underline{\pi} = (\pi_x)_{x \in \mathcal{S}}$, eli rivivektori $\underline{\pi} \in \mathbb{R}^{\mathcal{S}}$, jolle pätee*

$$\underline{\pi} = \underline{\pi}\, P \qquad \textit{eli} \qquad \pi_y = \sum_{x \in \mathcal{S}} \pi_x \, P_{x,y} \quad \forall y \in \mathcal{S}$$

ja

$$\pi_x \geq 0 \quad \forall x \in \mathcal{S}, \qquad \sum_{x \in \mathcal{S}} \pi_x = 1.$$

2.4.1. *Probabilistinen todistus stationaarisen jakauman olemassaololle*

Ylläoleva lineaarialgebrallinen todistus on lyhyt ja elegantti, mutta seuraava probabilistinen todistus antaa paremman intuition Markov-ketjuista. Se myös yleistyisi numeroituvasti äärettömän tila-avaruuden tapaukseen stationaarisen mitan (ei välttämättä todennäköisyysmitan) olemassaolotodistukseksi.

Tarkastelemme satunnaista ajanhetkeä T_x, jolla prosessi ensimmäistä kertaa alun jälkeen on tilassa $x \in \mathcal{S}$, eli

$$T_x = \min \big\{ t > 0 \; \big| \; X(t) = x \big\} .$$

Sanomme, että tila x on *palautuva*, jos alkujakaumalla $X(0) = x$ paluuhetki T_x on melkein varmasti äärellinen, eli

$$\mathsf{P}\big[T_x < \infty \; \big| \; X(0) = x \big] = 1.$$

Tehtävä V.3. Osoita, että Markov-ketjulla on ainakin yksi palautuva tila.

Tehtävä V.4. Osoita, että jos x on palautuva, niin $\mathsf{E}\big[T_x \, \big| \, X(0) = x\big] < \infty$.

Tehtävä V.5. Olkoon $X = (X(t))_{t \in \mathbb{Z}_{\geq 0}}$ Markov-ketju tila-avaruudella $\mathcal{S}$ ja $P \in \mathbb{R}^{\mathcal{S} \times \mathcal{S}}$ sen siirtymätodennäköisyysmatriisi. Olkoon x jokin palautuva tila. Määritellään kaikilla $y \in \mathcal{S}$

$$f(y) = \sum_{t=1}^{\infty} \mathsf{P}\Big[X(t) = y \text{ ja } t \leq T_x \, \Big| \, X(0) = x \Big].$$

(a) Osoita, että kaikilla $z \in \mathcal{S}$ pätee $\sum_{y \in \mathcal{S}} f(y) P_{y,z} = f(z)$.
(b) Päättele, että kaava

$$\pi_y = \frac{f(y)}{\mathsf{E}\big[T_x \, \big| \, X(0) = x\big]}$$

määrittelee Markov-ketjun stationaarisen jakauman $\underline{\pi} = (\pi_y)_{y \in \mathcal{S}}$.

Edellinen tehtävä siis tarjoaa vaihtoehtoisen todistuksen Propositiolle V.27, ja antaa lisäksi konkreettisen kuvailun stationaariselle jakaumalle: keskimääräinen osuus ajasta, jonka Markov-ketju viettää eri tiloissa johonkin palautuvaan tilaan tekemienstä peräkkäisten vierailujen välillä.

2.5. Suppeneminen kohti stationaarista jakaumaa

Stationaarisia mittoja on Proposition V.27 perusteella olemassa.

Lause V.28. *Olkoon $P = (P_{x,y})_{x,y \in \mathcal{S}}$ redusoitumaton, aperiodinen Markov-siirtymätodennäköisyysmatriisi äärellisellä tila-avaruudella $\mathcal{S}$, ja olkoon $\underline{\pi}$ sen stationaarinen mitta. Jos $X = (X(t))_{t \in \mathbb{Z}_{\geq 0}}$ on Markov-ketju, jonka siirtymätodennäköisyysmatriisi on P, niin pätee*

$$\mathsf{P}\big[X(t) = x \big] \underset{t \to \infty}{\longrightarrow} \pi_x \qquad \text{kaikilla } x \in \mathcal{S}.$$

Seurauksena saamme kaksi tärkeää johtopäätöstä redusoitumattomista Markov-ketjuista.

Seuraus V.29. *Redusoitumattomalla aperiodisella Markov-ketjulla on yksikäsitteinen stationaarinen jakauma $\underline{\pi}$. Mistä tahansa alkutilasta tai alkujakaumasta $\underline{\iota}$ lähtien, tällaisen Markov ketjun tilat $X(t)$ suppenevat, kun $t \to \infty$, kohti tätä stationaarista jakaumaa $\underline{\pi}$.*

Käytämme todistuksessa seuraavaa aputulosta.

Lemma V.30. *Redusoitumattomalle äärellistilaiselle Markov ketjulle $X = (X(t))_{t \in \mathbb{Z}_{\geq 0}}$ äärellisellä tila-avaruudella $\mathcal{S}$ pätee seuraavaa: mihin tahansa tilaan $x \in \mathcal{S}$ ensimmäisen saapumisen hetki $T_x = \inf \big\{ t > 0 \;\big|\; X(t) = x \big\}$ on odotusarvoltaan äärellinen, ja erityisesti melkein varmasti äärellinen.*

Todistus. Redusoitumattomuuden perusteella kaikilla $y \in \mathcal{S}$ on olemassa $r_y \geq 0$ siten, että $(P^r)_{y,x} > 0$. Seurauksena on olemassa $R \geq 0$ (esim. $R = \max\{r_y | y \in \mathcal{S}\}$) ja $q > 0$ (esim. $q = \min\{(P^{r_y})_{y,x} \,|\, y \in \mathcal{S}\}$) siten, että $\mathsf{P}[T_x > R \,|\, X(0) = y] \leq 1 - q$. Soveltamalla tätä arviota toistuvasti m kertaa, saamme $\mathsf{P}[T_x > mR \,|\, X(0) = y] \leq (1-q)^m$. Väitteet seuraavat tästä. □

Lauseen V.28 todistus. Todistamme väitteen koplaamalla sopivasti kaksi kopiota alkuperäisestä Markov ketjusta keskenään.

Olkoon $\underline{\pi}$ (jokin) stationaarinen jakauma ja $\underline{\iota}$ jokin jakauma joukolla $\mathcal{S}$. Olkoot prosessit $X = (X(t))_{t \in \mathbb{Z}_{\geq 0}}$ ja $Z = (Z(t))_{t \in \mathbb{Z}_{\geq 0}}$ riippumattomia Markov-ketjuja samalla siirtymätodennäköisyysmatriisilla P, ja vastaavilla alkujakaumilla $\underline{\iota}$ ja $\underline{\pi}$, eli

$$\mathsf{P}[X(0) = x] = \iota_x \qquad \text{ja} \qquad \mathsf{P}[Z(0) = x] = \pi_x.$$

Stationaarisuuden perusteella $\mathsf{P}[Z(t) = x] = \pi_x$, kaikilla t. Väitteemme on, että $\mathsf{P}[X(t) = x] \to \pi_x$ kun $t \to \infty$.

Parit $(X(t), Z(t))$, $t \in \mathbb{Z}_{\geq 0}$, muodostavat Markov-ketjun tila-avaruudella $\mathcal{S} \times \mathcal{S}$. Sen siirtymätodennäköisyydet ovat riippumattomuuden perusteella tulomuotoa

$$P_{(x,z),(y,w)} \;=\; \mathsf{P}\big[(X(t+1), Z(t+1)) = (y,w) \;\big|\; (X(t), Z(t)) = (x,z)\big] \;=\; P_{x,y}\, P_{z,w},$$

ja alkujakauma on

$$\mathsf{P}\big[(X(0), Z(0)) = (x,z)\big] \;=\; \iota_x\, \pi_z.$$

Propositiosta V.25 ja siirtymätodennäköisyyksien tulomuodosta nähdään, että tämä Markov-ketju $(X(t), Z(t))_{t \in \mathbb{Z}_{\geq 0}}$ on redusoitumaton.

Kiinnitetään nyt jokin tila $b \in \mathcal{S}$, ja merkitään

$$T = \inf \big\{ t \geq 0 \;\big|\; X(t) = b,\; Z(t) = b \big\}$$

ensimmäistä (satunnaista) ajanhetkeä, jolloin molemmat prosessit ovat tilassa b. Lemman V.30 perusteella $T < \infty$ melkein varmasti.

Idea on nyt huomata, että jos prosessit X ja Z ovat jo olleet samaan aikaan samassa tilassa b, on myös prosessin X jakauma helppo kontrolloida. Täsmällisemmin, laskemme

$$\begin{aligned}
\mathsf{P}\big[X(t) = x\big] &= \mathsf{P}\big[X(t) = x \text{ ja } t \geq T\big] + \mathsf{P}\big[X(t) = x \text{ ja } t < T\big] \\
&= \mathsf{P}\big[Z(t) = x \text{ ja } t \geq T\big] + \mathsf{P}\big[X(t) = x \text{ ja } t < T\big] \\
&= \mathsf{P}\big[Z(t) = x\big] - \mathsf{P}\big[Z(t) = x \text{ ja } t < T\big] + \mathsf{P}\big[X(t) = x \text{ ja } t < T\big].
\end{aligned}$$

Kaksi viimeistä termiä menevät selvästi nollaan, koska $\mathsf{P}\big[t < T\big] \to 0$, kun $t \to \infty$. Saamme halutun johtopäätöksen

$$\mathsf{P}\big[X(t) = x\big] \underset{t \to \infty}{\longrightarrow} \pi_x.$$

□