Σύντομες σημειώσεις σχετικά με

Κατασκευή και ιδιότητες αλυσίδων Markov με συνεχή χρόνο[1]

για χρήση στο μάθημα «Στοχαστικές Διαδικασίες»

Τακης Κωνσταντοπουλος

20/11/04

Όλες οι ασκήσεις που διατυπώνονται στο κείμενο αυτό πρέπει να θεωρηθούν σαν συνέχεια της θεωρίας, ακριβώς όπως και στο προηγούμενο φυλλάδιο που σας έδωσα. Η λύση τους αποτελεί αναπόσπαστο μέρος της κατανόησης του αντικειμένου. Σε αντίθεση με το προηγούμενο φυλλάδιο, δεν υποχρεούστε να τις παραδώσετε· όμως λάβετε υπ' όψη σας ότι θα σας βοηθήσουν στο τελικό διαγώνισμα.

1 Αμνησία σε διακριτό χρόνο

Έστω K μια τυχαία μεταβλητή με τιμές στο σύνολο των φυσικών αριθμών $\mathbb{N} = \{1, 2, , \ldots\}$ η οποία έχει την *ιδιότητα της αμνησίας*:

$$P(K > n + k | K > n) = P(K > k), \quad n, k \in \mathbb{N}.$$

Με αλλα λόγια, αν K παριστάνει τον (τυχαίο) χρόνο καταστροφής της Γης, το γεγονός $\{K > k\}$ ότι δεν έχει καταστραφεί η Γη μέχρι τώρα (n) είναι ανεξάρτητο από (δεν μας βοηθάει καθόλου στο να προβλέψουμε) το γεγονός ότι η Γη θα ζήσει k ακόμα χρόνια.

Πόσες τέτοιες τ.μ.[2] υπάρχουν; Αν θέσουμε $G(k) = P(K > k)$, βλέπουμε ότι η παραπάνω σχέση οδηγεί στην

$$G(n + k)/G(n) = G(k),$$

άρα $G(n + 1) = G(n)G(1)$, και άρα $G(n) = G(1)^n$. Θέτοντας $G(1) = \alpha$, έχουμε $G(n) = P(K > n) = \alpha^n$, $n \in \mathbb{N}$, και επομένως

$$P(K = n) = G(n - 1) - G(n) = \alpha^{n-1} - \alpha^n = \alpha^{n-1}(1 - \alpha).$$

Δηλαδή, δεδομένου του $\alpha = P(K = 1)$, προσδιορίζεται πλήρως (με μοναδικό τρόπο) η κατανομή της K. Η K λέγεται γεωμετρική τ.μ. Μόνο αυτή έχει την ιδιότητα της αμνησίας.

[1] Το κείμενο αυτό υπάρχει και στο `http://www.math.upatras.gr/~tkteach`

[2] τ.μ.=τυχαία μεταβλητή

Μια σύμβαση: Αν $\alpha = 0$ τότε $P(K = 1) = 1$ και αν $\alpha = 1$ τότε $P(K = \infty) = 1$. Έχουμε $EK = 1/(1-\alpha)$, και ο τύπος αυτός ισχύει για κάθε $\alpha \in [0,1]$.

2 Αμνησία και ιδιότητα Markov σε διακριτό χρόνο

Θυμηθείτε ότι η οικογένεια των τ.μ. $(X_n, n = 0,1,2,\ldots)$ έχει τη *Μαρκοβιανή ιδιότητα* αν το μέλλον $(X_k, k > n)$ είναι ανεξάρτητο από το παρελθόν $(X_k, k < n)$, δεδομένου του παρόντος X_n, για κάθε τιμή του n. Είναι δε *χρονικά ομογενής* αν οι δεσμευμένη κατανομή της X_n δεδομένης της X_{n-1} δεν εξαρτάται απο το n. Θέτουμε, ως συνήθως, $p_{i,j} = P(X_1 = j | X_0 = i)$. Για μια τέτοια διαδικασία ορίζουμε το χρόνο παραμονής σε μια κατάσταση i ως

$$\sigma_i = \inf\{n \geq 1 : X_n \neq i\}.$$

Ως συνήθως, γράφουμε $P_i(A)$ αντί για $P(A|X_0 = i)$. Ισχυριζόμαστε δε ότι, αν $X_0 = i$, τοτε η σ_i έχει την ιδιότητα της αμνησίας. Πράγματι,

$$\begin{aligned}
P_i(\sigma_i > n+k | \sigma_i > n) &= P_i(X_{n+1} = \cdots = X_{n+k} = i | X_0 = \cdots = X_n = i) && [\text{εξ ορισμού}] \\
&= P_i(X_{n+1} = \cdots = X_{n+k} = i | X_n = i) && [\text{Μαρκοβιανή ιδιότητα}] \\
&= P_i(X_1 = \cdots = X_k = i | X_0 = i) && [\text{Χρονική ομογένεια}] \\
&= P_i(\sigma_i > k) && [\text{εξ ορισμού}]
\end{aligned}$$

Επομένως η σ_i είναι γεωμετρική με $\alpha = P_i(\sigma_i = 1) = P_i(X_i = 1) = p_{i,i}$, δηλαδή,

$$P_i(\sigma_i = n) = p_{i,i}^{n-1}(1 - p_{i,i}), \quad n \in \mathbb{N}, \qquad E_i\sigma_i = 1/(1 - p_{i,i}).$$

Στην περίπτωση που $p_{i,i} = 0$, έχουμε ότι $P_i(\sigma_i = 1) = 1$.

Με βάση αυτή τη συζήτηση, μπορούμε να θεωρήσουμε τη διαδικασία από μια άλλη, ισοδύναμη, οπτική γωνία. Ας ορίσουμε

$$\begin{aligned}
T_1 &= \inf\{n > 0 : X_n \neq X_0\}, \\
T_2 &= \inf\{n > T_1 : X_n \neq X_{T_1}\}, \\
T_3 &= \inf\{n > T_2 : X_n \neq X_{T_2}\}, \ldots
\end{aligned}$$

Κάθε ένας από αυτούς τους τυχάιους χρόνους ικανοποιεί την ιδιότητα

Το γεγονός $\{T_k > n\}$ είναι συνάρτηση των $X_0, \ldots, X_n$.

Μια τυχαία μεταβλητή με αυτή την ιδιότητα λέγεται *χρόνος στάσης*. Όπως ξέρουμε, *κάθε διαδικασία Markov σε διακριτό χρόνο είναι και ισχυρά Markov*[3] δηλαδή, (α) για κάθε χρόνο στάσης T, το παρελθόν $(X_n, n < T)$ είναι ανεξάρτητο από το μέλλον $(X_k, k > T)$, δεδομένου του X_T και του γεγονότος ότι $T < \infty$, και (β) το $(X_k, k \ge T)$ έχει την ίδια κατανομή με το $(X_k, k \ge 0)$, δεδομένου ότι $T < \infty$.

ΑΣΚΗΣΗ 1. *Αποδείξτε, με αυστηρό τρόπο, την ισχυρή αυτή ιδιότητα Markov.*

Αν λοιπόν υποθέσουμε ότι $P(T_k < \infty) = 1$ για κάθε k, πράγμα που σημαίνει ότι $P(T_k < \infty \text{ για κάθε } k) = 1$, τότε εφαρμόζοντας αυτή την ιδιότητα στους χρόνους $0 = T_0, T_1, T_2, \ldots$, συμπεραίνουμε ότι

Η διαδικασία $(X_{T_0}, X_{T_1}, X_{T_2}, \ldots)$ είναι Markov και χρονικά ομογενής.

Ποιές είναι οι πιθανότητες μετάβασης; Απαντηση:

$$P(X_{T_{k+1}} = j | X_{T_k} = i) = P(X_{T_1} = j | X_0 = i) = P_i(X_{\sigma_i} = j)$$
$$= \sum_{n=1}^{\infty} P_i(X_1 = \cdots = X_{n-1} = i, X_n = j) = \sum_{n=1}^{\infty} p_{i,i}^{n-1} p_{i,j} = p_{i,j}/(1 - p_{i,i}), \quad j \ne i,$$
$$P(X_{T_{k+1}} = j | X_{T_k} = i) = 0, \quad j = i.$$

Κατασκευή: Μπορούμε επομένως να κατασκευάσουμε την αρχική αλυσίδα Markov ως εξής: Ξεκινώντας από, π.χ., $X_0 = i$, διαλέγουμε μια γεωμετρική τ.μ. σ_i με μέση τιμή $1/(1 - p_{i,i})$ και θέτουμε $X_n = i$, για $n < \sigma_i$. Διαλέγουμα ακολούθως μια νέα κατάσταση j με πιθανότητα

$$r_{i,j} = \begin{cases} p_{i,j}/(1 - p_{i,i}), & j \ne i \\ 0, & j = i \end{cases},$$

και θέτουμε $X_{\sigma_i} = j$. Διαλέγουμε μια γεωμετρική τ.μ. σ_j, ανεξάρτητη από οτιδήποτε άλλο, και θέτουμε $X_{\sigma_i + n} = j, 0 \le n < \sigma_j$, κ.λπ.

3 Αμνησία σε συνεχή χρόνο

Έστω T μια τ.μ. με την ιδιότητα της αμνησίας:

$$P(T > s + t | T > t) = P(T > s), \quad s, t > 0.$$

[3]Η λέξη «ισχυρά», εδώ, είναι επίρρημα.

ΑΣΚΗΣΗ 2. *Επαναλαμβάνοντας παραπάνω επιχειρήματα και χρησιμοποιώντας το γεγονός ότι η $t \mapsto P(T > t)$ είναι δεξιά συνεχής, να δείξετε ότι η T είναι μια απολύτως συνεχής τ.μ. με*

$$P(T > t) = e^{-\lambda t}, \quad t > 0,$$

για κάποιο $\lambda > 0$. Οι περιπτώσεις $\lambda = 0$ και $\lambda = \infty$ είναι επίσης δυνατές. Αν $\lambda = 0$ τότε $P(T = \infty) = 1$, και αν $\lambda = \infty$ τότε $P(T = 0) = 1$.

Λέμε ότι η T έχει εκθετική κατανομή με παράμετρο λ. Έχουμε,

$$ET = \int_0^\infty P(T > t)dt = 1/\lambda,$$

τύπος που ισχύει για κάθε $\lambda \in [0, \infty]$.

Με άλλα λόγια, αυτή η εκθετική κατανομή είναι η μόνη, σε συνεχή χρόνο, που έχει την ιδιότητα της αμνησίας.

4 Αμνησία και ιδιότητα Markov σε συνεχή χρόνο

Μια οικογένεια από τ.μ. $(X_t, t \geq 0)$ λέγεται Markov αν, για κάθε t, το $(X_s, s < t)$ είναι ανεξάρτητο από το $(X_s, s > t)$, δεδομένου του X_t. Λέγεται χρονικά ομογενής αν η δεσμευμένη κατανομή της X_t δεδομένης της X_s, για $t > s$, εξαρτάται μόνο από το $t - s$.

ΔΕΝ υποθέτουμε, προς το παρόν, ότι οι X_t παίρνουν τιμές σε αριθμήσιμο σύνολο S. Αντίθετα, αφήνουμε το S να είναι 'οτιδήποτε' θέλουμε. Στην περίπτωση αυτή, ΔΕΝ είναι εμφανές ότι Markov συνεπάγεται ισχυρά Markov. Θα δούμε παράδειγμα παρακάτω.

Ας ορίσουμε όμως τι εννοούμε. Κατ' αρχήν, μιά τ.μ. T με τιμές στο $[0, \infty]$, λέγεται χρόνος στάσης, αν, για κάθε t, το γεγονός $\{T > t\}$ προσδιορίζεται πλήρως από το $(X_s, \leq s \leq t)$.

Μια χρονικά ομογενής διαδικασία Markov (X_t) λέγεται ισχυρά Markov, αν, (α) για κάθε χρόνο στάσης T, το $(X_s, s < T)$ είναι ανεξάρτητο από το $(X_s, s > T)$, δεδομένου του X_T και του γεγονότος ότι $T < \infty$, και (β) το $(X_s, s \geq T)$ έχει την ίδια κατανομή με το $(X_s, s \geq 0)$, δεδομένου ότι $T < \infty$.

Θεωρούμε τώρα τους χρόνους

$$\sigma_x = \inf\{t > 0 : \; X_t \neq x\}.$$

ΑΣΚΗΣΗ 3. *Έχοντας υπ' όψη τα επιχειρήματα της Ενότητας 2, αποδείξτε ότι η σ_x έχει την ιδιότητα της αμνησίας, δεδομένου ότι $X_0 = x$.*

Άρα, με βάση το συμπέρασμα της Άσκησης 2, η σ_x είναι εκθετική. Έστω λ_x η παράμετρός της. Έχουμε $E_x \sigma_x = 1/\lambda_x$.

Συνοψίζοντας, έχουμε:
Αν μια χρονικά ομογενής διαδικασία Markov παραμείνει σε μιά κατάσταση για ένα χρονικό διάστημα διάφορο του μηδενός, τότε ο συνολικός χρόνος παραμονής πρέπει να είναι εκθετικά κατανεμημένος.

Αντιπαράδειγμα:

ΑΣΚΗΣΗ 4. *Αν T είναι μια θετική τ.μ., η διαδικασία*

$$X_t = \max\{t - T, 0\}, \quad t \geq 0,$$

είναι Markov αν και μόνο αν η T είναι εκθετική τ.μ. Επίσης, είναι χρονικά ομογενής.

Έστω λ η παράμετρός της. Υποθέτουμε ότι $0 < \lambda < \infty$, για να αποφύγουμε τετριμμένες περιπτώσεις. Τότε όμως, δεν είναι ισχυρά Markov, μιας και το μέλλον μετά το χρόνο στάσης T είναι $(X_{T+t} = t, t \geq 0)$, που, προφανώς, δεν έχει την ίδια κατανομή με το $(X_t, t \geq 0)$.

ΑΣΚΗΣΗ 5. *Πιο γενικά, να δειχτεί ότι αν $\sigma(t) := \inf\{s > 0 : X_{t+s} \neq X_t\}$ και $P(\sigma(t) > 0) > 0$, τότε, αφ' ενός μεν η $\sigma(t)$ είναι εκθετική και ανεξάρτητη από το $(X_s, s < t)$, δεδομένου του X_t, αφ' ετέρου δε ότι αν η διαδικασία είναι ισχυρά Markov, τότε $P($η X είναι συνεχής στο $t + \sigma(t)) = 0$.*

Κατασκευή: Θα κατασκευάσουμε αλυσίδες Markov σε συνεχή χρόνο και τιμές σε ένα πεπερασμένο σύνολο S. Δίνονται αριθμοί $\lambda_i, i \in S$, αυστηρά θετικοί και πεπερασμένοι, καθώς επίσης και πιθανότητες $r_{i,j}$ με $r_{i,i} = 0$, και $\sum_j r_{i,j} = 1$. Κατ' αρχήν, κατασκευάζουμε διαδικασία Markov $(Y_n, n = 0, 1, 2, \ldots)$, σε διακριτό χρόνο, με $P(Y_{n+1} = j | Y_n = i) = r_{i,j}$. Έπειτα, για κάθε $i \in S$, θεωρούμε οικογένεια $\underline{\sigma_i} := (\sigma_i(1), \sigma_i(2), \ldots)$, από ανεξάρτητες τ.μ. με κοιν·η εκθετική κατανομή με παράμετρο λ_i. Επίσης, θεωρούμε ότι οι οικογένειες $\underline{\sigma_i}, i \in S$, είναι ανεξάρτητες. Θεωρούμε τις τ.μ.

$$\sigma_{Y_0}(1), \sigma_{Y_1}(2), \sigma_{Y_2}(3), \ldots.$$

ΑΣΚΗΣΗ 6. *Να δείξετε ότι, δεδομένων των $(Y_0, Y_1, \ldots)$ οι $(\sigma_{Y_0}(1), \sigma_{Y_1}(2), \ldots)$ είναι ανεξάρτητες.*

Ορίζουμε

$$T_0 := 0, \quad T_n := \sigma_{Y_0}(1) + \cdots + \sigma_{Y_{n-1}}(n), \quad n \geq 1,$$

και, τέλος, θέτουμε

$$X_t = \sum_{n \geq 0} Y_n \; \mathbf{1}(T_n \leq t < T_{n+1}), \quad t \geq 0.$$

ΑΣΚΗΣΗ 7. *Να δειχτεί ότι η* $(X_t, t \geq 0)$ *είναι χρονικά ομογενής Markov και ισχυρά Markov. Να δειχτεί επίσης ότι αν* $X_{t-} := \lim_{\epsilon \downarrow 0} X_{t-\epsilon}$, *τότε η* $(X_{t-}, t \geq 0)$ *έχει την ίδια κατανομή με την* $(X_t, t \geq 0)$.

Ας συμβολίσουμε με

$$P_t(i,j) := P(X_{s+t} = j | X_s = i),$$

ποσότητες που δεν εξαρτώνται από το s λόγω χρονικής ομογένειας. Θα δείξουμε ότι

Λήμμα 1.

$$\frac{d}{dt} P_t(i,j) \Big|_{t=0} = \begin{cases} \lambda_i r_{i,j}, & j \neq i \\ -\lambda_i, & j = i. \end{cases}$$

Πριν αποδείξουμε αυτό, ας δούμε πρώτα τα εξής:

Α. Ένας χρήσιμος συμβολισμός του Landau: Γράφουμε $f(x) = o(g(x))$ όταν $x \to x_0$ και εννούμε ότι $\lim_{x \to x_0} f(x)/g(x) = 0$. Με άλλα λόγια, η $f(x)$ είναι μικρότερης τάξης από την $g(x)$ στη γειτονιά του x_0. Το x_0 μπορεί να είναι ένας πραγματικός αριθμός ή $\pm\infty$. Π.χ., $x = o(\sqrt{x})$ όταν $x \to 0$, ενώ $\log x = o(x)$ όταν $x \to \infty$. Ο συμβολισμός είναι κατάλληλος σε διάφορες περιπτώσεις. Π.χ., έστω ότι η f έχει παράγωγο $f'(a)$ στο σημείο a. Μπορούμε να εκφράσουμε αυτό ως $f(a+h) - f(a) = f'(a)h + o(h)$, όταν $h \to 0$. Ανάλογα, γράφουμε $f(x) = O(g(x))$, όταν $x \to x_0$ για να εκφράσουμε την έννοια ότι η $f(x)$ είναι το πολύ της ίδιας τάξης με την $g(x)$ στη γειτονιά του x_0. Αυστηρότερα, $f(x) = O(g(x))$, σημαίνει $\limsup_{x \to x_0} |f(x)/g(x)| < \infty$. Παραδείγματα: $e^{-t} = o(t^n)$ όταν $t \to \infty$, για κάθε $n > 0$. Δηλαδή, η e^{-t} πάει πιο γρήγορα στο 0, όταν το t τείνει στο ∞, από οποιαδήποτε θετική δύναμη του t. Από την άλλη μεριά, $e^{-t} = 1 - t + o(t)$, όταν $t \to 0$. Επίσης, $e^{-t} = 1 - t + t^2/2 + o(t^2)$, όταν $t \to 0$, και $e^{-t} = 1 - t + t^2/2 + O(t)$, όταν $t \to 0$.

Β. Αθροίσματα από ανεξάρτητες εκθετικές τ.μ.: Έστω τ_1, τ_2 ανεξάρτητες εκθετικές με παραμέτρους λ_1, λ_2 και άρα πυκνότητες $f_1(t) = \lambda_1 e^{-\lambda_1 t} \mathbf{1}(t \geq$

0), $f_2(t) = \lambda_2 e^{-\lambda_2 t}\mathbf{1}(t \geq 0)$, αντίστοιχα. Η πυκνότητα της $\tau_1 + \tau_2$ είναι η συνέλιξη

$$\begin{aligned}
f_1 * f_2(t) &= \int_{\mathbb{R}} f_1(t-s) f_2(s) ds \\
&= \int_{\mathbb{R}} \lambda_1 e^{-\lambda_1(t-s)}\mathbf{1}(t-s \geq 0)\lambda_2 e^{-\lambda_2 s}\mathbf{1}(s \geq 0) ds \\
&= \lambda_1 \lambda_2 e^{-\lambda_1 t} \int_0^t e^{(\lambda_1 - \lambda_2)s} ds \\
&= \lambda_1 \lambda_2 e^{-\lambda_1 t}(1 - e^{(\lambda_1-\lambda_2)t})/(\lambda_1 - \lambda_2) \\
&= \lambda_1 \lambda_2 (e^{-\lambda_1 t} - e^{-\lambda_2 t})/(\lambda_1 - \lambda_2), \text{ αν } \lambda_1 \neq \lambda_2,
\end{aligned}$$

ενώ

$$f_1 * f_2(t) = \lambda^2 t e^{-\lambda t}, \text{ αν } \lambda_1 = \lambda_2 = \lambda.$$

Έτσι έχουμε ότι

ΑΣΚΗΣΗ 8.

$$P(\tau_1 + \tau_2 \leq t) = o(t), \quad \text{όταν } t \to 0.$$

Συγκρίνατε αυτό με το $P(\tau_1 \leq t) = 1 - e^{-\lambda_1 t} = \lambda_1 t + o(t)$, *όταν* $t \to 0$. *Επίσης δείξτε ότι* $P(\tau_1 + \cdots + \tau_n \leq t) = o(t)$, *για ανεξάρτητες εκθετικές* $\tau_1, \ldots, \tau_n$. *Μήπως μπορείτε να βελτιώσετε το τελευταίο; Π.χ., μήπως ισχύει ότι* $P(\tau_1 + \cdots + \tau_n \leq t) = o(t^{n-1})$, *όταν* $t \to 0$; *Και γιατί αυτό είναι όντως βελτίωση;*

Γ. Ελάχιστα από ανεξάρτητες εκθετικές τ.μ.: Πάρτε και πάλι $\tau_1, \ldots, \tau_n$, ανεξάρτητες εκθετικές, με παραμέτρους $\lambda_1, \ldots, \lambda_n$. Ορίστε την

$$\tau := \min\{\tau_1, \ldots, \tau_n\}.$$

ΑΣΚΗΣΗ 9. *Δείξτε ότι η* τ *είναι εκθετική με παράμετρο* $\lambda_1 + \cdots + \lambda_n$. *Δείξτε επίσης ότι*

$$P(\tau = \tau_i) = \lambda_i/(\lambda_1 + \cdots + \lambda_n).$$

Απόδειξη του Λήμματος 1. Πρώτα για $j = i$, έχουμε $\frac{d}{dt}P_t(i,j)\big|_{t=0} = \lim_{\delta \to 0}(P_\delta(i,i) - 1)/\delta$. Έστω A_δ το γεγονός $\{X_t = i, 0 \leq t \leq \delta\}$. Έχουμε ότι $P_\delta(i,i) = P_i(X_\delta = i) = P_i(A_\delta) + P_i(X_\delta = i, A_\delta^c)$. Αλλά $P_i(A_\delta) = e^{-\lambda_i \delta} = 1 - \lambda_i \delta + o(\delta)$.

ΑΣΚΗΣΗ 10. *Χρησιμοποιώντας την Άσκηση 8 δείξτε ότι* $P_i(X_\delta = i, A_\delta^c) = o(\delta)$.

Άρα $P_\delta(i,i) - 1 = -\lambda_i\delta + o(\delta)$, που σημαίνει ότι η παράγωγος της $P_t(i,i)$ στο $t = 0$ ισούται με $-\lambda_i$. Από την άλλη μεριά, όταν $j \neq i$, έχουμε ότι $P_\delta(i,j) = P_i(X_\delta = j) = P_i(A_\delta) + P_i(X_\delta = j, A_\delta^c)$, όπου $A_\delta := \{\sigma_i(1) \leq \delta, X_{\sigma_i(1)} = j, \sigma_i(1) + \sigma_j(2) > \delta\}$. Αλλά

$$\begin{aligned} P_i(A_\delta) &= P_i(\sigma_i(1) \leq \delta, X_{\sigma_i(1)} = j) - P(\sigma_i(1) \leq \delta, X_{\sigma_i(1)} = j, \sigma_i(1) + \sigma_j(2) \leq \delta) \\ &= (1 - e^{-\lambda_i\delta})r_{i,j} - o(\delta). \end{aligned}$$

Ο λόγος που ο δεύτερος όρος είναι μικρότερης τάξης από δ είναι και πάλι η Άσκηση 8. Με ανάλογο τρόπο, έχουμε $P_i(X_\delta = j, A_\delta^c) = o(\delta)$, και άρα $P_\delta(i,j) = (1 - e^{-\lambda_i\delta})r_{i,j} + o(\delta) = \lambda_i\delta r_{i,j} + o(\delta)$, πράγμα που σημαίνει ότι $\frac{d}{dt}P_t(i,j)\big|_{t=0} = \lambda_i r_{i,j}$. □

5 Ημιομάδα μετάβασης και γεννήτορας

Κατασκευάσαμε μια χρονικά ομογενή διαδικασία Markov $(X_t, t \geq 0)$ με συνεχή χρόνο και τιμές σε ένα πεπερασμένο σύνολο S. Επαναλαμβάνουμε ότι τα δομικά συστατικά είναι:
(α) Οι παράμετροι λ_i.
(β) Οι πιθανότητες $r_{i,j}$.
Ορίσαμε

$$P_t(i,j) = P_i(X_t = j).$$

Από την ιδιότητα Markov έχουμε

$$P_{t+s}(i,j) = \sum_{k \in S} P_t(i,k) P_s(k,j).$$

Αυτές είναι οι λεγόμενες εξισώσεις των Chapman και Kolmogorov. Βολεύει να θεωρήσουμε τον πίνακα P_t με στοιχεία $P_t(i,j)$. Ο πίνακας αυτός λέγεται πίνακας μετάβασης σε χρόνο t. Παρατηρούμε ότι οι εξισώσεις των Chapman και Kolmogorov εκφράζονται πιο απλά με πολλαπλασιασμό πινάκων:

$$P_{t+s} = P_t P_s.$$

Προφανώς, ο P_0 είναι ο μοναδιαίος πίνακας. Από αλγεβρική σκοπιά, η οικογένεια $(P_t, t \geq 0)$ είναι ημιομάδα, λόγω του ότι είναι κλειστή ως προς πολλαπλασιασμό, έχει μοναδιαίο στοιχείο (αλλά δεν υπάρχουν κατ' ανάγκη αντίστροφα– εξ ου και το πρόθεμα 'ημι' στην 'ομάδα'). Δείξαμε ότι:

$$\frac{d}{dt}P_t(i,j)\bigg|_{t=0} = q_{i,j} = \begin{cases} \lambda_i r_{i,j}, & j \neq i \\ -\lambda_i, & j = i. \end{cases}$$

Είναι χρήσιμο να θεωρήσουμε τον πίνακα Q με στοιχεία $q_{i,j}$. Παρατηρείστε ότι:
(α) $q_{i,i} < 0$,
(β) $q_{i,j} \geq 0$ αν $j \neq i$,
(γ) $\sum_{j \in S} q_{i,j} = 0$.
Ο πίνακας Q λέγεται γεννήτορας της ημιομάδας (και θα δούμε γιατί λέγεται έτσι). Αυτά, όσον αφορά την παράγωγο στο $t = 0$. Αλλά και για $t \neq 0$, η κατάσταση είναι παρόμοια. Ας ονομάσουμε $\dot{P}_t(i,j)$ την παράγωγο στο t. Τότε

Θεώρημα 1. *Η παράγωγος $\dot{P}_t(i,j)$ υπάρχει για κάθε t και ισχύει ότι*

$$\dot{P}_t(i,j) = \sum_{k \in S} P_t(i,k) q_{k,j} = \sum_{k \in S} q_{i,k} P_t(k,j).$$

Με άλλα λόγια,

$$\dot{P}_t = P_t Q = Q P_t.$$

Απόδειξη. Χρησιμοποιείστε τις εξισώσεις των Chapman και Kolmogorov μαζί με το Λήμμα 1. □

Πόρισμα 1.

$$P_t = e^{tQ} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n.$$

Απόδειξη. Από τη βασική θεωρία των γραμμικών διαφορικών εξισώσεων. □

Συζήτηση χωρίς απόδειξη: Είδαμε ότι όλα αυτά ισχύουν για τη συγκεκριμένη κατασκευή που δώσαμε. Ανάποδα, αν $(X_t, t \geq 0)$ είναι μια χρονικά ομογενής διαδικασία Markov σε πεπερασμένο σύνολο S, τότε οι $P_t(i,j)$ ικανοποιούν όλα τα παραπάνω. Επίσης: δεδομένου ενός πίνακα Q με τις ιδιότητες (α)–(γ), υπάρχει μοναδική διαδικασία Markov με αυτόν τον πίνακα σαν γεννήτορα. Τέλος, σημειώνουμε ότι όλα αυτά ισχύουν και για αριθμήσιμα σύνολα S, κάτω από κάποιες αναλυτικές συνθήκες. (Θα δούμε πιο κάτω τι μπορεί να πάει στραβά για άπειρο S.)

6 Ρυθμοί μετάβασης

Η έννοια, σημασία και χρήση των αριθμών $q_{i,j}$ (που αποτελούν τα στοιχεία του πίνακα–γεννήτορα Q της διαδικασίας) είναι τόσο σπουδαία, που θα αφιερώσουμε μια ξεχωριστή ενότητα για περαιτέρω συζήτηση. Θα ονομάσουμε το $q_{i,j}$ *ρυθμό μετάβασης από την κατάσταση i στην κατάσταση j* και ο λόγος είναι ο εξής: Θα αποδειχτεί αργότερα ότι αν η διαδικασία εκτελεί άπειρες μεταβάσεις (από χρόνο 0 μέχρι χρόνο ∞) από την i στην j και αν $N_{i,j}(t)$ δηλώνει το πλήθος των μεταβάσεων αυτών μέχρι το χρόνο t, τότε

$$P(\lim_{t\to\infty} N_{i,j}(t)/t = q_{i,j}) = 1,$$

δηλαδή, πρακτικά, αν το t είναι πολύ μεγάλο μπορούμε να είμαστε σχεδόν σίγουροι ότι θα γίνουν $q_{i,j}t$ μεταβάσεις από τνν i στη j μέχρι το χρόνο t.

Στην πράξη (σε εφαρμογές), δίνονται[4] τα $q_{i,j}$ σαν πρωταρχικά στοιχεία της διαδικασίας, και, όπως είδαμε παραπάνω, όλα τα άλλα στοιχεία προκύπτουν από αυτά.

Ο γράφος της διαδικασία ορίζεται θέτοντας, για κάθε i, j με $q_{i,j} > 0$, μια διατεταγμένη ακμή από το i στο j με βάρος $q_{i,j}$.

Αν η διαδικασία είναι στην κατάσταση i σε κάποια χρονικ·η στιγμή, τότε θα παραμείνει στην κατάσταση αυτή ένα χρονικό διάστημα το οποιο είναι εκθετικά κατανεμημένο με παράμετρο $\lambda_i = -q_{-i,i} = \sum_{j\neq i} q_{i,j} > 0$. Στο τέλος του διαστήματος αυτού, θα πηδήσει στην κατάσταση j η οποία διαλεέγεται με πιθανότητα $q_{i.j}/\lambda_i$.

Αν θυμηθούμε τι είχμε πει για ελάχιστα ανεξάρτητων εκθετικών τ.μ. στην Ενότητα 4 τότε μπορούμε να πραγματοποιήσουμε τη διαδικασία με τον εξής τρόπο: Σε κάθε κατάσταση i αντιστοιχούμε τόσα «εκθετικά κατανεμημένα ρολόγια» όσες είναι οι καταστάσεις j με $q_{i,j} > 0$. Έστω $j_1, \ldots, j_k$ αυτές οι καταστάσεις. Λέγοντας «ρολόγια» εννοούμε τ.μ. $\tau_{j_1}, \ldots, \tau_{j_k}$, ανεξάρτητες και εκθετικά κατανεμημένες με παραμέτρους $q_{i,j_1}, \ldots, q_{i,j_k}$. Τα ρολόγια αυτά τα σκεφτόμαστε σαν ξυπνητήρια. Το πρώτο ξυπνητήρι θα χτυπήσει σε χρόνο ίσο με $\min(\tau_{j_1}, \ldots, \tau_{j_k})$ που είναι εκθετικά κατανεμημένος με παράμετρο $q_{i,j_1} + \cdots + q_{i,j_k} = \lambda_i$. Αν δε αυτό το ρολόι είναι το ρολόι j_m τότε θα οδηγηθούμε στην κατάσταση j_m, και όπως ξέρουμε, αυτό γίνεται με πιθανότητα $P(\tau = \tau_{j_m}) = q_{i,j_m}/\lambda_i$ (Άσκηση 9).

7 Διαδικασία Poisson

Ας θεωρήσουμε τη διαδικασά που ορίζεται από τους ρυθμούς μετάβασης $q_{i,i+1} = \lambda$, $i = 0, 1, \ldots$, και $q_{i,j} = 0$ για j διαφορετικό απο $i+1$ ή i. Με άλλα λόγια, ο γράφος της διαδικασίας είναι πάνω στο σύνολο $\mathbb{Z}_+$ και έχει ακμές από το

[4] Ή βγαίνουν από τα δεδομένα του προβλήματος

i στο $i+1$ με βάρος λ όλες, για $i = 0, 1, 2, \ldots$. Έστω N_t η κατάσταση της διαδικασίας τη στιγμή t.

Αν υποθέσουμε ότι $N_0 = 0$, τότε, από τα παραπάνω, είναι εμφανές ότι η διαδικασία παραμένει στο 0 για ενα εκθετικά κατανεμημένο χρόνο με παράμετρο λ, μετα πηδάει στο 1 και μένει εκεί εναν εκθετικά κατανεμημένο χρόνο με παράμετρο λ, ανεξάρτητο από το παρελθόν, μετα πηδάει στο 2, κ.λπ.

Ας ονομάσουμε $\tau_1, \tau_2, \ldots$ αυτούς τους ανεξάρτητους εκθετικούς (λ) χρόνους, και άς θέσουμε

$$T_0 = 0, \quad T_n = \tau_1 + \cdots + \tau_n, \quad n \geq 1.$$

ΑΣΚΗΣΗ 11. *Δείξτε ότι*

$$N_t = \sum_{n \geq 1} \mathbf{1}(T_n \leq t) = \sum_{n \geq 0} n \mathbf{1}(T_n \leq t < T_{n+1}).$$

Θα δείξουμε ότι

Θεώρημα 2.

$$P(N_t = n) = P(N_{s+t} = k + n | N_s = k) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad t, s \geq 0, n, k \in \mathbb{Z}_+.$$

Με άλλα λόγια, η τ.μ. N_t έχει *κατανομή Poisson* με παράμετρο λt. Θα δούμε ΔΥΟ αποδείξεις για το θεώρημα αυτό.

Πρώτη απόδειξη. Κατ' αρχήν, η ισότητα $P(N_t = n) = P(N_{s+t} = k + n | N_s = k)$ είναι προφανής λόγω χρονικής και χωρικής ομογένειας. Έστω $\varphi_t(n) = P(N_t = n)$. Από το Θεώρημα 1, ή–ακόμα καλύτερα–κάνοντας σκέψεις από τηην αρχή,

$$\begin{aligned}
\varphi_{t+\delta}(n) &= P(N_{t+\delta} = n) \\
&= P(N_{t+\delta} = n, N_t = n) + P(N_{t+\delta} = n, N_t = n - 1) + o(\delta) \\
&= e^{-\lambda\delta} P(N_t = n) + (1 - e^{-\lambda\delta}) P(N_t = n - 1) + o(\delta) \\
&= (1 - \lambda\delta + o(\delta))\varphi_t(n) + (\lambda\delta + o(\delta))\varphi_t(n-1) + o(\delta),
\end{aligned}$$

απ' όπου συμπεραίνουμε ότι

$$\begin{aligned}
\frac{1}{\delta}[\varphi_{t+\delta}(n) - \varphi_t(n)] &= \lambda[\varphi_t(n-1) - \varphi_t(n)] + o(\delta)/delta, \\
\dot{\varphi}_t(n) &= \lambda[\varphi_t(n-1) - \varphi_t(n)].
\end{aligned}$$

Αυτό είναι ένα σύστημα γραμμικών διαφορικών εξισώσεων, μία για κάθε n, και άρα άπειρες σε πλήθος. Ευτυχώς όμως η επόμενη εξαρτάται από το αποτέλεσμα της προηγούμενης[5] και άρα μπορούμε, πολλαπλασιάζοντας με $e^{\lambda t}$ παντού, να βρούμε ότι

$$\frac{d}{dt}(e^{\lambda t}\varphi_t(n)) = \lambda e^{\lambda t}\varphi_t(n-1).$$

Αν τώρα βαφτίσουμε το $e^{\lambda t}\varphi_t(n)$ με ένα καινούργιο όνομα, ας πούμε, π.χ.,

$$\psi_t(n) = e^{\lambda t}\varphi_t(n),$$

έχουμε

$$\frac{d}{dt}\psi_t(n) = \lambda\psi_t(n-1),$$
$$\psi_t(n) = \lambda\int_0^t \psi_u(n-1)du,$$

αφού $\psi_o(n) = 0$ πάντα. Αλλά,

$$\psi_t(0) = e^{\lambda t}\varphi_t(0) = e^{\lambda t}P(N_t = 0) = e^{\lambda}te^{-\lambda t} = 1,$$

και άρα, αναδρομικά,

$$\begin{aligned}
\psi_t(1) &= \lambda\int_0^t \psi_u(0)du = \lambda t,\\
\psi_t(2) &= \lambda\int_0^t \psi_u(1)du = (\lambda t)^2/2,\\
\psi_t(3) &= \lambda\int_0^t \psi_u(3)du = (\lambda t)^3/3!,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
\psi_t(n) &= (\lambda t)^n/n!,
\end{aligned}$$

και συνεπώς $\varphi_t(n) = e^{-\lambda t}\psi_t(n) = e^{-\lambda t}(\lambda t)^n/n!$, που είναι αυτό που θέλουμε. □

Σχόλιο: Η πρώτη απόδειξη είναι κατασκευαστική, δηλαδή ακόμα και να μην ξέραμε τον τύπο, θα τον βγάζαμε. Η παρακάτω απόδειξη είναι λιγότερο κατασκευαστική. (Αλλά δες και σχόλιο μετά.)

[5] Μήπως σας θυμίζει αυτό τη χωρική δομή της διαδικασίας;

Δεύτερη απόδειξη. Αφού η T_n είναι άθροισμα από n ανεξάρτητες ισόνομες τ.μ., έχουμε, για κάθε $\alpha > 0$,

$$Ee^{-\alpha T_n} = (Ee^{-\alpha \tau_1})^n = \left(\int_0^\infty e^{-\alpha t} \lambda e^{-\lambda t} dt\right)^n = (\lambda/(\lambda+\alpha))^n.$$

Χρησιμοποιώντας αυτό, βρίσκουμε

$$\begin{aligned}
\int_0^\infty e^{-\alpha t} P(N_t = n) dt &= E \int_0^\infty e^{-\alpha t} \mathbf{1}(T_n \le t < T_{n+1}) dt \\
&= E \int_{T_n}^{T_{n+1}} e^{-\alpha t} dt \\
&= E \alpha^{-1} (e^{-\alpha T_n} - e^{-\alpha T_{n+1}}) \\
&= \alpha^{-1} [(\lambda/(\lambda+\alpha))^n - (\lambda/(\lambda+\alpha))^{n+1}] \\
&= (\lambda/(\lambda+\alpha))^n [1 - (\lambda/(\lambda+\alpha))].
\end{aligned}$$

Από την άλλη μεριά, έχουμε

$$\int_0^\infty e^{-\alpha t} e^{-\lambda t} (\lambda t)^n / (n!) \, dt = (\lambda/(\lambda+\alpha))^n [1 - (\lambda/(\lambda+\alpha))].$$

Και τώρα χρησιμοποούμε ένα θεώρημα από την Ανάλυση που λέει ότι
Η μόνη συνάρτηση h που ικανοποιεί ότι $\int_0^\infty e^{-\alpha t} h(t) dt = 0$ για κάθε $\alpha = 0$ είναι η μηδενική. □

Σχόλιο: Αν $h : \mathbb{R}_+ \to \mathbb{C}$ είναι μια συνάρτηση που «δεν αυξάνει γρηγορότερα από ενα εκθετικό», τότε μπορύμε να ορίσουμε τη μιγαδική συνάρτηση $\hat{h} : \mathbb{C} \to \mathbb{C}$ με $\hat{h}(z) = \int_0^\infty e^{-zt} h(t) dt$, τουλάχιστον σε κάποιο ανοιχτό σύνολο της μορφής $\{z \in \mathbb{C} : \Re(z) < c\}$, και, όπως θα μάθετε στη Θεωρία των Συναρτήσεων, με μοναδική αναλυτική επέκλταση σε όλο το $\mathbb{C}$. Η $\hat{h}$ ονομάζεται *μετασχηματισμός Laplace* της h, η δε συνάρτηση $h \mapsto \hat{h}$ ονομάζεται μετασχηματισμός Laplace. Για την ακρίβεια, δεδομένης τηε $\hat{h}$ μπορούμε να ανακτήσουμε την h χρησιμοποιώντας τη σχέση

$$h(t) = \frac{1}{2\pi\sqrt{-1}} \int_C \hat{h}(z) e^{zt} dz,$$

όπου C οποιαδήποτε ευθεία της μορφής $C := \{z \in \mathbb{C} : \Re(z) = c\}$ περιλαμβάνει τους πόλους της $\hat{h}$ στα αριστερά της. Στην παραπάνω περίπτωση, έχουμε $h(z) = (\lambda/(\lambda+z))^n [1 - (\lambda/(\lambda+z))]$, με πόλο στο σημείο $z = -\lambda$.

ΑΣΚΗΣΗ 12. *Αποδείξτε τα εξής:*
1. Αν ξ_n είναι ανεξάρτητες τ.μ., όλες με $P(\xi_n = 1) = p$, $P(\xi_n = 0) = 1 - p$,

και ανεξάρτητες από τη διαδικασία Poisson N που εισαγάγαμε παραπάνω, τότε η διαδικασίες

$$N_t^1 := \sum_{n=1}^{\infty} \mathbf{1}(T_n \le t, \xi_n = 1), \quad t \ge 0,$$
$$N_t^0 := \sum_{n=1}^{\infty} \mathbf{1}(T_n \le t, \xi_n = 0), \quad t \ge 0,$$

είναι κι αυτές Poisson, με παραμέτρους λp, $\lambda(1-p)$, αντίστοιχα, και είναι ανεξάρτητες μεταξύ τους.
2. Αν $(N_t^1, t \ge 0)$, $(N_t^2, t \ge 0)$, είναι ανεξάρτητες διαδικασίες Poisson με παραμέτρους λ_1, λ_2, αντίστοιχα, τότε η $N_t := N_t^1 + N_t^2$ είναι κι αυτή Poisson με παράμετρο $\lambda_1 + \lambda_2$.
3. Αν η $(N_t, t \ge 0)$ είναι διαδικασία Poisson με παράμετρο λ, και $\alpha > 0$, τότε η $(N_{\alpha t}, t \ge 0)$ είναι διαδικασία Poisson με παράμετρο $\alpha\lambda$.

8 Άλλες δομικές ιδιότητες της διαδικασίας Poisson:

Συμβολισμός: Έστω $x_1, \ldots, x_n$ πραγματικοί αριθμοί. Θέτουμε $x_{(1)} = \min\{x_1, \ldots, x_n\}$, $x_{(2)} = \min(\{x_1, \ldots, x_n\} - \{x_{(1)}\})$, $x_{(3)} = \min(\{x_1, \ldots, x_n\} - \{x_{(1)}, x_{(2)}\}, \ldots$.

ΑΣΚΗΣΗ 13. *Να αποδειχτεί ότι η δεσμευμένη κατανομή του $(T_1, \ldots, T_n)$, δεδομένου ότι $N_t = n$ είναι η κατανομή του $(U_{(1)}, \ldots, U_{(n)})$, όπου $U_1, \ldots, U_n$ είναι ανεξάρτητες, όλες ομοιόμορφες στο $[0, t]$.*

ΑΣΚΗΣΗ 14. *Να αποδειχτεί ότι αν $t_0 < t_1 < t_2 < \cdots$, οι τ.μ. $N_{t_1} - N_{t_0}$, $N_{t_2} - N_{t_1}$, ..., είναι ανεξάρτητες.*

ΑΣΚΗΣΗ 15. *Να υπολογιστεί η κατανομή*

$$P(N_{t_0} = n_0, N_{t_1} = n_1, \ldots, N_{t_k} = n_k).$$

Επεκτείνοντας το συμβολισμό, ορίζουμε, για $A \subseteq [0, \infty)$,

$$N(A) = \sum_{n \ge 1} \mathbf{1}(T_n \in A),$$

την τ.μ. που δίνει το πλήθος των σημείων T_n που βρίσκονται στο σύνολο A. Χρησιμοποιώντας την Άσκηση 14, μπορούμε να πειστούμε ότι αν $A, B \subseteq [0, \infty)$, και $A \cap B = \emptyset$, τότε $N(A)$, $N(B)$ είναι ανξάρτητες τ.μ., μάλιστα δε

$$P(N(A) = n) = e^{-\lambda|A|}(\lambda|A|)^n/n!, \quad n = 0, 1, 2, \ldots,$$

δηλαδή Poisson με παράμετρο $\lambda|A|$, όπου $|A|$ δηλώνει το μήκος του A.

ΑΣΚΗΣΗ 16. *Δείξτε ότι*

$$P(\lim_{t\to\infty} N_t/t = \lambda) = 1.$$

Υπόδειξη: Έχουμε ότι $E\tau_1 = 1/\lambda$, αφού η τ_1 είναι εκθετική με παράμετρο λ. Αφού $T_n = \tau_1 + \cdots + \tau_n$, από τον Ισχυρό Νόμο των Μεγάλων Αριθμών, $P(\lim_{n\to\infty} T_n/n = 1/\lambda) = 1$. Παρατηρείστε ότι $T_{N_t} \le t < T_{N_t+1}$, για κάθε t (γιατί;) Μετά αποδείξτε ότι $P(\lim_{t\to\infty} N_t = \infty) = 1$ και συνδυάστε τα παραπάνω.

9 Διαδικασία Poisson σε Ευκλείδιο χώρο

Θεωρείστε τον $\mathbb{R}^d$ και ονομάστε $|A|$ τον «d-διάστατο όγκο» του συνόλου A. Θέλουμε να φτιάξουμε ένα ΤΥΧΑΙΟ, αριθμήσιμο, σύνολο Π το οποίο να είναι όσο το δυνατό πιο τυχαίο. Τι εννοούμε με αυτό; Θέλουμε να έχει τη δομική ιδιότητα: Αν παρατηρήσουμε τα σημεία του Π που βρίσκονται σε ένα σύνολο A να μην μπορούμε να πάρουμε καμμία απολύτως πληροφορία για τα σημεία του Π που βρίσκονται σε ένα άλλο σύνολο B που δεν τέμνει το A. Αν ονομάσουμε $\Pi(A)$ το πλήθος των σημείων του Π που βρίσκονται στο A, με άλλα λόγια, θέλουμε οι τ.μ. $\Pi(A)$, $\Pi(B)$ να είναι ανεξάρτητες.

Θεώρημα 3 (χωρίς απόδειξη). *Αν, για κάθε $A_1, \ldots, A_k$, ξένα ανά δύο, έχουμε ότι $\Pi(A_1), \ldots, \Pi(A)k)$ είναι ανεξάρτητες τ.μ., και αν $E\Pi(A) = |A|$, τότε η $\Pi(A)$ είναι Poisson τ.μ. με παράμετρο $|A|$.*

ΟΡΙΣΜΟΣ: Το ΤΥΧΑΙΟ ΣΥΝΟΛΟ Π λέγεται διαδικασία Poisson στον $\mathbb{R}^d$.

Φυσικά, αντί $E\Pi(A) = |A|$. μπορούμε να πολλαπλασιάσουμε με μια θετικής σταθερά λ, και να έχουμε διαδικασία Poisson με παράμετρο $\lambda|A|$.

Γεννήτριες συναρτήσεις. Στη Θεωρία Πιθανοτήτων, οι γεννήτριες συναρτήσεις παίζουν σημαντικό ρόλο, ιδίως όταν μπλέκεται μέσα και η ανεξαρτησία. Για μια τ.μ. X, θεωρούμε τη γεννήτρια

$$\varphi(\theta) = Ee^{\theta X},$$

όπου το θ είναι πραγματικός (ή μιγαδικός) αριθμόσ–που ανήκει σε κατάλληλο σύνολο, εξαρτώμενο από τις ιδίοτητες της X κάθε φορά· αλλά αυτές είναι ανλυτικές λεπτομέρειες που θα αγνοήσουμε. Για δύο τ.μ., X_1, X_2, η γεννήτρια είναι συνάρτηση δύπ μεταβλητών:

$$\varphi(\theta_1, \theta_2) = Ee^{\theta_1 X_1 + \theta_2 X_2}.$$

Αυτό γενικεύεται για n τ.μ.–δηλ. για ένα τυχαίο δυάνυσμα $X = (X_1, \dots, X_n)$:

$$\varphi(\theta) = Ee^{\langle \theta, X \rangle},$$

όπου $\theta = (\theta_1, \dots, \theta_n)$, και $\langle \theta, X \rangle$ δηλώνει το εσωτερικό γινόμενο $\theta_1 X_1 + \cdots + \theta_n X_n$. Αν το θ είναι δοσμένο δυάνυσμα στον $\mathbb{R}^n$, τότε το εσωτερικό γινόμενο $\mathbb{R}^n \ni x \mapsto \langle \theta, x \rangle \in R$ δεν είναι παρά μια γραμμική συνάρτηση $\mathbb{R}^n \to \mathbb{R}$. Οποιαδήποτε γραμμική συνάρτηση $\mathbb{R}^n \to \mathbb{R}$ είναι της μορφής «εσωτερικό γινόμενο».

Έτσι, πιο γενικά, αν έχουμε ένα άπειρο δυάνυσμα, πρέπει να βρούμε τι εννοούμε με εσωτερικό γινόμενο, δηλαδή τι είναι εκείνες οι πραγματικές γραμμικές συναρτήσεις του απείρου δυανύσματος. Αν π είναι ένα το πολύ αριθμήσιμο υποσύνολο του $\mathbb{R}^d$, με στοιχεία $x_1, x_2, \dots$, τότε, γιά κάθε συνάρτηση $f : \mathbb{R}^d \to \mathbb{R}$, χρησιμοποιούμε το συμβολισμό

$$\int f(x) d\pi(x) = \sum f(x_i).$$

Βλέπουμε ότι αυτός ο αριθμός, σαν συνάρτηση της συνάρτησης f, είναι γραμμικός. Αυτή ακριβώς είναι η κατάλληλη έννοια όσον αφορά το εσωτερικό γινόμενο.

Άρα, αν θέλουμε να ορίσουμε τη γεννήτρια συνάρτηση μιας διαδικασίας Poisson Π θα πρέπει να την θεωρήσουμε σαν συνάρτηση μιας συνάρτησης f, με

$$\varphi(f) = E \exp \int f(x) d\Pi(x).$$

Το εκπληκτικό είναι ότι έχουμε κλειστό τύπο για την $\varphi(f)$:

Θεώρημα 4.

$$\varphi(f) = E \exp \int f(x) d\Pi(x) = \exp \lambda \int (e^{f(x)} - 1) dx$$

Απόδειξη. (στο περίπου) Ας πάρουμε $f(x) = c\mathbf{1}(x \in A)$, δηλαδή $f(x) = c$ αν $x \in A$ και $f(x) = 0$ αν $x \notin A$. Τότε

$$\int c\mathbf{1}_A(x) d\Pi(x) = c\Pi(A) = \ c\times\text{πλήθος σημείων του } \Pi \text{ στο } A.$$

Αφού η $\Pi(A)$ είναι Poisson με παράμετρο $\lambda|A|$ τότε

$$\varphi(c\mathbf{1}_A) = Ee^{c\Pi(A)} = \exp(\lambda|A|(e^c - 1)) = \exp \lambda \int (e^{c\mathbf{1}_A} - 1) dx.$$

Αν η f είναι της μορφής

$$f = \sum_{j=1}^{n} c_j \mathbf{1}_{A_j},$$

όπου τα $A_1, \ldots A_n$ είναι ξένα μεταξύ τους, τότε

ΑΣΚΗΣΗ 17. *(Χρησιμοποιώντας την ανεξαρτησία μεταξύ των $\Pi(A_1), \ldots, \Pi(A_n)$) δείξτε ότι ισχύει και πάλι ο τύπος.*

Θα τελειώσουμε την απόδειξη λέγοντας ότι «κάθε» συνάρτηση f προσεγγίζεται από συναρτήσεις της προηγούμενης μορφής και, «παίρνοντας όρια», μπορούμε να καταλήξουμε στο συμπέρασμα. Αυτό όμως χρειάζετε να καταλάβετε γενική θεωρία ολοκληρωσης, με άλλα λόγια θεωρία μέτρου. □

ΑΣΚΗΣΗ 18. *Να υπολογιστεί η $\varphi(f)$ στο σημείο f, όπου $f(x) = e^{-||x||^2}$, $||x||^2 = x_1^2 + \cdots + x_d^2$.*

Θεώρημα 5 (Πολύ κομψό και όμορφο). *Αν το τυχαίο σύνολο Π ικανοποιεί τη σχέση*

$$E \exp \int f(x) d\Pi(x) = \exp \lambda \int (e^{f(x)} - 1) dx,$$

για «κάθε» f, τότε το Π πρέπει να είναι διαδικασία Poisson.

10 Διαδικασίες Poisson και διαδικασίες Markov

Θεωρείστε και πάλι ε'να πεπερασμένο σύνολο S και μια διαδικασία Markov $(X_t, t \geq 0)$ με τιμές στο S και ρυθμούς $q_{i,j}$.

Θα δούμε πως μπορύμε να χρησιμοποιήσουμε μια διαδικασία Poisson για να κατασκευάσουμε την X.

Έστω λ ένας αριθμός που είναι μεγαλύτερος από κάτηε $\lambda_i = -q_{i,i} = \sum_{j \neq i} q_{i,j}$. Θέτουμε

$$p_{i,j} = \begin{cases} q_{i,j}/\lambda, & i \neq j, \\ 1 - \lambda_i/\lambda, & i = j. \end{cases}$$

Εύκολα βλέπουμε ότι $p_{i,j} \geq 0$ και $\sum_j p_{i,j} = 1$.

Ας θεωρήσουμε λοιπόν μια διαδικασία Markov $(Z_n, n \in \mathbb{Z}_+)$, σε διακριτό χρόνο, με αυτές τις $p_{i,j}$ ως πιθανότητες μετάβασης. Γνωρίζουμε ότι υπάρχει συνάρτηση $h : S \times [0, 1] \to S$ έτσι ώστε

$$Z_{n+1} = h(Z_n, \xi_n),$$

όπου (ξ_n) είναι ανεξάρτητες ομοιόμορφες στο $[0,1]$. Έστω επίσης $(N_t, t \geq 0)$ μια διαδικασία Poisson με παράμετρο λ, ανεξάρτητη από τα (ξ_n). Ονομάζουμε, ως συνήθως, με $T_1, T_2, \ldots$ τους χρόνους κατά τους οποίους η N αλλάζει τιμή. Δηλαδή, $N_t = n$ αν $T_n \leq t < T_{n+1}$, $n \geq 1$.

Αν $f : S \to \mathbb{R}$, τότε

$$\begin{aligned} f(X_t) - f(X_0) &= f(X_{T_{N_t}}) - f(X_0) \\ &= [f(X_{T_{N_t}}) - f(X_{T_{N_t-1}})] + \cdots + [f(X_{T_2}) - f(X_{T_1})] + [f(X_{T_1}) - f(X_0)] \\ &= \sum_{n \geq 0} [f(X_{T_{n+1}}) - f(X_{T_n})] \mathbf{1}(T_n \leq t). \end{aligned}$$

(Εδώ θεωρήσαμε ότι $0 = T_0$.) Χρησιμοποιώντας το συμβολισμό της προηγούμενης ενότητας και τη συνάρτηση h, γράφουμε τη σχέση αυτή ως

$$f(X_t) - f(X_0) = \int_0^t [f(h(X_{s-}, \xi_{N_{s-}})) - f(X_{s-})] \, dN_s, \tag{1}$$

όπου $X_{s-} := \lim_{\epsilon \downarrow 0} X_{s-\epsilon}$.

Αφού η (1) ισχύει για κάθε f, η σχέση αυτή μας δίνει μια συγκεκριμένη αναπαράσταση της διαδικασίας Markov.[6]

Ας θυμηθούμε τώρα ότι με Q συμβολίζουμε τον πίνακα με στοιχεία $q_{i,j}$. Αν θεωρήσουμε την f σαν ένα δυάνυσμα-στήλη με στοιχεία $f(i)$, τότε με Qf είναι το δυάνυσμα-στήλη με στοιχεία

$$Qf(i) = \sum_j q_{i,j} f(j).$$

ΑΣΚΗΣΗ 19. *Αποδείξτε ότι*

$$Qf(i) = \sum_j [f(j) - f(i)] q_{i,j}.$$

Παίρνοντας μέσες τιμές στην (1) έχουμε:

$$\begin{aligned} Ef(X_t) - Ef(X_0) &= E \int_0^t [f(h(X_{s-}, \xi_{N_{s-}})) - f(X_{s-})] \, dN_s \\ &= E \int_0^t \sum_j [f(j) - f(X_{s-})] \frac{q_{i,j}}{\lambda} \, \lambda ds \\ &= E \int_0^t Qf(X_{s-}) \, ds = E \int_0^t Qf(X_s) \, ds. \end{aligned}$$

[6] Αρκεί, π.χ., να ξέρουμε την $f(X_t)$ για κάτηε $f = \delta_i$, όπου $\delta_i(x) := \mathbf{1}(x = i)$, $i \in S$.

Με άλλα λόγια, η

$$M_t := f(X_t) - f(X_0) - \int_0^t Qf(X_s)\, ds$$

έχει μέση τιμή μηδέν.

ΑΣΚΗΣΗ 20. *Επιπλέον, αποδείξτε με πιό αυστηρό τρόπο απο το «ψευτο-επιχείρημά» μου ότι*

$$E(M_t - M_s | X_u, 0 \le u \le s) = 0, \quad s < t.$$

Μιά διαδικασία (M_t) που ικανοποιεί την τελευταία σχέση λέγεται martingale.

Η κατασκευή (1) λέγεται ομογενοποίηση.

11 Διαδικασίες Poisson και Markov – Μέρος II

Υπάρωουν πολλοί τρόποι να χρησιμοποιήσει κανείς διαδικασίες Poisson για να φτιάξει διαδικασίες Markov. Η προσέγγιση της προηγούμενης ενότητας πάσχει από το γεγονός ότι χρησιμοποίησε και άλλα στοιχεία, πέραν μιας διαδικασίας Poisson, δηλ., τις τ.μ. (ξ_n).

Ας δούμε πως να κάνουμε χωρίς αυτές σε ένα παράδειγμα.

Έστω η διαδικασία Markov (X_t) με $q_{i,i+1} = \lambda$, $q_{i+1,i} = \mu$, $i \in \mathbb{Z}_+$, ενώ $q_{i,j} = 0$ δε οποιαδήποτε άλλη περίπτωση με $j \ne i$. Έστω A, B δύο ανεξάρτητες διαδικασίες Poisson με παραμέτρους λ, μ, αντίστοιχα.

ΑΣΚΗΣΗ 21. *Αποδείξτε ότι οι παρακάτω σχέσεις ορίζουν τη διαδικασία* (X_t)*:*

1.

$$X_t = X_0 + A_t - \int_0^t \mathbf{1}(X_{s-} > 0)\, dB_s.$$

2.

$$X_t = X_0 + A_t - B\left(\int_0^t \mathbf{1}(X_s > 0)\, ds\right).$$

Η διαδικασία αυτή λέγεται διαδικασία γέννησησ-θανάτου ή και ουρά τύπου M/M/1.