Vorlesung 12a

Markovketten (Teil 3)

Gleichgewichtsverteilungen

1. Begriffsbildung

Sei P eine Übergangsmatrix auf S

und ρ eine (Start-)Verteilung auf S.

Dann gilt

$$\mathbf{P}_\rho(X_0 = a_0, X_1 = a_1) = \rho(a_0)P(a_0, a_1), \text{ also}$$

$$\mathbf{P}_\rho(X_0 = a_0) = \rho(a_0),$$

$$\mathbf{P}_\rho(X_1 = a_1) = \sum_{a_0 \in S} \rho(a_0)P(a_0, a_1).$$

Für welche Startverteilung ρ ist X_1 so verteilt wie X_0?

Eine Verteilung π auf S heißt

Gleichgewichtsverteilung zur Übergangsmatrix P,

wenn eine der folgenden äquivalenten Bedingungen erfüllt ist:

(G1) $$\sum_{a \in S} \pi(a) P(a,b) = \pi(b) \ , \quad b \in S \ .$$

(G2) $$\mathbf{P}_\pi(X_1 = b) = \mathbf{P}_\pi(X_0 = b) \ , \quad b \in S$$

d.h. unter $\mathbf{P}_\pi$ haben X_0 und X_1 dieselbe Verteilung.

Reversible Gleichgewichtsverteilungen

Hinreichend für

(G2)

$$\mathbf{P}_\pi(X_1 = b) = \mathbf{P}_\pi(X_0 = b)\ , \quad b \in S$$

ist die Bedingung

(R)

$$\mathbf{P}_\pi(X_0 = a, X_1 = b) = \mathbf{P}_\pi(X_0 = b, X_1 = a)\ , \quad a, b \in S$$

Denn dann ist unter $\mathbf{P}_\pi$

das Paar (X_0, X_1) so verteilt wie (X_1, X_0),

also insbesondere X_0 so verteilt wie X_1.

Gleichbedeutend mit

(R)

$$\mathbf{P}_\pi(X_0 = a, X_1 = b) = \mathbf{P}_\pi(X_0 = b, X_1 = a)\ , \quad a, b \in S$$

ist

$$\pi(a)P(a,b) = \pi(b)P(b,a)\ , \quad a, b \in S\ .$$

π heißt dann *reversible Gleichgewichtsverteilung* zu P.

2. Beispiele:

Ein Beispiel einer nicht reversiblen Gleichgewichtsverteilung:

Zyklische Irrfahrt auf $S = \{a, b, c\}$, mit

$$P(a,b) = P(b,c) = P(c,a) := p,$$

$$P(b,a) = P(c,b) = P(a,c) := 1 - p$$

Die uniforme Verteilung auf S

ist eine Gleichgewichtsverteilung zu P.

Nur für $p = 1/2$ ist sie reversibel.

Die Gleichgewichtsverteilung der einfachen Irrfahrt auf dem Würfel $S = \{0, 1\}^3$:

Von jedem $a \in S$ geht man in einem Schritt
zu einem rein zufällig ausgewählten Nachbarn.
(Zwei Elemente von S heißen *benachbart*,
wenn sie sich in genau einer Komponente unterscheiden.)
Für benachbarte Knoten a und b ist hier $P(a, b) = 1/3$.

Die uniforme Verteilung auf S
ist reversible Gleichgewichtsverteilung.

Eine wichtige Beispielklasse:

Die einfache Irrfahrt

auf einem ungerichteten, zusammenhängenden Graphen

mit endlicher Knotenmenge S

Von jedem $a \in S$ geht man in einem Schritt

zu einem rein zufällig ausgewählten Nachbarn:

$$P(a,b) = \frac{1}{g(a)},$$

mit b Nachbar von a, $g(a) := \#$ Nachbarn von a

Ansatz: $\pi(a) := \frac{1}{c} g(a)$

Diese Verteilung π erfüllt die die Reversibilitätsbedingung (R),

denn für benachbarte Knoten a, b gilt:

$$\frac{1}{c} g(a) \frac{1}{g(a)} = \frac{1}{c} g(b) \frac{1}{g(b)}$$

Man kann zeigen (hier ohne Beweis): Es gibt (unter den gegebenen Voraussetzungen) nur **eine** Gleichgewichtsverteilung.

Fazit: Die Gewichte der Knoten

unter der Gleichgewichtsverteilung

sind proportional zur Anzahl der Nachbarn der Knoten.

3. Das Ehrenfest-Modell

veröffentlicht 1909 von Paul und Tatjana Ehrenfest, konzipiert als Spielzeugmodell für Boltzmanns Statistische Mechanik:

d Teilchen sind verteilt auf eine linke und eine rechte Urne:
ℓ Teilchen links, r Teilchen rechts.
In jedem Schritt wird rein zufällig eines aus den d ausgewählt und in die andere Urne verfrachtet.
Hat diese Dynamik eine Gleichgewichtsverteilung, und wenn ja, wie sieht sie aus?

Zur Illustration betrachten wir hier nur den Fall $d = 3$. Die Übergangsw'keiten für die *Anzahl links* sind dann (für $\ell = 0, 1, 2, 3$):

$$P(\ell, \ell + 1) = \frac{3 - \ell}{3}, \qquad P(\ell, \ell - 1) = \frac{\ell}{3}.$$

Ein eleganter Weg zur Antwort führt über ein *Feinmodell*:

Die Teilchen werden nummeriert mit $1, 2., 3$.

$$a_i = \begin{cases} 1 & \text{falls das Teilchen mit Nr. } i \text{ in linker Urne,} \\ 0 & \qquad\ldots\ldots\qquad \text{in rechter Urne.} \end{cases}$$

$$a := (a_1, a_2, a_3) \in \{0, 1\}^3.$$

Dynamik des Feinmodells:

Eine Nummer $i \in \{1, 2, 3\}$ wird rein zufällig ausgewählt

und das a_i wird "geflippt"

(von 0 nach 1 bzw. von 1 nach 0).

Das ergibt die Irrfahrt auf der Menge der Würfelecken

$\{0, 1\}^3$.

Diese hat ein reversibles Gleichgewicht:

die uniforme Verteilung.

Vom Feinmodell zum Ehrenfest-Modell kommt man durch
“Zählen der Teilchen links”:

$$l(a) := a_1 + a_2 + a_3.$$

Ist $Z = (Z^{(1)}, Z^{(2)}, Z^{(3)})$ uniform verteilt auf $\{0,1\}^d$,
dann ist $Z^{(1)} + Z^{(2)} + Z^{(3)}$ Binomial($3, \frac{1}{2}$)-verteilt.

Zur Probe: Die Binomial$(3, \frac{1}{2})$-Verteilung

ist ein reversibles Gleichgewicht für das Ehrenfest-Modell

mit $d = 3$ Kugeln:

$$2^{-3}\binom{3}{\ell}\,\frac{3-\ell}{3} = 2^{-3}\binom{3}{\ell+1}\,\frac{\ell+1}{3}. \;\Box$$

Was 3 recht ist, ist einem allgemeinen d billig -

siehe Buch S. 109-110:

Die Binomial$(d, \frac{1}{2})$-Verteilung

ist ein reversibles Gleichgewicht

für das Ehrenfest-Modell mit d Kugeln.

Zur Vorbereitung auf V12b “Schätzen mit Verlass”

rekapitulieren wir Inhalte aus der

Vorlesung 8a “ZGWS und Mittelwerte”

7. Populationsmittelwert und Stichprobenmittelwert

Denken wir an eine Liste (eine “Population”)

von reellen Daten

$$w_1, \ldots, w_g$$

Angenommen man möchte den Populationsmittelwert

$$\mu := \frac{1}{g} \sum_{j=1}^{g} w_j.$$

schätzen,

und zwar aus den Werten einer

aus der Population gezogenen Stichprobe

$$x_1, \ldots, x_n.$$

Als *Schätzwert* für μ bietet sich an:

$$m := \frac{1}{n}(x_1 + \ldots + x_n)$$

Wie zuverlässig ist diese Schätzung?

Goldene Idee der Statistik:

Man fasst $x_1, \ldots, x_n$ auf als Ergebnis eines rein zufälligen Ziehens aus der Population:

$$X_1 := w_{J_1},\ X_2 := w_{J_2}, \ldots$$

mit $J_1, J_2, \ldots$ rein zufällige Wahl aus $\{1, \ldots, g\}$ ("Ziehen mit Zurücklegen").

$$m = \frac{1}{n}(x_1 + \ldots + x_n)$$

fasst man also auf

als *eine* Realisierung (*einen* Ausgang)

der **Zufallsvariable**

$$M := \bar{X} := \frac{1}{n}(X_1 + \ldots + X_n)$$

(Mittelwert der zufälligen Stichprobe $(X_1, \ldots, X_n)$)

Es gilt jedenfalls:

$$\mathbf{E}[M] = \mu$$

Der Erwartungswert des Stichprobenmittelwertes

ist gleich dem Populationsmittelwert.

8. Populationsvarianz und Varianz des Stichprobenmittelwertes

Der Populationsmittelwert war gleich $\mathbf{E}[X_1] = \mu$.

Wir haben

$$\mathbf{Var}[X_1] = \frac{1}{g} \sum_{j=1}^{g} (w_j - \mu)^2 =: \sigma^2$$

Diese Zahl σ^2 heißt auch die **Populationsvarianz**.

Wir hatten

$$M = \frac{1}{n}(X_1 + \ldots + X_n),$$

$$\mathbf{E}[M] = \mu.$$

$$M = \frac{1}{n}(X_1 + \ldots + X_n)$$

In unserer Vorstellung sind die X_i unabhängig,

also ergibt sich

$$\mathbf{Var}[M] = \frac{\sigma^2}{n}.$$

Die Standardabweichung des Stichprobenmittelwertes M

ist also

$$\sigma_M = \frac{\sigma}{\sqrt{n}}.$$

9. Approximative Verteilung des Stichprobenmittelwertes

Wie ist (für nicht zu kleines n)

der Stichprobenmittelwert M verteilt?

Der Zentrale Grenzwertsatz gibt eine Antwort:

In der oben beschriebenen Situation gilt

M ist approximativ N$(\mu, \frac{\sigma^2}{n})$-verteilt.

10. Die Stichprobenvarianz als Schätzung für die Populationsvarianz

Ein Problem in der Praxis: Im Allgem. kennt man σ^2 nicht.

Auch σ^2 muss man dann schätzen.

Zwei Vorschläge für die

(aus der Stichprobe) **geschätzte** (Populations-)**Varianz**:

(i) die *Stichprobenvarianz*

$$\hat{\sigma}^2 := \frac{1}{n} \sum_{i=1}^{n} (x_i - m)^2.$$

(ii) die *modifizierte Stichprobenvarianz*

$$s^2 := \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2$$

Es gibt theoretische Begründung für beide Vorschläge

(vgl. Buch S. 124, S. 138).

Wir halten uns hier an den Vorschlag (ii):

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - m)^2$$

Die Standardabweichung

des Stichprobenmittelwertes $M = \bar{X}$ ist $\frac{\sigma}{\sqrt{n}}$.

Die geschätzte Standardabweichung

des Stichprobenmittelwertes $M = \bar{X}$ ist

$$\boxed{s/\sqrt{n} =: f}$$

Diese Größe nennen wir auch den *Standardfehler*.

M ist approximativ $\mathrm{N}(\mu, \frac{\sigma^2}{n})$-verteilt.

Und (gut für die Praxis):

M ist approximativ $\mathrm{N}(\mu, f^2)$-verteilt.

Auf den Spuren von Aufgabe 33 stellen wir fest:

$$I := [M - 2f, M + 2f]$$

ist ein zufälliges Intervall,

welches den Populationsmittelwert μ

mit Wahrscheinlichkeit ≈ 0.95 überdeckt.

(Denn das Ereignis $\{|M - \mu| \leq 2f\}$

hat Wahrscheinlichkeit ≈ 0.95.)

I heißt approximatives 95% Konfidenzintervall für μ.