

## T-61.281 Luonnollisen kielen tilastollinen käsittely

Harjoitus 6, ti 2.3.2004, 8:30-10:00 Markov-ketjut ja kätkeyt Markov-mallit, Versio 1.0

1. Kauniissa Turun kaupungissa on tutkittu sään vaihtelua. Joka päivä 12:00 on Sampalinnanmäeltä katsottu taivaalle ja kirjattu onko päivä aurinkoinen ( $S_1$ ), pilvinen ( $S_2$ ) vai sateinen ( $S_3$ ). Näistä havainnoista on laskettu seuraavat siirtymätodennäköisyydet:

$$A = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.4 & 0.5 & 0.1 \\ 0.3 & 0.3 & 0.4 \end{bmatrix}$$

Tässä taulukon alkio  $a_{ij}$  kuvaa siirtymätodennäköisyyttä tilasta  $S_i$  tilaan  $S_j$ . Esimerkiksi todennäköisyys, että pilvistä päivää seuraa aurinkoinen on 0.4.

- a) Piirrä Turun säätila Markov-ketjuna
  - b) Tänään on pilvistä. Millä todennäköisyydellä seuraava viiden päivä sääsekvenssi seuraa: huomenna sataa, ylihuomenna on pilvistä, sitten tulee 3 aurinkoista päivää ?
  - c) Kuinka monta aurinkoista päivää peräkkäin on keskimäärin Turussa ?
2. Olet kolmen päivän lomalla Hawajilla. Sinua rupeaa suunnattomasti kiinnostamaan, millainen sää Turussa on. Paikallisissa lehdissä on kuitenkin kerrottu vain Turun seudun lämpötila. Osaat kuitenkin arvioida esim. millä todennäköisyydellä  $P(x < -5^\circ C \mid q = S_1) = b_1(x \leq -5^\circ C)$  aurinkoisena päivänä on yli 5 astetta pakkasta. Tarkemmin ilmaistuna

$$\begin{aligned} b_1(x \leq -5^\circ C) &= 0.8 \\ b_1(-5^\circ C \leq x \leq 5^\circ C) &= 0.05 \\ b_1(x \geq 5^\circ C) &= 0.15 \\ b_2(x \leq -5^\circ C) &= 0.1 \\ b_2(-5^\circ C \leq x \leq 5^\circ C) &= 0.7 \\ b_2(x \geq 5^\circ C) &= 0.2 \\ b_3(x \leq -5^\circ C) &= 0.3 \\ b_3(-5^\circ C \leq x \leq 5^\circ C) &= 0.4 \\ b_3(x \geq 5^\circ C) &= 0.3 \end{aligned}$$

Lähtiessäsi lomalle Turusta, siellä oli aurinkoista.

- a) Kolme päivää kuluu, loma loppuu ja on aika palata takaisin. Haluat arvioida, millainen sää palatessasi on, jotta osaat pukeutua paluumatkalle oikein. Lämpötilat Turussa olivat  $7^\circ C$ ,  $3^\circ C$  ja  $-8^\circ C$ . Laske siis kunkin säätilan todennäköisyydet paluupäivänäsi. Ratkaisussa suositellaan käytettäväksi forward-algoritmia, brutaalilla voimalla kaikkien ratkaisujen läpikäynti lienee hieman työlästä.
- b) Lentokoneessa matkaalla takaisin pohdiskelit aikasi kuluksi, minkälainen sää on lomalla ollessasi Turussa ollut. Etsi siis todennäköisin tilasekvenssi. Viterbihaulla voi tämän tehdä vähimmällä määrällä laskutoimituksia.