

1. Tapahtumat ja todennäköisyys

- ★ Perusjoukko (sample space) Ω on mv. (mahd. ääretön) joukko
- ★ Ω :n alkiot ovat alkeistapauksia ja niiden osajoukkoa $E \subseteq \Omega$ nimitetään *tapahtumaksi* (event)
- ★ Aina ei olla kiinnostuneita kaikista potenssijoukon $\mathcal{P}(\Omega)$ tapahtumista, vaan ainoastaan hyvinmääritellystä tapahtumien osakokoelmasta
- ★ σ -kenttä (Ω, \mathcal{F}) sisältää osajoukkojen kokoelman (σ -algebran) $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ s.e.
 1. $\emptyset \in \mathcal{F}$
 2. $E \in \mathcal{F} \Rightarrow \bar{E} \in \mathcal{F}$, missä $\bar{E} = \Omega \setminus E$
 3. $E_1, E_2, \dots \in \mathcal{F} \Rightarrow E_1 \cup E_2 \cup \dots \in \mathcal{F}$

11

- ★ σ -kentän (Ω, \mathcal{F}) todennäköisyysmitta $\text{Pr}: \mathcal{F} \rightarrow \mathbb{R}^+$ on kuvaus s.e.
 - ◇ $0 \leq \text{Pr}(E) \leq 1$ kaikille $E \in \mathcal{F}$
 - ◇ $\text{Pr}(\Omega) = 1$
 - ◇ Toisensa poissulkeville tapahtumille E_1, E_2, \dots ,
 $\text{Pr}(\bigcup_i E_i) = \sum_i \text{Pr}(E_i)$
- ★ Yhdessä $(\Omega, \mathcal{F}, \text{Pr})$ määrittävät *todennäköisyysavaruuden* (probability space)
- ★ Useimmiten tarkastelemme *diskreettejä* todennäköisyysavaruuksia, joissa Ω on äärellinen tai numeroituvasti ääretön
- ★ Tällöin kaikkien alkeistapauksen $\omega \in \Omega$ todennäköisyydet voidaan antaa ja sallittujen tapahtumien σ -algebra $\mathcal{F} = \mathcal{P}(\Omega)$

12

ESIMERKKEJÄ

- ★ Yhden nopan heitossa on kuusi mahdollista alkeistapausta silmäluvun mukaan
- ★ Kunkin silmäluvun todennäköisyys painottamattomalla nopalla on $1/6$
- ★ Alkeistapaukset ovat toisensa poissulkevia, joten niiden todennäköisyydet määräävät todennäköisyyden $\text{Pr}(E)$ kaikille $E \in \mathcal{P}(\Omega)$
- ★ Esim. $\text{Pr}(\text{silmäluku} \leq 3) = 3 \cdot 1/6 = 3/6$
- ★ Tämä määrää perusjoukon *tasaisen* todennäköisyysavaruuden, missä $\text{Pr}(E) = |E|/6$ kaikilla $E \subseteq \Omega$
- ★ Esim. $\text{Pr}(\text{pariton silmäluku}) = 3/6$

13

- ★ Kahden korkeintaan astetta d olevan polynomin $F(x)$ ja $G(x)$ ekvivalenssin tarkistaminen voidaan tehdä deterministisesti $\Theta(d^2)$ ajassa muuntamalla ne kanoniseen muotoon ja vertaamalla tekijöiden (kokonaisluku)kertoimia
- ★ Nopeampi satunnaisalgoritmi perustuu huomioon, että $F(y) - G(y) = 0$ kaikilla $x = y$ kun $F(x) = G(x)$
- ★ Toisaalta kun $F(x) \neq G(x)$, niin polynomilla $F(x) - G(x)$ on korkeintaan d juurta
- ★ Jos x valitaan joukosta $\{1, \dots, md\}$, $m \in \mathbb{N}$, niin on vähint. $(m-1)d$ mahdollisuutta löytää arvo, jolla $F(x) - G(x) \neq 0$

14

- ★ Vedetään siis luku $r \in \{1, \dots, md\}$ tasaisen jakauman mukaan
- ★ Lasketaan $F(r)$ ja $G(r)$ ajassa $O(d)$
 - ◊ Jos $F(r) = G(r)$, tulosta "samat"
 - ◊ Muuten tulosta "eri"
- ★ Vain jos $F(x) \neq G(x)$ voi edellinen menetelmä antaa väärän vastauksen
- ★ Algoritmin virhe on siis *yksipuolinen* (**one-sided error**)
- ★ Virheellinen vastaus annetaan kun r on polynomin $F(x) - G(x)$ juuri
- ★ Juuria on korkeintaan d ja arvovaihtoehtoja md , joten virheen todennäköisyys on kork. $d/md = 1/m$
- ★ Valitsemalla m sopivasti voidaan säätää oikean vastauksen todennäköisyyttä:
 $m = 100 : 99\%$ tai $m = 1000 : 99,9\%$

15

Yhdisteen todennäköisyys

- ★ Suoraan määritelmien perusteella

Lemma 1.1: Mille tahansa kahdelle tapahtumalle E_1 ja E_2 pätee $\Pr(E_1 \cup E_2) =$

$$\Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2).$$

- ★ Yksinkertainen mutta varsin käyttökelpoinen todennäköisyyden aksiomien seuraus on

Lemma 1.2 (union bound): Mille tahansa tapahtumien äärelliselle tai numeroituvalle jonolle E_1, E_2, \dots pätee

$$\Pr\left(\bigcup_{i \geq 1} E_i\right) \leq \sum_{i \geq 1} \Pr(E_i).$$

- ★ Tässä tapahtumien ei siis tarvitse olla toisensa poissulkevia

17

- ★ Ylinumeroituvan \mathbb{R} kaikille osajoukoille ei voida määritellä todennäköisyyksiä rikkomatta aksiomia
- ★ Useimmiten riittää *Borel-joukkojen* käsittely
- ★ Olk. \mathcal{F} suppein \mathbb{R} :n σ -algebra, joka sisältää kaikki suljetut välit $[a, b]$, $a, b \in \mathbb{R}$
- ★ \mathcal{F} :n alkiot ovat Borel-joukot
- ★ Välin $[a, b]$ todennäköisyydeksi määritellään välin $[a, b] \cap [0, 1]$ pituus

$$\Pr([a, b]) = \min(b, 1) - \max(a, 0)$$
- ★ Tämä on osavälin $[0, 1]$ tasainen todennäköisyyssmitta

16

- ★ Lemman 1.1 yleistyminen n :lle tapahtumalle on

Lemma 1.3 (inclusion-exclusion principle): Olk. E_1, \dots, E_n mielivaltaiset n tapahtumaa.

$$\begin{aligned} \Pr\left(\bigcup_{i=1}^n E_i\right) &= \sum_{i=1}^n \Pr(E_i) - \sum_{i < j} \Pr(E_i \cap E_j) \\ &\quad + \sum_{i < j < k} \Pr(E_i \cap E_j \cap E_k) \\ &\quad - \dots + (-1)^{\ell+1} \sum_{i_1 < \dots < i_\ell} \Pr\left(\bigcap_{r=1}^{\ell} E_{i_r}\right) + \dots \end{aligned}$$

18

Bonferronin epäytälöt

- ★ Yhdisteen tarkka tn. voidaan ilmaista kompaktimmin muodossa

$$\Pr\left(\bigcup_{i=1}^n E_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{|J|=k} \Pr\left(\bigcap_{j \in J} E_j\right)$$

- ★ Lakemalla summa vain rajaan $\ell < n$ saakka, saadaan vuorotellen ylä- ja alarajoja

- ★ ℓ pariton:

$$\Pr\left(\bigcup_{i=1}^{\ell} E_i\right) \leq \sum_{k=1}^{\ell} (-1)^{k+1} \sum_{|J|=k} \Pr\left(\bigcap_{j \in J} E_j\right)$$

- ★ ℓ parillinen:

$$\Pr\left(\bigcup_{i=1}^{\ell} E_i\right) \geq \sum_{k=1}^{\ell} (-1)^{k+1} \sum_{|J|=k} \Pr\left(\bigcap_{j \in J} E_j\right)$$

19

Riippumattomuus

- ★ Tapahtumat E ja F ovat riippumattomia, joss

$$\Pr(E \cap F) = \Pr(E) \cdot \Pr(F)$$

- ★ Yleistäen: tapahtumat E_1, \dots, E_k ovat toisistaan riippumattomia, joss mille tahansa osajoukolle $I \subseteq [1, k]$

$$\Pr\left(\bigcap_{i \in I} E_i\right) = \prod_{i \in I} \Pr(E_i)$$

- ★ Tapahtuman E todennäköisyys ehdolla F , $\Pr(F) > 0$, on

$$\Pr(E | F) = \frac{\Pr(E \cap F)}{\Pr(F)}$$

- ★ Jos $\Pr(F) > 0$, niin E ja F ovat riippumattomia joss $\Pr(E | F) = \Pr(E)$

20

- ★ Toinen mahdollisuus parantaa polynomien ekvivalenssin tunnistamisen onnistumistn:ä on testata onko $F(r) = G(r)$ useilla arvoilla r
- ★ Toista $i = 1, \dots, k$ kertaa
 - ◇ Vedä satunnainen $r_i \in \{1, \dots, md\}$
 - ◇ Jos $F(r_i) \neq G(r_i)$, tulosta "eri"
- ★ Tulosta "samat"
- ★ Kun luvut r_i valitaan takaisinpanolla, on jokaisella kierroksella juuren löytämistn. kork. $1/m$
- ★ Tulos on virheellinen vain kun jokainen k :sta kierroksesta löytää juuren
- ★ Virhetodennäköisyys on kaikkiaan siis kork. $(1/m)^k$
- ★ Virhetn. pienenee eksponentiaalisesti kierrosten lukumäärän suhteen

21

- ★ Ilman takaisinpanoa toimiminen pienentää virheen tn:ttä hieman
- ★ Jos kierroksilla $1, \dots, j-1$ on löydetty juuri, niin mahdollisia juuria on enää $d - (j-1)$ jäljellä kun mahdollisia valintoja on vielä $md - (j-1)$
- ★ Olk. tapahtuma E_i "kierroksella i valittu r_i on polynomin $F(x) - G(x)$ juuri"

$$\Pr(E_j | E_1 \cap \dots \cap E_{j-1}) \leq \frac{d - (j-1)}{md - (j-1)}$$

- ★ Yli kaikkien $k < d$ kierroksen virhetn:ttä rajoittaa

$$\begin{aligned} \Pr(E_1 \cap \dots \cap E_k) &\leq \prod_{j=1}^k \frac{d - (j-1)}{md - (j-1)} \\ &\leq \left(\frac{1}{m}\right)^k \end{aligned}$$

22

Kokonaistodennäköisyys

Lause 1.6: Olk. E_1, \dots, E_n kokoelma toisensa poissulkevia perusjoukon Ω tapahtumia s.e. $\bigcup_{i=1}^n E_i = \Omega$. Tällöin

$$\begin{aligned}\Pr(B) &= \sum_{i=1}^n \Pr(B \cap E_i) \\ &= \sum_{i=1}^n \Pr(B | E_i) \Pr(E_i).\end{aligned}$$

- ★ Tapahtuman B todennäköisyys voidaan siis jakaa osiin
- ★ Ns. *viivästetyn valinnan* periaate (**principle of deferred decision**)
- ★ Jos esim. satunnaisvektorissa $\vec{r} = (r_1, \dots, r_n)$ on komponentteina n satunnaismuuttujaa r_i , niin voi olla hyödyllistä ajatella

23

- ◇ Jonkin/joidenkin arvojen olevan kiinnitettyjä tietyllä hetkellä,
- ◇ muiden s-muuttujien arvojen kiinnittämistä viivästyttävän ja
- ◇ todennäköisyyksien olevan ehdollistettuja kiinnitetyille arvoille

MATRIISITULO

- ★ Annettuna kolme binääristä $n \times n$ -neliomatriisia \mathbf{A} , \mathbf{B} ja \mathbf{C} , halutaan tarkkistaa päteekö $\mathbf{AB} = \mathbf{C}$
- ★ Suoraviivainen deterministinen algoritmi vaatii $\Theta(n^3)$ ajan ja erikoistuneempi n. $\Theta(n^{2.37})$ ajan
- ★ Valitaan satunnainen vektori $\vec{r} = (r_1, \dots, r_n) \in \{0, 1\}^n$ tasaisen jakouman mukaan

24

- ★ Kukin r_i voidaan valita riippumattomasti ja tasaisesti joukosta $\{0, 1\}$, sillä tällöin jokaisella 2^n vektorilla \vec{r} on tn. 2^{-n} tulla valituksi
- ★ Lasketaan $\mathbf{B}\vec{r}$ ja sitten $\mathbf{A}(\mathbf{B}\vec{r})$ sekä $\mathbf{C}\vec{r}$
- ★ Jos $\mathbf{A}(\mathbf{B}\vec{r}) \neq \mathbf{C}\vec{r}$, tulosta "erisuuret"
- ★ Muuten tulosta "yhtäsuuret"
- ★ Tarvitaan kolme matriisiin ja vektorin tuloa, jotka vaativat $\Theta(n^2)$ ajan
- ★ Virhetilanteessa $\mathbf{D} = \mathbf{AB} - \mathbf{C} \neq 0$
 - ◇ Tällöin $\mathbf{AB}\vec{r} = \mathbf{C}\vec{r}$ implikoi, että $\mathbf{D}\vec{r} = 0$
 - ◇ Koska $\mathbf{D} \neq 0$, niin siinä on oltava nollasta poikkeava alkio
 - ◇ Voidaan tarkastella tilannetta, jossa tuo alkio on d_{11}

25

- ◇ Koska $\mathbf{D}\vec{r} = 0$, niin
$$\sum_{j=1}^n d_{1j}r_j = d_{11}r_1 + \sum_{j=2}^n d_{1j}r_j = 0$$
- ◇ Toisin ilmaisten

$$r_1 = \frac{-\sum_{j=2}^n d_{1j}r_j}{d_{11}}$$

- ◇ Valitaan nyt \vec{r} :n komponenttiarvot riippumattomasti ja satunnaisesti järjestyksessä r_n, \dots, r_1
- ◇ Juuri ennen arvon r_1 kiinnittämistä yo. kaavan oikean puolen s-muuttujien arvot on kiinnitetty
- ◇ Kork. toinen r_1 :n mahdollisista arvoista voi toteuttaa yhtälön
- ◇ Tämän ja samalla ekvivalenssin $\mathbf{AB}\vec{r} = \mathbf{C}\vec{r}$ tn. on kork. $1/2$

26

- ★ Edellisen perusteella siis

Lause 1.4: Jos $\mathbf{AB} \neq \mathbf{C}$ ja kun \vec{r} on valittu tasaisen jakauman mukaan joukosta $\{0, 1\}^n$, niin

$$\Pr(\mathbf{AB}\vec{r} = \mathbf{C}\vec{r}) \leq \frac{1}{2}$$

- ★ Algoritmin virhe on yksipuolinen \Rightarrow oikean vastauksen tn:ttä voidaan parantaa toistamalla sitä useilla satunnaisilla \vec{r}
- ★ Takaisinpanolla valitut k vektoria pudottavat virhetn:n arvoon 2^{-k} mutt'eivät kasvata ajoaikaa kuin luokkaan $\Theta(kn^2)$
- ★ Virhetn. voidaan painaa mielivaltaisen alas algoritmin asymptoottisen aikavaativuuden kasvamatta

27

Bayesin sääntö

Lause 1.7: Olk. E_1, \dots, E_n toisensa poissulkevia. Tällöin

$$\begin{aligned} \Pr(E_j | B) &= \frac{\Pr(E_j \cap B)}{\Pr(B)} \\ &= \frac{\Pr(B | E_j) \Pr(E_j)}{\sum_i \Pr(B | E_i) \Pr(E_i)} \end{aligned}$$

- ★ Säännöllä usein päivitetään uskomuksia, uusien havaintojen jälkeen
 - ◇ $\Pr(E_j)$ on teorian E_j a priori -tn., uskomme siihen ennen havaintoja
 - ◇ $\Pr(B | E_j)$ mittaa kuinka hyvin E_j "selittää" havainnon B
 - ◇ $\Pr(E_j | B)$ on teorian E_j a posteriori -tn., uskomme siihen havaintojen B jälkeen

28

- ★ Annettuna 3 kolikkoa, joista 2 tiedetään olevan tasapainoisia ja 1 painotettu s.e. kruunan tn. on $2/3$

- ★ Heitetään mv. järjestyksessä olevia kolikkoja tuloksena $B = (1: \text{kruuna}, 2: \text{kruuna}, 3: \text{klaava})$

- ★ Mikä on tn., että ensimmäinen kolikko on painotettu?

- ★ Olk. $E_i =$ "is kolikko on painotettu"

$$\Pr(E_1) = \Pr(E_2) = \Pr(E_3) = \frac{1}{3}$$

$$\Pr(B | E_1) = \Pr(B | E_2) = \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{6}$$

$$\Pr(B | E_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{12}$$

$$\Pr(E_1 | B) = \frac{\Pr(B|E_1)\Pr(E_1)}{\sum_{i=1}^3 \Pr(B|E_i)\Pr(E_i)} = \frac{2}{5}$$

- ★ Havainnon perusteella tn., että 1. kolikko on painotettu kasvaa $\frac{1}{3} \rightarrow \frac{2}{5}$

29

Minimileikkausalgoritmi

- ★ Tarkastellaan yhtenäistä, suuntaamatonta moniverkkoa (**multigraph**) $G = (V, E)$, jossa on n solmua

- ★ Moniverkossa voi olla useita kaaria kahden solmun välillä

- ★ Verkon G leikkaus on joukko kaaria $C \subseteq E$, joiden poistaminen johtaa siihen että $(V, E \setminus C)$ ei ole enää yhtenäinen

- ★ Minimileikkaus on pienimmän määrän kaaria sisältävä leikkaus

- ★ Tarkastelemme yksinkertaista satunnaisalgoritmia minimileikkauksen löytämiseksi

30

- ★ **Toista:** valitse kaari satunnaisesti (tasaisten jakauman mukaan) ja yhdistä solmut, jotka kaari yhdistää
- ★ Jos kahden solmun välillä on useita kaaria, niin ne kaikki säilytetään
- ★ Yhdistettävien solmujen väliset kaaret puolestaan hävitetään; verkossa ei ole silmukoita solmusta itseensä
- ★ Kahden solmun yhdistäminen pienentää G :n solmujen lukumäärää yhdellä
- ★ Keskeinen havainto: solmujen yhdistäminen ei pienennä minimileikkauksen kokoa
 - ◇ Mikä tahansa muutetun verkon leikkaus on myös alkuperäisen verkon leikkaus

31

- ★ Solmujen yhdistämistä jatketaan kunnes jäljellä on enää kaksi solmua
- ★ Niiden väliset kaaret ovat siis verkon G leikkaus (ei kuitenkaan välttämättä minimaalinen)
- ★ Olk. k minimileikkauksen koko
- ★ Kiinnitetään jokin minimileikkaus C
- ★ Verkossa G on oltava vähintään $kn/2$ kaarta, sillä muuten siinä on solmu, jonka aste on vähemmän kuin k .
- ★ Tällöin solmuun liittyvät kaaret olisivat leikkaus, jossa on vähemmän kuin k kaarta.

32

- ★ Rajoitamme (alhaalta) todennäköisyyden, ettei yhtään C :n kaarista valita solmujen yhdistämisessä, jolloin lopulta jäljelle jäävät kaaret olisivat täsmälleen leikkauksen C muodostavat kaaret
- ★ Olk. tapahtuma E_i , $1 \leq i \leq n-2$, "askelella i ei valita C :hen kuuluvaa kaarta"
- ★ Ensimmäisellä kierroksella todennäköisyys, että satunnaisesti valittu kaari kuuluu C :hen on korkeintaan $k/(nk/2) = 2/n$, joten $\Pr(E_1) \geq 1 - 2/n$
- ★ Jos E_1 toteutuu, niin 2. kierroksella on jäljellä vähintään $k(n-1)/2$ kaarta, joten $\Pr(E_2 | E_1) \geq 1 - 2/(n-1)$

33

- ★ Kierroksella i solmuja on jäljellä $n-i+1$ ja kaaria siten vähintään $k(n-i+1)/2$. Näin ollen

$$\Pr(E_i | \cap_{j=1}^{i-1} E_j) \geq 1 - \frac{2}{n-i+1}$$

- ★ Tapahtumien leikkauksen todennäköisyydelle pätee laskusääntö

$$\Pr(\cap_{i=1}^k E_i) = \Pr(E_1) \cdot \Pr(E_2 | E_1) \cdots \Pr(E_k | \cap_{i=1}^{k-1} E_i)$$

- ★ Nyt voimme laskea tn:n, ettei yhtään C :n kaarta valita algoritmissa:

$$\begin{aligned} \Pr\left(\bigcap_{i=1}^{n-2} E_i\right) &\geq \prod_{i=1}^{n-2} \left(1 - \frac{2}{n-i+1}\right) \\ &= \left(\frac{n-2}{n}\right) \left(\frac{n-3}{n-1}\right) \cdots \left(\frac{1}{3}\right) \\ &= \frac{2}{n(n-1)} \end{aligned}$$

34

★ Yksittäisen (mahdollisesti ainoan) minimileikkauksen löytymistä. Siis on suurempi kuin $2/n^2$, mutta algoritmin tulostama leikkaus ei välttämättä ole minimaalinen

$$\star \left(1 + \frac{t}{n}\right)^n \leq e^t \quad \forall t, n \in \mathbb{R}^+$$

$$\star 1 - x \leq e^{-x}$$

★ Toistamalla minimileikkausalgoritmia $(n-1)n \ln n$ kertaa saadaan todennäköisyydeksi, ettei minimileikkausta löydy yhdelläkään suorituksista, korkeintaan

$$\begin{aligned} \left(1 - \frac{2}{n(n-1)}\right)^{n(n-1) \ln n} &\leq e^{-2 \ln n} \\ &= \frac{1}{n^2} \end{aligned}$$

35

★ Riippumattomin toistoin epäonnistumisen todennäköisyyttä saatiin pienennettyä huomattavasti (ajankäytön kustannuksella)

★ Algoritmilta voi kuitenkin edelleen jäädä minimileikkaus löytämättä

★ Deterministiset (verkon vuohon perustuvat) algoritmit ovat yleensä huomattavasti monimutkaisempia

★ Edellä esitetyn algoritmin variantin odotusarvoinen aikavaativuus on huomattavasti pienempi kuin parhaan vuohon perustuvan algoritmin

36