Johdatus todennäköisyyslaskentaan
Kevät 2014

Luento 4 / 13

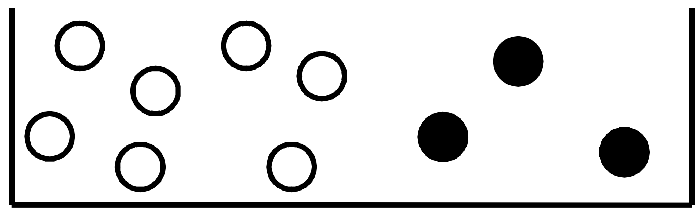
Jukka Kohonen

Matematiikan ja tilastotieteen laitos

Helsingin yliopisto

Ilmoitusasioita

- Mikroluokan C128 varaustilanne löytyy seuraavasti:
  laitoksen sivu → Hallinto → ATK-luokan (C128) varaustilanne
- Mikroluokkaa voi käyttää silloin kun se on varattu **tämän kurssin ohjaukseen** (ks. kurssisivu).
- Luokkaa voi myös käyttää **omatoimiseen opiskeluun** silloin **kun se ei ole varattu opetukseen**.
- Matlab löytyy myös jostain muusta mikroluokasta (otan selvää)
- Jos tietokonetehtäviä haluaa tehdä omalla koneella, voi käyttää ilmaista Octave-ohjelmaa https://www.gnu.org/software/octave/
  (jotkut Matlab-koodit toimivat siinä sellaisenaan, mutta eivät kaikki)



OTANTA
TAKAISINPANOLLA JA ILMAN

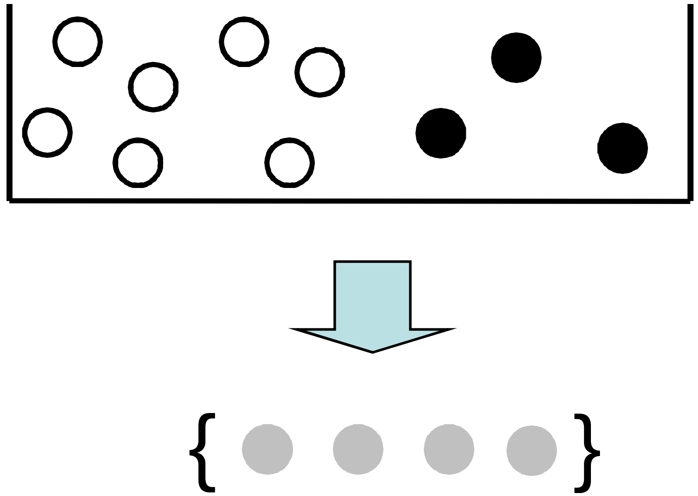
Otanta ”ilman takaisinpanoa”

- Populaatio, jossa N alkiota (palloa, ihmistä tms.), kahdenlaisia (”valkoinen”, ”musta”)
- Poimitaan umpimähkään (= symmetrisesti) n-osajoukko eli kombinaatio eli **otos**

- Merkitään tapahtuma $\boldsymbol{A_k}$ **= ”otoksessa on *k* valkoista palloa”**
- Tapahtumaan A_k sisältyy monta alkeistapausta (mitkä valkoiset, mitkä mustat pallot otoksessa)

N=10 palloa yhteensä

K=7 valkoista N-K=3 mustaa



otos n=4 palloa

Montako valkoista?

Suotuisat alkeistapaukset

k valkoista valitaan K:sta

n-k mustaa valitaan N-K:sta

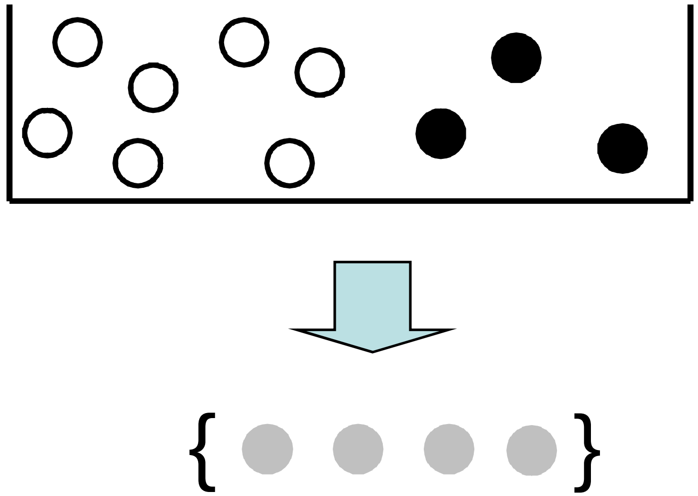
$$P(A_k) = \frac{\binom{K}{k} \times \binom{N-K}{n-k}}{\binom{N}{n}}$$

Kaikki alkeistapaukset
= erilaiset n-otokset

Otanta ”ilman takaisinpanoa”

N=10 palloa yhteensä

K=7 valkoista N-K=3 mustaa



otos n=4 palloa

$$P(A_2) = \frac{\binom{7}{2} \times \binom{3}{2}}{\binom{10}{4}} = \frac{21 \times 3}{210} = 0{,}3$$

Samalla kaavalla

P(”0 valkoista”) = 0 (miksi?)

P(”1 valkoinen”) ≈ 0,033

P(”2 valkoista”) = 0,300

P(”3 valkoista”) = 0,500

P(”4 valkoista”) ≈ 0,167

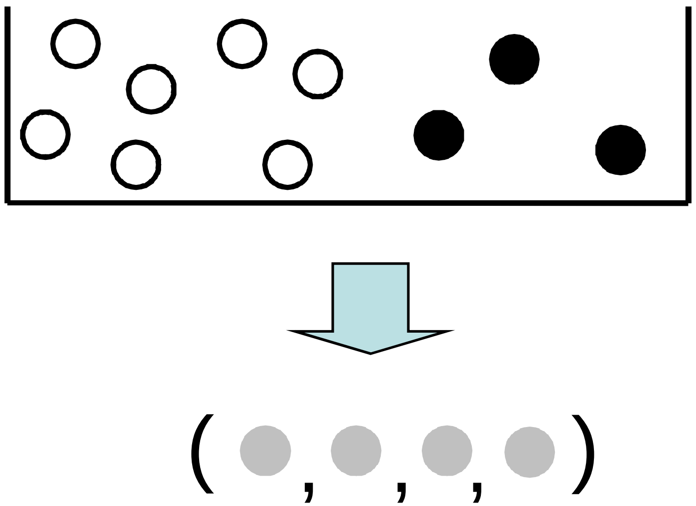
Otanta ”takaisinpanolla”

- Populaatio kuten edellä. Poimitaan umpimähkään yksi pallo, **pannaan takaisin**, otetaan umpimähkään toinen pallo jne.
- Ei välttämättä konkreettista ”takaisinpanoa”, olennaista on että **poimitaan joka kerta uudestaan samasta populaatiosta** (esim. puhelinluettelosta ihmisiä)

- Sama pallo voi tulla otokseen monta kertaa! Osajoukko ei mielekäs malli otokselle.
- Ajatellaan otos järjestetyksi n-jonoksi <u>jossa sama alkio saa esiintyä monta kertaa</u>.
- Montako valkoista palloa otoksessa?

N=10 palloa yhteensä

K=7 valkoista N-K=3 mustaa



Suotuisat alkeistapaukset

k:n valkoisen sijainti n-jonossa

k valkoista valitaan K:sta

n-k mustaa valitaan N-K:sta

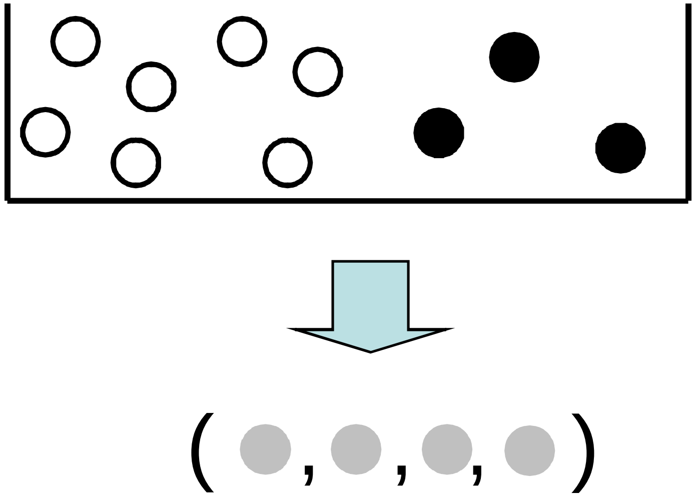
$$P(A_k) = \binom{n}{k} \frac{K^k (N-K)^{n-k}}{N^n}$$

Kaikki alkeistapaukset = erilaiset n-jonot. Joka kerta samat N vaihtoehtoa!

Otanta ”takaisinpanolla”

N=10 palloa yhteensä

K=7 valkoista *N*-*K*=3 mustaa



Suotuisat alkeistapaukset

2 valkoisen sijainti 4-jonossa

2 valkoista valitaan 7:stä

2 mustaa valitaan 3:sta

$$P(A_2)=\binom{4}{2}\frac{7^2\times 3^2}{10^4}\approx 0{,}265$$

Kaikki alkeistapaukset
= erilaiset 4-jonot.
Joka kerta
samat 10 vaihtoehtoa!

Yleensä eri todennäköisyys

7 valkoista ja 3 mustaa, otos n=4

k	P("k valkoista") ilman takaisinpanoa	P("k valkoista") takaisinpanolla
0	0,000	0,008 (miksi > 0 ?)
1	0,033	0,076
2	0,300	0,265
3	0,500	0,412
4	0,167	0,240

Esimerkki: Hyvin suuri populaatio

- N = 5 000 000 (suomalaiset)
- K = 500 000 (helsinkiläiset)
- n = 3 (otoskoko) $n << K$ ja $n << N\text{-}K$

- Millä tn:llä otokseen osuu tasan yksi helsinkiläinen?
- Ilman takaisinpanoa

$$P(A_1) = \frac{\binom{500\,000}{1} \times \binom{4\,500\,000}{2}}{\binom{5\,000\,000}{3}} \approx 0{,}243\,000\,094$$

- Takaisinpanolla

$$P(A_1) = \binom{3}{1} \frac{500\,000^1 \times 4\,500\,000^2}{5\,000\,000^3} = 0{,}243\,000\,000$$

Lotossa P(tasan *k* oikein)

- Asiakas valitsee 7 numeroa. Kone arpoo 7 numeroa.
- Montako samaa tulee millä todennäköisyydellä?

= Otanta ilman takaisinpanoa. 39 palloa,
  - 7 palloa on ”mustia” (joiden numero on asiakkaan rivissä)
  - 32 palloa on ”valkoisia” (joiden numero ei ole asiakkaan rivissä)
  - Kone nostaa 7 palloa, montako niistä ”mustia”?
- Esim. P(asiakkaalla on viisi oikein) =

Otoksessa oltava 5 mustaa palloa (7:stä)

$$\frac{\binom{7}{5}\binom{32}{2}}{\binom{39}{7}}$$

Otoksessa oltava 2 valkoista palloa (32:sta)

Perusjoukko: 7 pallon otos 39:stä

P(tasan *k* oikein)

k	Suotuisia rivejä (kpl)		Tn
0	3 365 856		0,219
1	6 343 344		0,412
2	4 228 896		0,275
3	1258 600		0,081 8
4	173 600		0,011 3
5	10 416		0,000 677
6	224		0,000 014 6
7	1		0,000 000 065
yhteensä	15 380 937		1,000

Huom. Tuloksen ”7 oikein” pieni tn ei johdu siitä, että kyseinen rivi olisi muita epätodennäköisempi – jokaisen yksittäisen rivin tn on sama 1/ 15380937. Olennaista on eri tapahtumille suotuisien rivien **lukumäärä**. (Kuvittele kaikkien rivien luettelo ja jaa rivit ryhmiin ”*k* oikein”.)

EHDOLLINEN TODENNÄKÖISYYS

Nostetaan 2 korttia, P(ässäpari)?

Tapa 1, variaatiot (vrt. Tuominen s. 37)

- Perusjoukko: 2 kortin jonot, $(52)_2 = 2652$ kpl.
- Suotuisat tapaukset: 2 ässän jonot, $(4)_2 = 12$ kpl.
- P(ässäpari) = 12 / 2652 = **1/221**

Tapa 2, kombinaatiot, käytiin 3. luennolla, tulos oli sama

Tapa 3: ehdollinen tn ja kertolaskusääntö (lause 1.7.4)
Merk. A = ”1. kortti on ässä”, B = ”2. kortti on ässä”.

- P(A) = 4/52
- **P(B | A)** = 3/51
- $P(A \cap B)$ = **P(A) · P(B | A)** = **1/221**

Vrt. osamäärä- ja kertolaskulausekkeita tavoissa 1 ja 3!

Riippuvuus

- Entä mikä on $P(B \mid A^c)$?
- Jos 1. kortti ei ollut ässä, jäljellä on 51 korttia joissa 4 ässää
- Ilmeisestikin $P(B \mid A^c) = 4/51$,
  eli enemmän kuin $P(B \mid A) = 3/51$

- **Jos tiedämme, onko tapahtuma A totta, todennäköisyytemme B:lle muuttuu**
- Tätä kutsutaan tn-laskennassa **riippuvuudeksi** tapahtumien välillä.

Riippuvat ja riippumattomat

Tapahtumat eli väitteet A ja B ovat todennäköisyyslaskennan mielessä

riippuvia, jos A:n totuusarvon tietäminen **auttaa** arvioimaan B:n totuusarvoa

Toisin sanoen: B:n mahdollisuudet (olla tosi tai epätosi)
jakautuvat **eri** tavalla A:n ollessa tosi ja A:n ollessa epätosi

Jos tapahtumat **eivät ole riippuvia**, ne ovat **riippumattomia**. Tällöin toisen tietäminen **ei auta** arvioimaan toisen totuusarvoa. Matemaattisesti tämä voidaan ilmaista yhtälönä:

Riippumattomille tapahtumille pätee

$$\mathbf{P(B \mid A) = P(B \mid A^c)}$$

Tämä ei päde yleisesti mille tahansa tapahtumille A ja B.

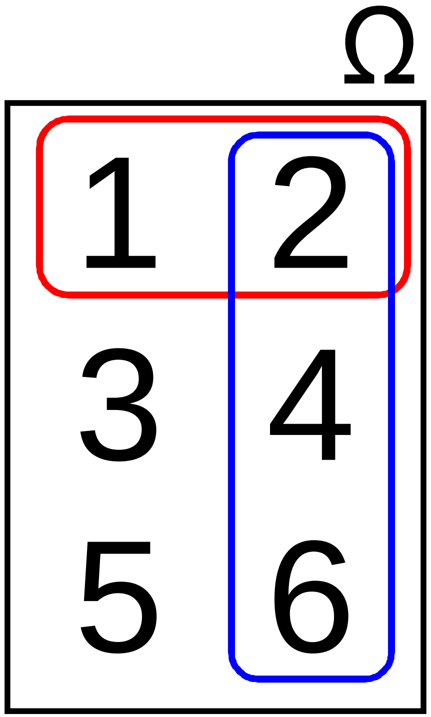
Tapahtuma vs. koe

- Tn-slangi: **tapahtuma** tarkoittaa **väitettä**, joka on tosi tai epätosi.
  - Joukko-opillisesti: **Tapahtuma on jokin joukko** alkeistapauksia.
  - Kokeessa toteutuu tasan yksi alkeistapauksista.
  - Tapahtuma $A = \{\omega_1, \omega_2, \omega_3\}$ vastaa väitettä ”toteutui jokin alkeistapauksista $\omega_1, \omega_2, \omega_3$”
- Esim. **heitetään yhtä noppaa** ja tulos on X. **Tämä ei ole tapahtuma** vaan **”koe”.**

- **Tapahtumia** ovat tässä yhdessä nopanheitossa väitteet
  - **A** **”X on pieni”** eli ”X on 1 tai 2” eli joukkona {1,2}
  - **B** **”X on parillinen”** eli ”X on 2 tai 4 tai 6” eli joukkona {2,4,6}
  - **C** **”X on tasan kaksi”** eli joukkona {2}
  - ... kaikki joukot, jotka alkeistapauksista 1,...,6 voi muodostaa **(montako?)**

- Yhdestä <u>kokeesta</u> voidaan siis luetella vaikka kuinka monta eri <u>tapahtumaa</u>, joista osa tapahtuu (on tosia) ko. kokeessa ja osa ei.

Esim. yksi noppa

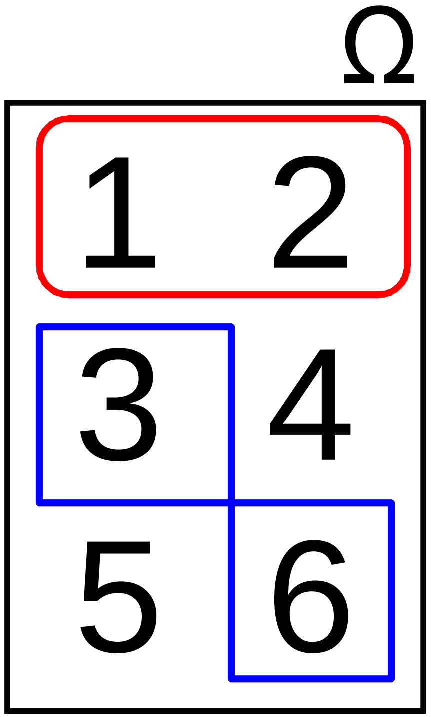


P(pieni) = 2/6 = 1/3
P(parillinen) = 3/6 = 1/2
P(pieni ja parillinen) = 1/6

Parillisia on **sama osuus** (1/2) pienistä kuin koko perusjoukosta.

Tapahtumat ”pieni” ja ”parillinen” ovat **riippumattomat**.

Esim. yksi noppa



P(pieni)	= 2/6	= 1/3
P(3:lla jaollinen)	= 2/6	= 1/3
P(pieni ja 3:lla jaollinen)	= 0/6	= 0

3:lla jaollisia on **eri osuus** (0) pienistä kuin koko perusjoukosta (1/3).

Tapahtumat ”pieni” ja ”3:lla jaollinen” ovat **riippuvat**.

Esim. kaksi noppaa

Ω

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

P(eka noppa on 4) = 6/36 = 1/6

P(toka noppa on 6) = 6/36 = 1/6

P(eka 4 ja toka 6) = 1/36

Tapahtumat ”eka 4” ja ”toka 6” ovat **riippumattomat**.

TOISTOKOE

Tn saada *k* onnistumista

- *n* koetta, jokaisessa sama onnistumis-tn
- Kokeiden tulokset riippumattomia (oletus!)
- Esimerkiksi:
  - Kolikonheitto, tuleeko kruuna
  - Nopanheitto, tuleeko kuutonen
  - Otanta takaisinpanolla, tuleeko punainen pallo
- Lause 1.9.1: onnistumisten lukumäärän tn saadaan binomikertoimella

  Miksi: ajatellaan alkeistapausten luetteloa!

  Ks. kurssisivulta ”Esimerkki toistokokeesta”