Johdatus todennäköisyyslaskentaan
kevät 2014

Luento 2 / 13

Jukka Kohonen
Matematiikan ja tilastotieteen laitos
Helsingin yliopisto

Klassinen eli symmetrinen tn

Oletus: On n vaihtoehtoa (**alkeistapausta**) joista **toteutuu tasan yksi**.

– nopanheitto	$\Omega = \{1, 2, 3, 4, 5, 6\}$	$n = 6$
– kolikonheitto	Ω = {kruuna, klaava}	$n = 2$
– kortti pakasta	Ω = {♥A, ♥2, ♥3, ... }	$n = 52$

Idea: Ajattelemme (jostain syystä), että jokainen **perusjoukon** Ω alkio eli alkeistapaus on ”yhtä todennäköinen”. Jokin niistä tapahtuu ”umpimähkään”.

Määritelmä:

Tapahtuma on mikä tahansa osajoukko $A \subset \Omega$, ja sen **todennäköisyys** on $P(A) = n(A) / n(\Omega)$.

”suotuisien alkeistapausten **lukumääräinen osuus**”

Huom. Kunkin alkeistapauksen tn on $1 / n(\Omega)$.

Symmetrisiä vai ei?

- Vaikka olisimme luetelleet ***n*** vaihtoehtoa, ja tietäisimme, että tasan yksi niistä toteutuu, ei siitä seuraa, että ne ovat välttämättä yhtä todennäköisiä.

Heitetään nasta kovalle alustalle. Tiedetään, että nasta tulee joko
(A) piikki alaspäin tai
(B) piikki ylöspäin.
Mahdollisuuksia on kaksi ($n = 2$). Varmasti toteutuu jompikumpi niistä. Ovatko todennäköisyydet = ½ ?

- Symmetria on oletus, jonka järkevyyttä on arvioitava erikseen. **Jos *n*** alkeistapausta ovat symmetrisiä, on kunkin todennäköisyys **1/*n***.

Tn-slangia

termi	selitys	esim
koe (satunnaiskoe)	Asetelma, jossa joidenkin tapahtumien todennäköisyyksiä tarkastellaan	nopanheitto
tapahtuma	Mikä tahansa asia (väite), jonka totuutta yritetään arvioida todennäköisyydellä	”Tulos on parillinen”
alkeistapaus	”Pienin” mahdollinen tapahtuma, josta olemme kiinnostuneet	”Tulos on 6” 6
suotuisa	Sellainen alkeistapaus, jossa tietty tapahtuma toteutuu	alkeistapaukset 2, 4 ja 6 ovat suotuisia
perusjoukko	Kaikkien mahdollisten alkeistapausten joukko; **tasan yksi niistä toteutuu**	{1, 2, 3, 4, 5, 6}
symmetriset	Yhtä todennäköiset	”Reilun” nopan alkeistapaukset
ei symmetriset	Ei yhtä todennäköiset	nastanheitto

Joukko-oppi ↔ logiikka

Esim. Nopanheitto $\Omega = \{1, 2, 3, 4, 5, 6\}$, määritellään tapahtumia

$A =$	$\{1, 2\}$	= ”(silmäluku on) pieni”	**P(*A*) = $^2/_6 = {}^1/_3$**
$B =$	$\{3, 4\}$	= ”keskisuuri”	**P(*B*) = $^2/_6 = {}^1/_3$**
$C =$	$\{5, 6\}$	= ”iso”	**P(*C*) = $^2/_6 = {}^1/_3$**
$D =$	$\{1, 3, 5\}$	= ”pariton”	**P(*D*) = $^3/_6 = {}^1/_2$**
$E =$	$\{2, 4, 6\}$	= ”parillinen”	**P(*E*) = $^3/_6 = {}^1/_2$**

Unioni (∪) vastaa sanaa ”tai” (∨) (symbolien ∪ ↔ ∨ samannäköisyys ei ole sattumaa!)

$A \cup B = \{1,2\} \cup \{3,4\} = \{1,2,3,4\}$ = ”(silmäluku on) pieni **tai** keskisuuri”

Leikkaus (∩) vastaa sanaa ”ja” (∧)

$A \cap E = \{1,2\} \cap \{2,4,6\} = \{2\}$ = ”(silmäluku on) pieni **ja** parillinen”

Komplementti vastaa sanaa ”ei”:

$E^C = \{2, 4, 6\}^C = \{1, 3, 5\}$ = ”ei parillinen” = ”pariton”

Eri syntakseja

Joukko-oppi	Logiikka	MATLAB-operaattori
$A \cup B$ ”unioni”	$A \lor B$ ”tai”	A **|** B
$A \cap B$ ”leikkaus”	$A \land B$ ”ja”	A **&** B
A^c ”komplementti”	$\neg A$ ”ei”	**not**(A)

Ilmeisiä perushavaintoja

Additiivisuus: Jos $A \cap B = \emptyset$, eli tapahtumat erilliset, niin

$$P(A \cup B) = P(A) + P(B).$$

Todistus: Unionissa on **yhteensä n(A) + n(B) alkiota**, ja niiden lukumääräinen osuus perusjoukosta on siis $(n(A)+n(B)) / n(\Omega)$, josta väite seuraa.

Mutta jos tapahtumat eivät ole erilliset, niin niillä on yhteisiä alkioita: unionin alkioiden lukumäärälle tuleekin kaava $n(A) + n(B) - n(A \cap B)$.

Lisää ilmeisiä perushavaintoja

Todennäköisyyden ala- ja yläraja:

$0 \le P(A) \le 1$ (kaikille tapahtumille A)

Komplementin todennäköisyys:

$P(A^C) = 1 - P(A)$ Esim: P(*pariton*) = 1 – P(*parillinen*).

Nämä (yms.) seuraavat *symmetriselle* todennäköisyydelle suoraan **joukko-opin perusominaisuuksista** (kertaa ne!) ja alkioiden lukumäärien laskemisesta.

Mutta monet samat kaavat pätevät yleisemminkin, ei-symmetriselle todennäköisyydelle. Tätä käsitellään pian tarkemmin **todennäköisyyden aksioomien** jälkeen.

Yhdistetty satunnaiskoe

- Suoritetaan peräkkäin esim. kaksi (erilaista tai samanlaista) satunnaiskoetta, esim. kaksi nopanheittoa.
- Koko juttu voidaan ajatella yhdeksi kokeeksi, jonka tuloksena on **jono** yksittäisten kokeiden tuloksisa.
- Yhdistetyn kokeen perusjoukko on siis yksittäisten kokeiden perusjoukkojen **karteesinen tulo**.

KOMBINATORIIKKA

SUURTEN LUKUMÄÄRIEN LASKEMISEN TAITO

Kombinatoriikkaa

- Symmetrisessä mallissa tapahtuman tn on helppo laskea, jos osataan laskea suotuisat alkeistapaukset.
- Joskus, varsinkin yhdistetyssä satunnaiskokeessa, alkeistapauksia on **hyvin monta** (tuhansia, miljardeja, enemmän).
- Kombinatoriikassa lasketaan näitä suuria lukumääriä.

Kombinatoriikkaa

Kymmenen henkilöä, numeroituna 1,…,10, asettuu jonoon umpimähkään. Millä tn he ovat aakkosjärjestyksessä?

Alkeistapaus on tietty järjestys, esim.

(2, 5, 1, 10, 3, 4, 7, 9, 5, 8).

Eri alkeistapauksia on hyvin monta. **Luetteleminen** ei kannata.

Osaamme kuitenkin **laskea**, että niitä on 10! = 3 628 800 kappaletta.

Isot symmetriset tn-avaruudet

Pääkysymykset ovat:

1. Mitkä ovat **alkeistapaukset**? Montako niitä on?
   - voiko ne esim. **luetella** kaikki, tai
   - voiko **kuvitella** niiden luettelemisen jonkinlaisen järjestyksen tai valintaprosessin kautta
2. Mitkä ovat tapahtuman **suotuisat alkeistapaukset**? Millaisia ne ovat? Montako niitä on?

Kun näihin on vastattu, niin tapahtuman *A* tn on

$$\frac{n(A)}{n(\Omega)} \quad \begin{array}{l} \leftarrow \textbf{suotuisien alkeistapausten määrä} \\ \leftarrow \textbf{kaikkien alkeistapausten määrä} \end{array}$$

Helpommin sanottu kuin tehty

Perusjoukko voi olla melko suuri tai **hyvin suuri**

- Kahden erilaisen kolikon heittotuloksia 4 kpl

  {(0,0), (0,1), (1,0), (1,1)}
- Kahden erilaisen nopan heittotuloksia 36 kpl

  {(1,1), (1,2), ..., (1,6), (2,1), (2,2), ..., (6,6)}
- Kolmen erilaisen nopan heittotuloksia 216 kpl

  ei jaksa luetella
- Korttipakan järjestyksiä $52! \approx 8 \cdot 10^{67}$

  enemmän kuin maapallossa on atomeita

Tuloperiaate

- Tarkasteltavat oliot (alkeistapaukset) voidaan muodostaa tekemällä ***k*** **peräkkäistä valintaa**
- *i*:nnessä vaiheessa valitaan $\boldsymbol{n_i}$:stä vaihtoehdosta siten, että **lukumäärä ei riipu siitä aiemmin valittiin**

→ Erilaisia olioita on

$\boldsymbol{n_1 \cdot n_2 \cdot \ldots \cdot n_k}$ kappaletta.

Kolmen nopan heitto

- Alkeistapauksina järjestetyt jonot (a,b,c), missä a, b ja c ovat kokonaislukuja joukosta {1,2,3,4,5,6}
  - 1. luku voidaan valita 6 tavalla $n_1 = \mathbf{6}$
  - 2. luku voidaan valita 6 tavalla $n_2 = \mathbf{6}$
  - 3. luku voidaan valita 6 tavalla $n_3 = \mathbf{6}$
- Alkeistapauksia on $6 \cdot 6 \cdot 6 = 216$ kpl

- Tässä oli $n_1 = n_2 = n_3$

  eli joka vaiheessa oli **sama määrä vaihtoehtoja.**
  Kertolaskun voisi laskea suoraan potenssina 6^3.

Kolmen ihmisen jono (permutaatio, s. 23)

- A, B, C asettuvat jonoon. Kukin ihminen on jonossa vain yhdessä kohtaa! Jono (A,A,A) eli AAA on mahdoton.
- Tuloperiaate, mutta **vaihtoehdot vähenevät** vaihe vaiheelta (joka kerta 1:llä)
- 1. paikassa voi olla A, B tai C $n_1 = \mathbf{3}$
- 2. paikassa jompikumpi jäljelle jääneistä $n_2 = \mathbf{2}$
- 3. paikassa on viimeiseksi jäänyt $n_3 = \mathbf{1}$
- Jonoja on 3·2·1 = 6 erilaista. Luetellaan ne:

{ABC, ACB,
BAC, BCA,
CAB, CBA}

Osajono (variaatio, Tuominen s. 26)

Kymmenestä ihmisestä ABCDEFGHIJ kolme tulee huoneeseen jonossa.

Montako erilaista jonoa on?

- 1. ihminen on joku 10:stä
- 2. ihminen on joku 9:stä (ei sama kuin 1. ihminen)
- 3. ihminen on joku 8:sta (ei kumpikaan aiemmista kahdesta)

Laskeva tulo $10 \cdot 9 \cdot 8 = \mathbf{(10)_3} = 720$

= alkaa 10:stä, kolme tekijää, pienenee aina 1:llä

Huom 1. Kussakin vaiheessa valintavaihtoehdot riippuvat siitä, mitä aiemmin valittiin, mutta vaihtoehtojen määrä ei riipu

Huom 2. Jos osajonoon tulee kaikki 10 ihmistä, saadaan permutaatio

Vertaa alkeistapauksia

	3 noppaa	**3 ihmisen jono (permutaatio)**	**3 ihmisen osajono 10:stä (variaatio)**
Vaihtoehtoja vaiheittain	$6 \cdot 6 \cdot 6$	$3 \cdot 2 \cdot 1$	$10 \cdot 9 \cdot 8$
Lyhyempi merkintä	6^3 **potenssi**	$(3)_3 = 3!$ **kertoma**	$(10)_3$ **laskeva tulo**
Yhteensä	216 kpl	6 kpl	720 kpl
Alkeistapausten luettelo = perusjoukko	{111, 112, 113, ... 121, 122, 123, ... 211, 212, ... 611, ..., 666}	{ABC, ACB, BAC, BCA, CAB, CBA}	{ABC, ABD, ABE, ..., ABJ, ACB, ACD, ..., ACJ, ... BAC, BAD, ... DAB, ..., EAB, …, JIH}

Entäs ne suotuisat

- **Perusjoukon** koon laskeminen on usein helppoa.
  - Usein joku kombinatoriikan perusmenetelmä (permutaatiot tms.)

- **Suotuisien** alkeistapausten laskeminen voi olla ”hiukan” vaikeampaa.
  - Pitää olla tarkkana, että kaikki mahdollisuudet tulee laskettua (ja jokainen vain kerran).
  - Usein ei onnistu suoraan yhdellä lausekkeella (kertoma, binomikerroin tms) vaan joudutaan käyttämään tuloperiaatetta ja (joskus hyvinkin yksityiskohtaista) päättelyä ja vaihtoehtojen läpikäyntiä.

Kaksi noppaa, P(molemmat parillisia)?

- **Alkeistapaukset** = 2-jonot joukosta {1,...,6}, niitä on 36 kpl: {(1,1),(1,2),...,(6,5),(6,6)}
- **Suotuisat** alkeistapaukset: Eka noppa on joku joukosta {2,4,6}, samoin toinen, suotuisia ovat siis {(2,2),(2,4),(2,6),...,(6,6)}
  - **Tuloperiaate**, suotuisia on 3·3 = 9 kpl

- Todennäköisyys = **9 / 36** = 1/4

Kaksi noppaa, P(summa = 6)?

- Samat 36 alkeistapausta kuin edellä
- **Suotuisia** ovat {(1,5),(2,4),(3,3),(4,2),(5,1)} eli **5** kpl
- Todennäköisyys = 5/36

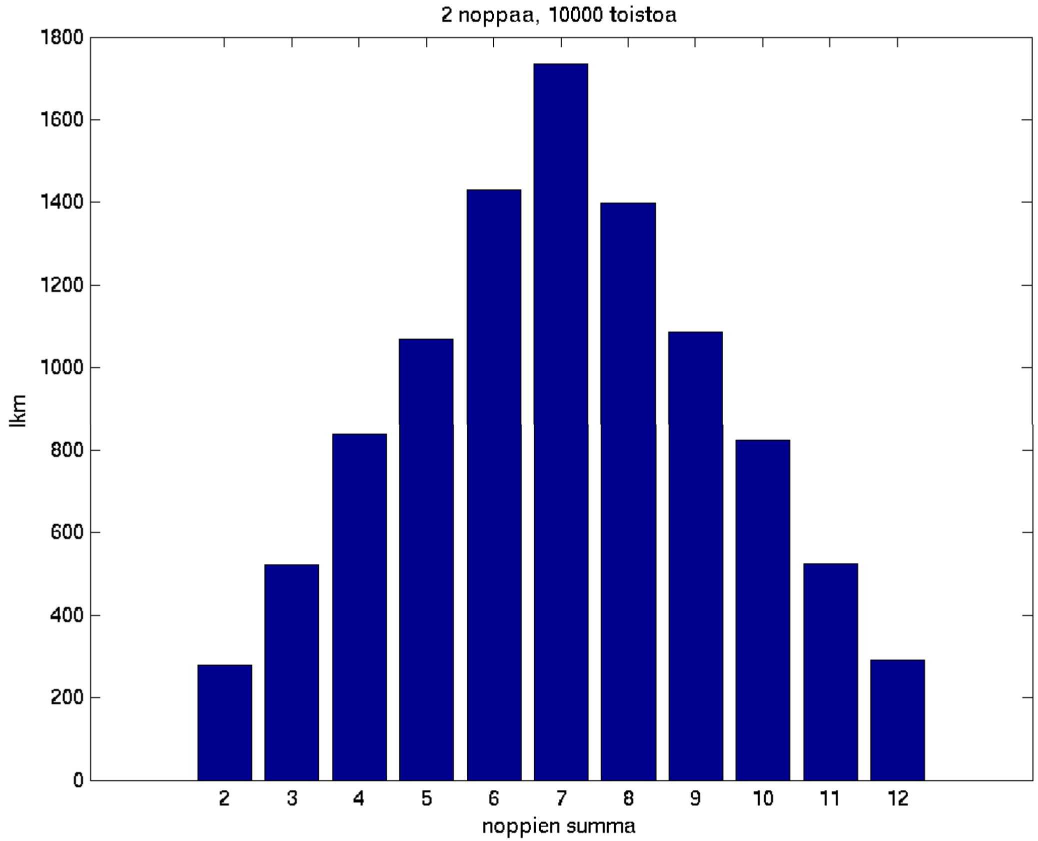
- Huom. Noppien summa on jokin 11 luvusta {2,3,4,...,12}, mutta nämä mahdollisuudet eivät ole keskenään yhtä todennäköiset!

Kaksi noppaa, eri summien tn:t

Noppien summa	**Suotuisat alkeistapaukset**	**Suotuisia kpl**	**tn**
2	”11”	1	1 / 36
3	”12”, ”21”	2	2 / 36
4	”13”, ”22”, ”31”	3	3 / 36
5	”14”, ”23”, ”32”, ”41”	4	4 / 36
6	”15”, ”24”, ”33”, ”42”, ”51”	5	5 / 36
7	”16”, ”25”, ”34”, ”43”, ”52”, ”61”	6	6 / 36
8	”26”, ”35”, ”44”, ”53”, ”62”	5	5 / 36
9	”36”, ”45”, ”54”, ”63”	4	4 / 36
10	”46”, ”55”, ”64”	3	3 / 36
11	”56”, ”65”	2	2 / 36
12	”66”	1	1 / 36
yhteensä		36	1

Kaksi noppaa, summan kokeellinen jakauma

```
n = 10000;
a = noppa(n);
b = noppa(n);
s = a+b;
pylvas(s)
```



10 ihmisen jono, onko A **ennen** B:tä?

- Jonoja on 10! = 3.6 miljoonaa kpl, ikävä läpikäydä yksitellen
- Millaisia ovat suotuisat jonot?
- Jonossa 10 paikkaa.
  - A on jossain 10 paikasta
  - B on jossain, jäljellä 9 paikkaa, **mutta niistä vain osa on suotuisia**
  - Muut 8 ovat jossain (jäljellä 8, 7, ..., 1 paikkaa)
- Jatkuu ...

Onko A ennen B:tä (jatkuu)

- **Jos** A on 1. paikalla, B saa olla missä vain (9 vaihtoehtoa)
- **Jos** A on 2. paikalla, B saa olla paikalla 3...10 (8 vaihtoehtoa)
- ...
- **Jos** A on 9. paikalla, B täytyy olla 10. paikalla (1 vaihtoehto)
- **Jos** A on 10. paikalla, B ei voi olla missään (0 vaihtoehtoa)

Summaperiaate: A:n ja B:n paikat voidaan valita 9+8+7+6+5+4+3+2+1+0 = 45 tavalla.

Sen jälkeen loppujen 8 ihmisen paikat mielivaltaisesti 8! = 40320 tavalla.

Suotuisia jonoja (tuloperiaate) on 45 · 40320 = 1814400kpl

Todennäköisyys = 1814400 / 10! = **0.5**

Olisiko saman voinut päätellä lyhyemmin jonkin symmetrian nojalla?

Jos korttipakka sekoitetaan, millä tn pataässä on ennen ruutuviitosta?

Osajoukot (kombinaatiot s. 26)

4 ihmistä ABCD kättelevät toisiaan. Montako kättelyä tapahtuu? (2 ihmisen osajoukot)

Jos luetellaan 2-osajonot, $4 \cdot 3 = 12$ kpl

AB, AC, AD,

BC, BA, BD,

CA, CB, CD,

DA, DB, DC

Osajoukot (kombinaatiot s. 26)

4 ihmistä ABCD kättelevät toisiaan. Montako kättelyä tapahtuu? (2 ihmisen osajoukot)

Jos luetellaan 2-osajonot, $4 \cdot 3 = 12$ kpl

Punaisella merkityt ovat sama osajoukko
$\{A,B\} = \{B,A\}$

AB, AC, AD,

BC, BA, BD,

CA, CB, CD,

DA, DB, DC

Osajoukot (kombinaatiot s. 26)

4 ihmistä ABCD kättelevät toisiaan. Montako kättelyä tapahtuu? (2 ihmisen osajoukot)

Jos luetellaan 2-osajonot, $4 \cdot 3 = 12$ kpl

AB, AC, AD,

BC, BA, BD,

CA, CB, CD,

DA, DB, DC

Punaisella merkityt ovat sama osajoukko
{A,B} = {B,A}

Sinisellä merkityt ovat sama osajoukko
{A,C} = {C,A}

Jne: Jokaista 2-osajoukkoa vastaa kaksi eri 2-osajonoa. Osajoukkojen määrä on vain puolet eli 6 kpl

Osajoukot (kombinaatiot s. 26)

4 ihmistä ABCD kättelevät toisiaan. Montako kättelyä tapahtuu? (2 ihmisen osajoukot)

AB, AC, AD,

BC, BD,

CD

Karsittu luettelosta tuplat pois,
nyt kukin osajoukko on kerran

Osajoukkojen määrä on 6 kpl

Osajoukot yleisesti

n:stä alkiosta valitaan *k*-osajoukkoja (k-kombinaatioita):

valinta voidaan tehdä näin:

- Muodostetaan ensin kaikki *k*-osajonot: $(n)_k$ kpl
- Kutakin osajoukkoa vastaa *k*! eri osajonoa, ja valitaan niistä vain yksi

Osajoukkoja on siis näin monta (miksi?)

$$\frac{(n)_k}{k!} = \frac{n(n-1)...(n-k+1)}{k!} = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

Vertailutaulukko (täydennetty)

	3 noppaa	**3 ihmisen jono (permutaatio)**	**3 ihmisen osajono 10:stä (variaatio)**	**3 ihmisen osajoukko 10:stä (kombinaatio)**
Vaihtoehtoja vaiheittain	$6 \cdot 6 \cdot 6$	$3 \cdot 2 \cdot 1$	$10 \cdot 9 \cdot 8$	$\dfrac{10 \cdot 9 \cdot 8}{3 \cdot 2 \cdot 1}$
Lyhyempi merkintä	6^3 **potenssi**	$(3)_3 = 3!$ **kertoma**	$(10)_3$ **laskeva tulo**	$\binom{10}{3}$ **binomikerroin**
Yhteensä	216 kpl	6 kpl	720 kpl	120 kpl
Alkeis-tapausten luettelo = perusjoukko	{111, 112, 113, ... 121, 122, 123, ... 211, 212, ... 611, ..., 666}	{ABC, ACB, BAC, BCA, CAB, CBA}	{ABC, ABD, ABE, ..., ABJ, ACB, ACD, ..., ACJ, ... BAC, BAD, ... DAB, ..., EAB, …, HIJ, HJI, …, JIH}	{{A,B,C}, {A,B,D}, ..., {H,I,J}}

Montako lottoriviä on olemassa?

- 39:stä numerosta valitaan 7-osajoukko

$$\binom{39}{7} = \frac{(39)_7}{7!} = \frac{39 \cdot 38 \cdot 37 \cdot 36 \cdot 35 \cdot 34 \cdot 33}{7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 15\,380\,937$$

Montako lottoriviä, joissa on kolmonen?

- Kolmonen on pakko ottaa mukaan osajoukkoon, siinä ei ole valinnanvaraa
- Lopuista 38:sta numerosta valitaan 6-osajoukko:

$$\binom{38}{6} = \frac{(38)_6}{6!} = \frac{38 \cdot 37 \cdot 36 \cdot 35 \cdot 34 \cdot 33}{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 2\,760\,681$$

Montako lottoriviä, joissa on 3 ja 8?

- 3 ja 8 on pakko ottaa mukaan osajoukkoon, siinä ei ole valinnanvaraa
- Lopuista 37:sta numerosta valitaan 5-osajoukko:

$$\binom{37}{5} = \frac{(37)_5}{5!} = \frac{37 \cdot 36 \cdot 35 \cdot 34 \cdot 33}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 435\,897$$

Tn saada lotossa 4 oikein?

- Asiakas on valinnut jotkut 7 numeroa, esim. {2, 3, 10, 11, 26, 28, 33}
- Kone arpoo jotkut 7 numeroa
- Tn, että riveissä on tasan 4 samaa?
- Perusjoukossa on (39 yli 7) riviä, kone arpoo **yhden rivin**
- Suotuisia: Koneen pitää arpoa
  - Jotkut 4 numeroa asiakkaan 7:stä, (7 yli 4)
  - Jotkut 3 numeroa niistä 39-7 = 32:sta, joita asiakas ei valinnut (muuten tulee enemmän kuin 4 oikein)
  - Tuloperiaate: (7 yli 4) · (32 yli 3) suotuisaa riviä
- Tn = (7 yli 4) · (32 yli 3) / (39 yli 7)

Pokerikäsi, tn saada 4 samaa

- 52 korttia (4 maata à 13 korttia)
- 5 kortin käsi = osajoukko: (52 yli 5) erilaista
- Suotuisien käsien lukumäärän laskeminen:
  - Kädessä oltava jotakin numeroarvoa 4 kpl – tämä numeroarvo voidaan valita 13 tavalla
  - Käden viides kortti on joku jäljellä olevista 48 kortista
  - Tuloperiaate, suotuisia käsiä $13 \cdot 48 = 624$ kpl (oliko siinä varmasti kaikki?)
- Tn on 624 / (52 yli 5) ≈ 0.000240

Montako osajoukkoa (s. 31)

- Montako *mielivaltaisen kokoista* osajoukkoa on n alkion joukolla?
- Voisi luetella erikseen erikokoiset osajoukot
- Tai sitten alkioittain: 1. alkio joko on tai ei ole osajoukossa, 2. alkio samoin jne.
- Näin syntyy kaikki eri osajoukot **(onko näin?)**
- Tuloperiaate: n vaihetta, jokaisessa 2 vaihtoehtoa

  $\rightarrow$ 2^n eri osajoukkoa

Joukon {a,b,c} osajoukot

• { }	0-osajoukot	1 kpl
• {a},{b},{c}	1-osajoukot	3 kpl
• {a,b}, {a,c}, {b,c}	2-osajoukot	3 kpl
• {a,b,c}	3-osajoukot	1 kpl
Yhteensä	$2^3 =$	8 kpl

Tapausten luetteleminen (matlab)

Jonot (järjestykset)

>> **perms**('ABCD')

```
ans =
DCBA
DCAB
DBCA
DBAC
DABC
DACB
CDBA
CDAB
CBDA
CBAD
CABD
CADB
BCDA
BCAD
BDCA
BDAC
BADC
BACD
ACBD
ACDB
ABCD
ABDC
ADBC
ADCB
```

k-osajoukot

>> **nchoosek**('ABCD',2)

```
ans =
AB
AC
AD
BC
BD
CD
```

Lisää esimerkkejä

- Tuomisen kirja s. 23-
- Mika Koskenojan esimerkkikokoelma

  http://www.helsinki.fi/~koskenoj/johdtn/esimerkit.pdf