Todennäköisyyslaskenta

Osa 1: Todennäköisyys ja sen laskusäännöt

- Satunnaiskokeet, otosavaruudet ja tapahtumat

- Todennäköisyyden määritteleminen

Satunnaiskokeet

Reaalimaailman ilmiö on **satunnaisilmiö**, jos:

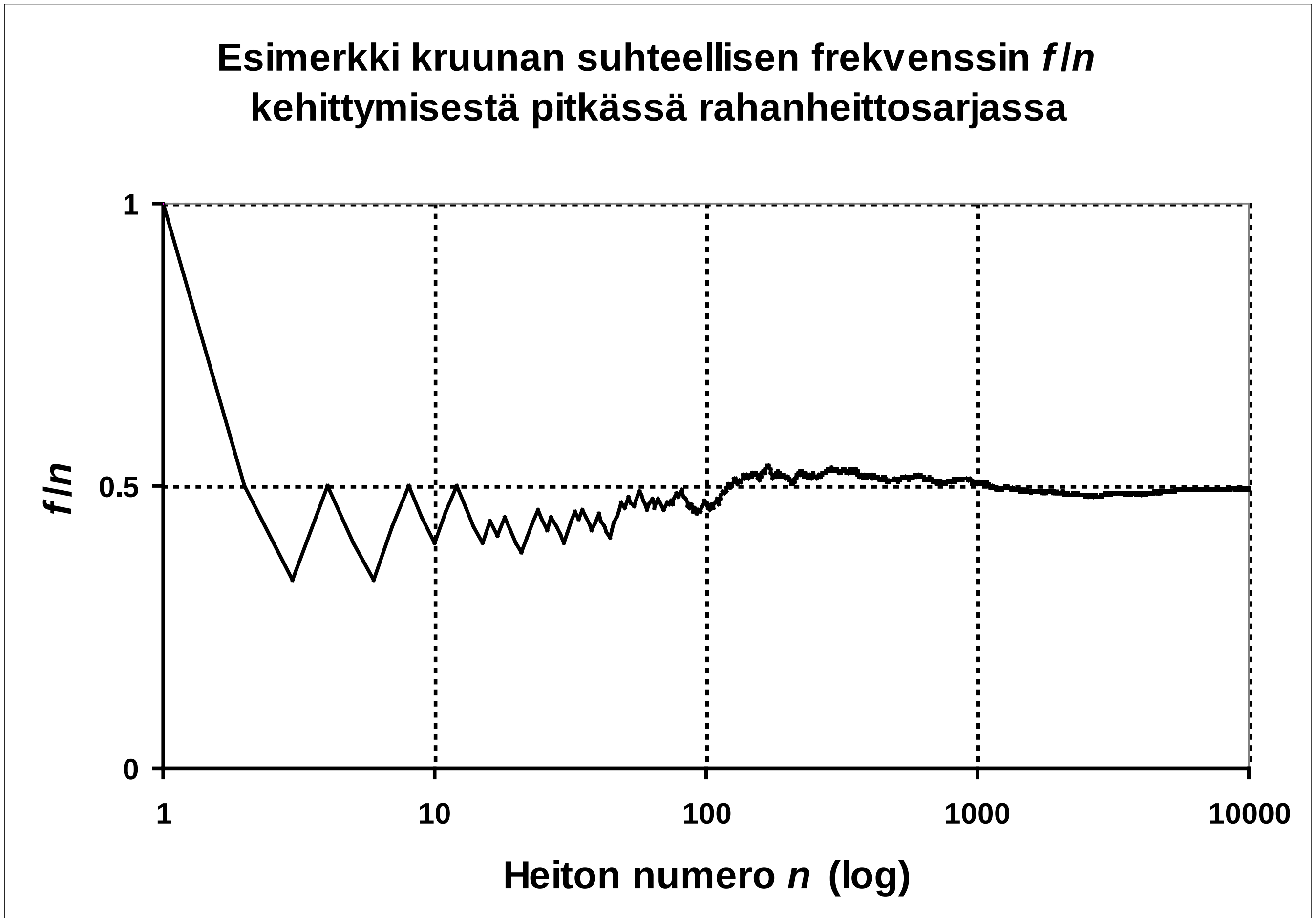
(i) Ilmiöllä on useita erilaisia *vaihtoehtoisia tuloksia*.

(ii) Ilmiön alkutilan perusteella *ei voida* tarkasti *ennustaa* ilmiön lopputilaa eli sitä, mikä mahdollisista *tulosvaihtoehdoista realisoituu* eli *toteutuu*.

(iii) Vaikka ilmiön lopputilaa ei voida ennustaa tarkasti, tulosvaihtoehtojen *suhteellisten frekvenssien* eli *osuuksien nähdään* ilmiön toistuessa *käyttäytyvän säännönmukaisesti*.

Kutsumme *satunnaisilmiötä* usein **satunnaiskokeeksi** ja *satunnaisilmiön esiintymiskertaa* **koetoistoksi**.

Tyypillinen esimerkki



Todennäköisyyslaskennan peruskäsitteet

(i) Sanomme, että satunnaisilmiön tulosvaihtoehto on **alkeistapahtuma**, jos satunnaisilmiötä *ei voida tai haluta* ”*purkaa*” sitä alkeellisempiin tulosvaihtoehtoihin

(ii) Kutsumme satunnaisilmiön *kaikkien* alkeistapahtumien muodostamaa joukkoa **otosavaruudeksi**.

Esimerkki - otosavaruus ja alkeistapahtumat:

- Satunnaiskoe (1):

  Nopan heitto

- Otosavaruus:

  $S = \{1,2,3,4,5,6\}$

- Alkeistapahtumat:

  $s_1 = 1$

  $s_2 = 2$

  $s_3 = 3$

  $s_4 = 4$

  $s_5 = 5$

  $s_6 = 6$

Esimerkki - otosavaruus ja alkeistapahtumat:

- Satunnaiskoe (2):

  Päivittäisen sademäärän mittaaminen

- Otosavaruus:

$$S = \{x : 0 \leq x \leq M\}$$

- Satunnaiskoe (3):

  Koneen osan tutkiminen onko viallinen

- Otosavaruus:

  S = {V,K}, V= viallinen K=kelvollinen

- Alkeistapahtumat:

  s_1 = V

  s_2 = K

Todennäköisyyslaskennan peruskäsitteet

(iii) Satunnaisilmiön **tapahtumat** ovat satunnaisilmiön alkeistapahtumien muodostamia otosavaruuden osajoukkoja.

(Muistutus joukko-opin peruskäsitteistä löytyy esimerkkikokoelma 1:n alussa kohdassa **Joukko-oppi**)

Tapahtumat

- Jos siis A on jokin otosavaruuden S *tapahtuma*, niin

  $$A \subset S$$

  eli

  $$s \in A \Rightarrow s \in S$$

  jossa s on tapahtumaan A kuuluva *alkeistapahtuma*.

- **Kun sanomme, että tapahtuma *A* sattuu, tarkoitamme aina sitä, että jokin tapahtumaan *A* kuuluva alkeis-tapahtuma *s* sattuu**.

Otosavaruus, alkeistapahtumat ja tapahtumat: Esimerkki nopanheitosta

- Satunnaiskoe:

  Nopanheiton tulos
- Otosavaruus:

  Silmälukujen joukko $S = \{1, 2, 3, 4, 5, 6\}$
- Alkeistapahtumat:

  Silmäluvut 1, 2, 3, 4, 5, 6
- Esimerkki tapahtumasta:

  A = ”Silmäluku on parillinen” = $\{2, 4, 6\}$

Otosavaruus, alkeistapahtumat ja tapahtumat: Esimerkki 2 nopanheitosta 1/3

- Satunnaisilmiö:

  Tulokset kahdesta nopanheitosta

- Otosavaruus S:

  Silmälukuparien (i, j) (36 kpl) joukko, jossa

  i = 1. nopanheiton silmäluku, $i = 1, 2, 3, 4, 5, 6$

  j = 2. nopanheiton silmäluku, $j = 1, 2, 3, 4, 5, 6$

$$
\begin{aligned}
S = \{ & (1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ (1,6), \\
& (2,1),\ (2,2),\ (2,3),\ (2,4),\ (2,5),\ (2,6), \\
& (3,1),\ (3,2),\ (3,3),\ (3,4),\ (3,5),\ (3,6), \\
& (4,1),\ (4,2),\ (4,3),\ (4,4),\ (4,5),\ (4,6), \\
& (5,1),\ (5,2),\ (5,3),\ (5,4),\ (5,5),\ (5,6), \\
& (6,1),\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6) \}
\end{aligned}
$$

Otosavaruus, alkeistapahtumat ja tapahtumat: Esimerkki 2 nopanheitosta 2/3

- Otosavaruuden alkiot voidaan esittää seuraavana *taulukkona*:

(*i*, *j*)	***j* = tulos 2. nopanheitosta**					
***i* = tulos 1. nopan-heitosta**	**1**	**2**	**3**	**4**	**5**	**6**
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Otosavaruus, alkeistapahtumat ja tapahtumat: Esimerkki 2 nopanheitosta 3/3

- Esimerkki tapahtumasta:

  A = ”Kummallakin nopalla saadaan sama silmäluku”
  = $\{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$

- Esimerkiksi:

  $$(2,2) \in A$$

  $$(6,1) \notin A$$

  $$A \subset S$$

Uusien tapahtumien muodostaminen

- Mistä tahansa otosavaruuden tapahtumista voidaan muodostaa uusia tapahtumia.
- Nämä uudet tapahtumat voidaan ilmaista sanoin tai kirjoittaa joukko-opin operaatioitten avulla.
- Käytämme **Venn-diagrammia** kuvaamaan otosavaruuksia ja niihin liittyviä tapahtumia.

Johdetut tapahtumat

- Olkoot A ja B johonkin satunnaiskokeeseen liittyviä *tapahtumia*.
- Tarkastelemme seuraavia **tapahtumista *A* ja *B* johdettuja tapahtumia**:

(i) Tapahtuma A *ei* satu.

(ii) Tapahtuma A *tai* B sattuu:

Tapahtuma A sattuu *tai* tapahtuma B sattuu *tai* molemmat sattuvat.

(iii) Tapahtuma A sattuu *ja* tapahtuma B sattuu.

(iv) Tapahtuma A sattuu, *mutta* tapahtuma B *ei* satu.

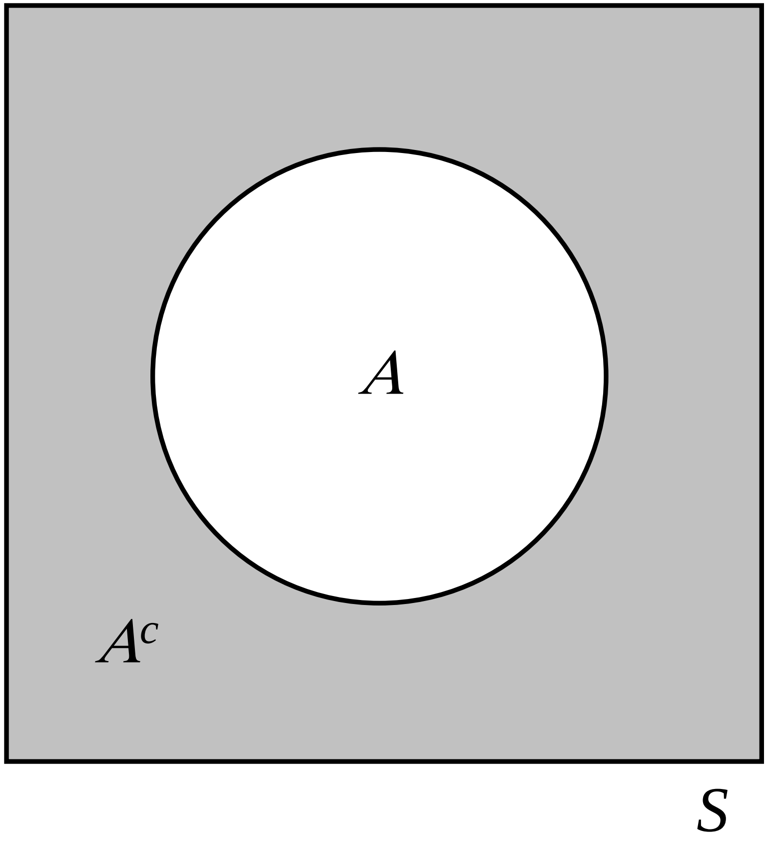
Tapahtuman *A* komplementti

- Olkoon $A \subset S$ otosavaruuden S tapahtuma.
- Tapahtuman A *komplementtitapahtuma*

  A^c = ”*A ei satu*” = *ei*-A

  on niiden alkeistapahtumien joukko, jotka *eivät* kuulu joukkoon A:

  $A^c = \{s \in S \mid s \notin A\}$



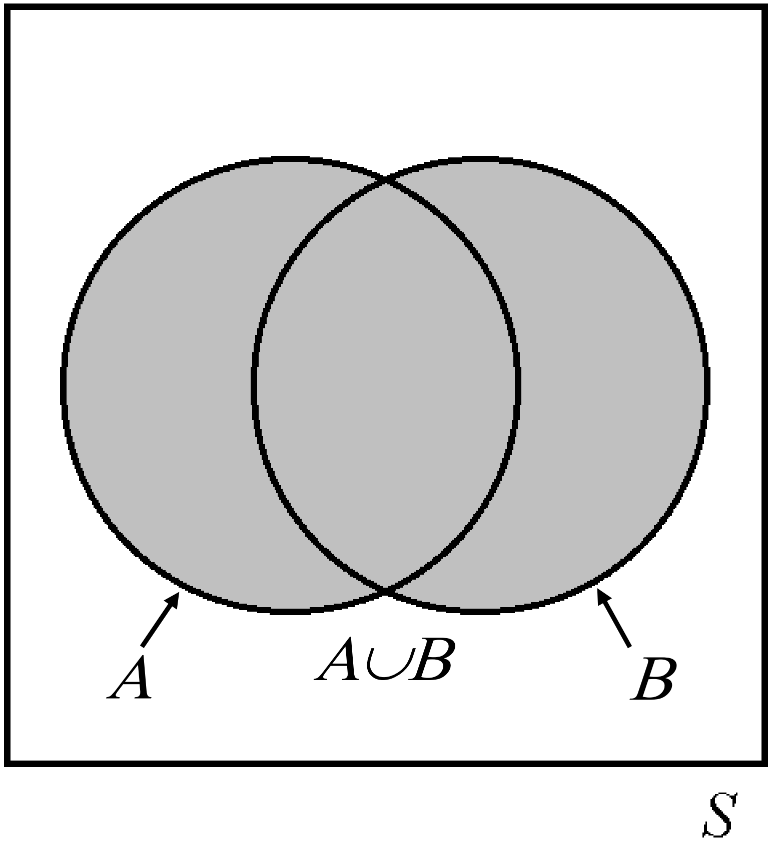
Tapahtumien *A* ja *B* yhdiste

- Olkoot $A \subset S$ ja $B \subset S$ otos-avaruuden S tapahtumia.
- Yhdistetty tapahtuma

$A \cup B$ = ”*A sattuu tai B sattuu tai molemmat sattuvat*”

on niiden alkeistapahtumien joukko, jotka kuuluvat joukkoon A *tai* joukkoon B *tai* molempiin:

$$A \cup B = \{s \in S \mid s \in A \text{ tai } s \in B\}$$



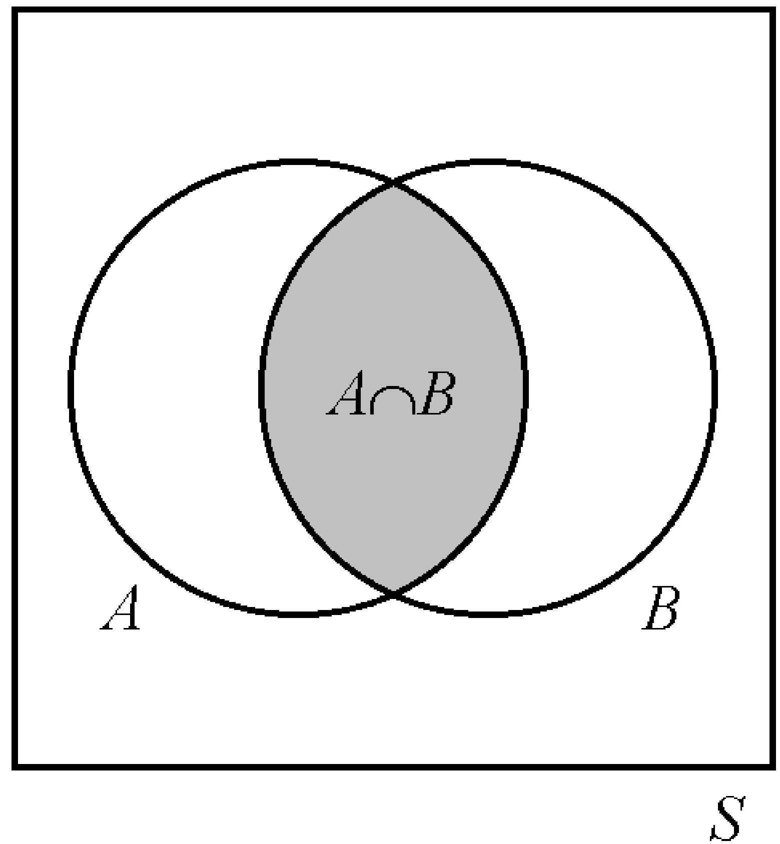
Tapahtumien *A* ja *B* leikkaus

- Olkoot $A \subset S$ ja $B \subset S$ otosavaruuden S tapahtumia.
- Yhdistetty tapahtuma

  $A \cap B$ = ”*A ja B sattuvat*”

  on niiden alkeistapahtumien joukko, jotka kuuluvat joukkoon A *ja* joukkoon B:

  $A \cap B$
  $= \{s \in S \mid s \in A \text{ ja } s \in B\}$



Esimerkki

- Satunnaiskoe: Nopanheiton tulos
- Otosavaruus: $S = \{1, 2, 3, 4, 5, 6\}$
- Tapahtumia:

A = ”Silmäluku on suurempi kuin 3” = $\{4,5,6\}$

B = ”Silmäluku on pienempi kuin 6” = $\{1,2,3,4,5\}$

C = ”Silmäluku on parillinen” = $\{2, 4, 6\}$

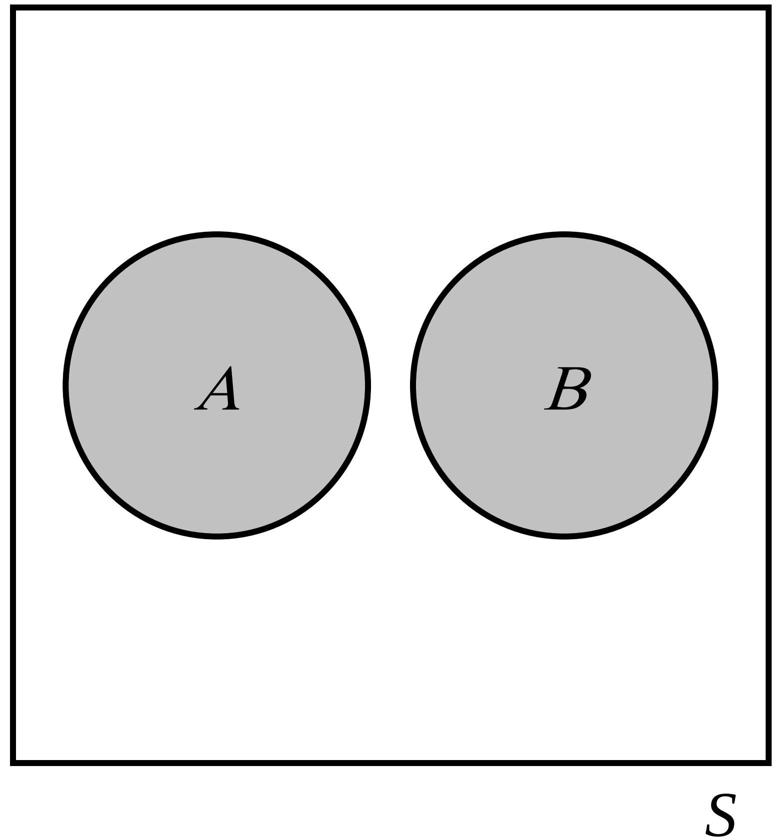
$A \cap C$ = ”Silmäluku on suurempi kuin 3 ja parillinen” = $\{4,6\}$

A^c = ”Silmäluku ei ole suurempi kuin 3” = $\{1,2,3\}$

Toisensa poissulkevat tapahtumat

- Tapahtumat A ja B ovat **toisensa poissulkevia**, jos A ja B *eivät voi sattua samanaikaisesti*.
- Tapahtumat A ja B ovat toisensa poissulkevia, jos ne ovat otosavaruuden S osajoukkoina *pistevieraita* eli

  $A \cap B = \varnothing$



Toisensa poissulkevat tapahtumat: Esimerkki eduskunnasta

- Satunnaiskoe: Kansanedustaja valitaan satunnaisesti kaikkien kansanedustajien joukosta.

- Olkoon

  A = ”Edustaja kuuluu vasemmistoliittoon”

  B = ”Edustaja kuuluu kokoomukseen”

- Tällöin tapahtumat A ja B ovat toisensa poissulkevia, koska yksikään kansanedustaja ei voi olla kahden puolueen jäsen, mikä merkitsee sitä, että joukot A ja B ovat pistevieraita:

  $A \cap B = \varnothing$

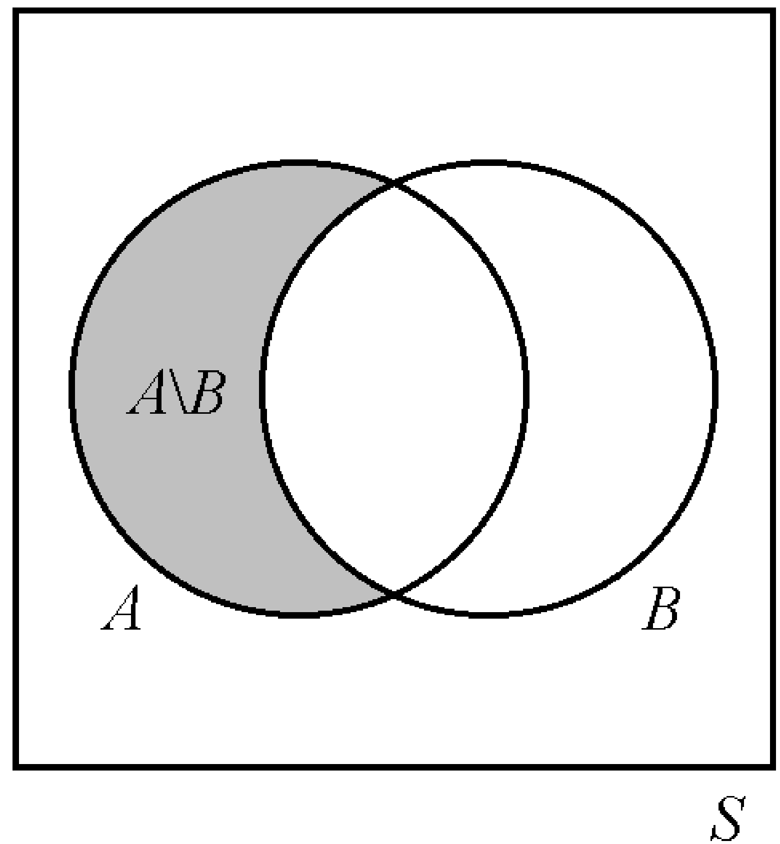
Tapahtumien *A* ja *B* erotus

- Olkoot $A \subset S$ ja $B \subset S$ otos-avaruuden S tapahtumia.
- Yhdistetty tapahtuma

  $A \backslash B$ = *"A sattuu, mutta B ei satu"*

  on niiden alkeistapahtumien joukko, jotka kuuluvat joukkoon A, *mutta eivät* kuulu joukkoon B:

  $A \backslash B$
  $= \{s \in S \mid s \in A \text{ ja } s \notin B\}$
  $= A \cap B^c$



Todennäköisyyslaskenta

Osa 1: Todennäköisyys ja sen laskusäännöt

- Satunnaiskokeet, otosavaruudet ja tapahtumat
- **Todennäköisyyden määritteleminen**

Todennäköisyyden määritteleminen

- Käsittelemme satunnaiskokeita joten kokeen alkutilan perusteella *ei voida* tarkasti *ennustaa* mikä mahdollisista *tulosvaihtoehdoista toteutuu*.

- Haluamme kuitenkin osata jotenkin sanoa millaiset mahdollisuudet on sille, että satunnaiskokeen joku tapahtuma sattuu. Todennäköisyys mittaa tätä mahdollisuutta.

Todennäköisyyden aksiomaattinen määrittely

- Matemaattisesti kelvollisen yleisen määritelmän todennäköisyydelle esitti venäläinen matematiikko *A. N. Kolmogorov* 1930-luvun alussa.
- *Kolmogorovin aksioomien* mukaan todennäköisyyslaskenta on *matemaattisen mittateorian* osa.

- Todennäköisyyden naiivit määritelmät (ks kohdat **klassinen ja empiirinen todennäköisyys** muutaman kalvon päästä) voidaan sijoittaa – sopivasti muotoiltuina – Kolmogorovin aksioomajärjestelmään todennäköisyyden käsitteen *tulkintoina* tai *kuvauksina*.

Todennäköisyyden aksioomat

Olkoon S *otosavaruus*, jossa satunnaisilmiötä tarkastellaan.

(i) $$\boxed{\Pr(S) = 1}$$

(ii) *Jokaisen* **tapahtuman** $A \subset S$ **todennäköisyys** Pr(A) on reaaliluku välillä [0,1]:

$$\boxed{0 \leq \Pr(A) \leq 1}$$

(iii) Jos $A_1, A_2, A_3, \ldots$ ovat *toisensa poissulkevia tapahtumia* *(* $A_i \cap A_j = \varnothing, i \neq j$), sitten:

$$\boxed{\Pr(A_1 \cup A_2 \cup A_3 \ldots) = \Pr(A_1) + \Pr(A_2) + \Pr(A_3) + \ldots..}$$

Klassinen todennäköisyys: Määritelmä

- Tarkastellaan satunnaiskoetta, johon liittyy n kpl yhtä todennäköisiä vaihtoehtoja:

  $S = \{s_1, s_2, \dots , s_n\}$ *äärellinen otosavaruus*.

- Tarkastellaan *tapahtumaa* $A \subset S$, johon kuuluu k alkeis-tapahtumaa.

- Tällöin tapahtuman A **klassinen todennäköisyys** on

$$\Pr(A) = \frac{k}{n}$$

Klassinen todennäköisyys - Esimerkki

- Heitetään noppaa.
- Tällöin otosavaruus on

$$S = \{1, 2, 3, 4, 5, 6\}.$$

- Jos oletetaan, että noppa on *virheetön*, niin

$$\Pr(i) = \frac{1}{6}, \text{ kaikille silmäluvuille } i = 1, 2, 3, 4, 5, 6$$

- Olkoon tapahtuma

$$A = \{5, 6\} \subset S.$$

- Tapahtumalle A *suotuisien* alkeistapahtumien lukumäärä $k = 2$.
- Siten tapahtuman A todennäköisyys on

$$\Pr(A) = \frac{2}{6} = \frac{1}{3} \approx 0.333$$

- Tulosvaihtoehtojen *lukumäärien laskeminen* on usein epätriviaali tehtävä ja apuna tarvitaan *kombinatoriikaksi* kutsuttua matematiikan osa-aluetta; ks. **Klassinen todennäköisyys ja kombinatoriikka** Esimerkkikokoelmassa 1.

Klassinen todennäköisyys: Ongelmat määritelmässä

- Klassisen todennäköisyyden määritelmä *ei anna* mahdollisuutta puhua sellaisten tapahtumien toden-näköisyyksistä, *joihin liittyvät tulosvaihtoehdot eivät ole yhtä todennäköisiä*.

- Määritelmä *ei anna* mahdollisuutta puhua sellaisten tapahtumien todennäköisyyksistä, joihin liittyy *äärettömän monta tulosvaihtoehtoa*.

Empiirinen todennäköisyys

- Tarkastellaan *satunnaiskoetta*, jota voidaan *toistaa* siten, että seuraavat ehdot pätevät:
  - (i) Kokeen olosuhteet *säilyvät muuttumattomina* koetoistosta toiseen.
  - (ii) Koetoistot *ovat riippumattomia*.

Empiirinen todennäköisyys : Määritelmä

- *Toistetaan satunnaiskoetta n* kertaa.
- Oletetaan, että tapahtuma *A* sattuu koetoistojen aikana *f* kertaa.
- Jos tapahtuman *A suhteellinen frekvenssi*

$$\frac{f}{n}$$

lähestyy (jossakin mielessä) jotakin kiinteätä lukua *p* koetoistojen lukumäärän *n* kasvaessa rajatta, on *p* tapahtuman *A* **empiirinen todennäköisyys**.

Empiirinen todennäköisyys: Esimerkki 1/3

- Eräässä *kyselytutkimuksessa* selvitettiin miten suomalaiset suhtautuvat Suomen mahdolliseen NATO-jäsenyyteen.
- Tutkimus perustui *satunnaisotokseen*, johon poimittiin *arpomalla*

  1800

  suomalaista.
- *Otoksessa*

  1080

  henkilöä ilmoitti vastustavansa NATO-jäsenyyttä.

Empiirinen todennäköisyys: Esimerkki 2/3

- Tutkimusta voidaan kuvata satunnaisilmiönä seuraavalla tavalla:

  Satunnaiskoe:

  Poimitaan yksi suomalainen arpomalla otokseen.

  Koetoistojen lukumäärä (otoskoko):

  $n = 1800$

  Tapahtuma A:

  Otokseen poimittu suomalainen vastustaa Suomen NATO-jäsenyyttä.

  Tapahtuman *A frekvenssi* koetoistojen joukossa:

  $f = 1080$

  Tapahtuman *A suhteellinen frekvenssi*:

  $$\frac{f}{n} = \frac{1080}{1800} = 0.6$$

Empiirinen todennäköisyys: Esimerkki 3/3

- Jos otoksen poiminnassa käytettiin arvontaa, voidaan olettaa, että tapahtuman A suhteellinen frekvenssi säilyy *stabiilina*, jos *otoskokoa kasvatetaan* tai *otantaa toistetaan*.
- Jos oletus tapahtuman A suhteellisen frekvenssin stabiiliudesta pätee, havaittua suhteellista frekvenssiä 0.6 on järkevää kutsua *todennäköisyydeksi*, että satunnaisesti valittu suomalainen vastustaa Suomen NATO-jäsenyyttä.
- Siten tapahtuman A *empiirinen todennäköisyys* on

$$\Pr(A) = 0.6$$

otoksesta saatujen tietojen perusteella.

Empiirinen todennäköisyys: Ongelmat määritelmässä 1/2

- Empiirinen todennäköisyys on *empiirinen käsite* siinä mielessä, että tulosvaihtoehdon suhteellisen frekvenssin f/n määrääminen vaatii satunnaiskokeen *toistamista* ja *havaintojen keräämistä* satunnaiskokeen tuloksista.
- Tulosvaihtoehdon *empiiristä todennäköisyyttä ei voida* – nimestään huolimatta – *määrätä kokeellisesti*, koska suhteellisen frekvenssin tilastollisen stabiliteetin todentaminen vaatisi satunnaiskokeen toistamista *äärettömän* monta kertaa.
- Käsite *ei anna* mahdollisuutta puhua sellaisten tapahtumien todennäköisyyksistä, *joista ei ole havaintoja*.

Empiirinen todennäköisyys: Ongelmat määritelmässä 2/2

- Empiirisen todennäköisyyden määritelmässä esiintyvä suhteellisen frekvenssin raja-arvo *ei ole hyvin määritelty*:

  Mikään ei takaa, että määritelmässä esiintyvä raja-arvo *on olemassa*.

- Empiiristä todennäköisyyttä voidaan pikemminkin pitää *tilastollisesti stabiilisti käyttäytyvän suhteellisen frekvenssin ominaisuutena* kuin *todennäköisyyden määritelmänä*.

Todennäköisyyslaskenta

Osa 1: Todennäköisyys ja sen laskusäännöt

-Peruslaskusäännöt todennäköisyydelle

- Ehdollinen todennäköisyys ja riippumattomuus

Todennäköisyyslaskennan peruslaskutoimitukset ja -säännöt

- *Todennäköisyyslaskennalla tarkoitetaan* usein sellaisten **laskutoimitusten** ja **-sääntöjen kokoelmaa**, joiden avulla voidaan määrätä jonkin satunnaisilmiön tapahtumista *joukko-opin operaatioiden avulla johdettujen* **uusien tapahtumien** todennäköisyydet.

- Todennäköisyyslaskennan laskusäännöt *voidaan todistaa Kolmogorovin aksioomajärjestelmässä*.

- *Tässä ei anneta todistuksia* mutta *säännöt tehdään ilmeisiksi Venn-diagrammien* ja *esimerkkien avulla*.

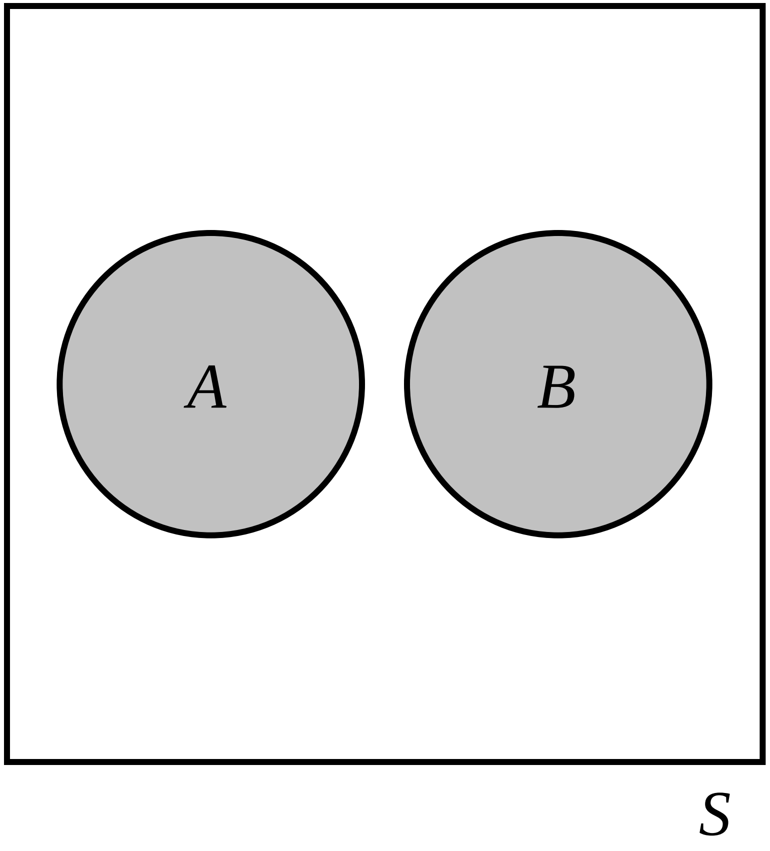
Yhteenlaskusääntö toisensa poissulkeville tapahtumille

- Olkoon tapahtuman A todennäköisyys $\Pr(A)$.
- Olkoon tapahtuman B todennäköisyys $\Pr(B)$.
- Jos A ja B ovat *toisensa poissulkevia*, **yhdisteen**

  $A \cup B$ = ”*A tai B sattuu*”

  todennäköisyys on

  $\Pr(A \cup B) = \Pr(A) + \Pr(B)$



Peruslaskusäännöt todennäköisyydelle

Yhteenlaskusääntö toisensa poissulkeville tapahtumille: Esimerkki eduskunnasta

- Vuoden 2011 eduskunnassa (200 kansanedustajaa yhteensä):

$$n_{Sosialistit} = n_{SDP} + n_{Vas} = 42 + 14 = 56$$

- Todennäköisyys, että satunnaisesti valittu edustaja oli sosialisti:

$$\begin{aligned}\Pr(Sosialisti) &= \Pr(SDP) + \Pr(vas)\\ &= \frac{42}{200} + \frac{14}{200}\\ &= \frac{56}{200}\\ &= 0.28.\end{aligned}$$

Yleistetty yhteenlaskusääntö pareittain toisensa poissulkeville tapahtumille

- Olkoot $A_1, A_2, \ldots, A_k$ **pareittain toisensa poissulkevia**.
- Tällöin $A_i \cap A_j = \emptyset$, kun $i \neq j$.
- Olkoon tapahtuman A_i todennäköisyys

  $\Pr(A_i),\ i = 1, 2, \ldots, k.$
- Tällöin **yhdisteen**

  ”A_1 *tai* A_2 *tai … tai* A_k *sattuu*”

  todennäköisyys on

  $$\Pr(A_1 \cup A_2 \cup \cdots \cup A_k) = \Pr(A_1) + \Pr(A_2) + \cdots + \Pr(A_k)$$

Mahdottoman tapahtuman todennäköisyys:

- Todennäköisyyden *aksioomien* (i) ja (iii) mukaan

$$1 = \Pr(S) = \Pr(\emptyset \cup S) = \Pr(\emptyset) + \Pr(S) = \Pr(\emptyset) + 1$$

josta välttämättä seuraa

$$\boxed{\Pr(\emptyset) = 0}$$

Komplementtitapahtuman todennäköisyys 1/2

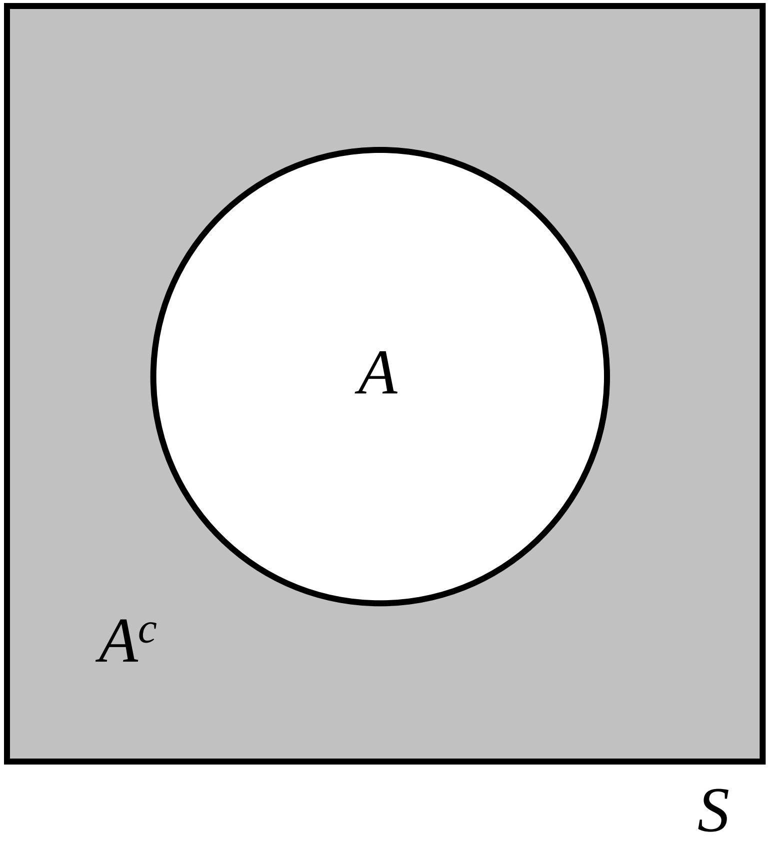
- Olkoon $A \subset S$ otos-avaruuden S tapahtuma.
- Olkoon

  A^c

  $=$ "*A ei satu*"

  $= \{s \in S \mid s \notin A\}$

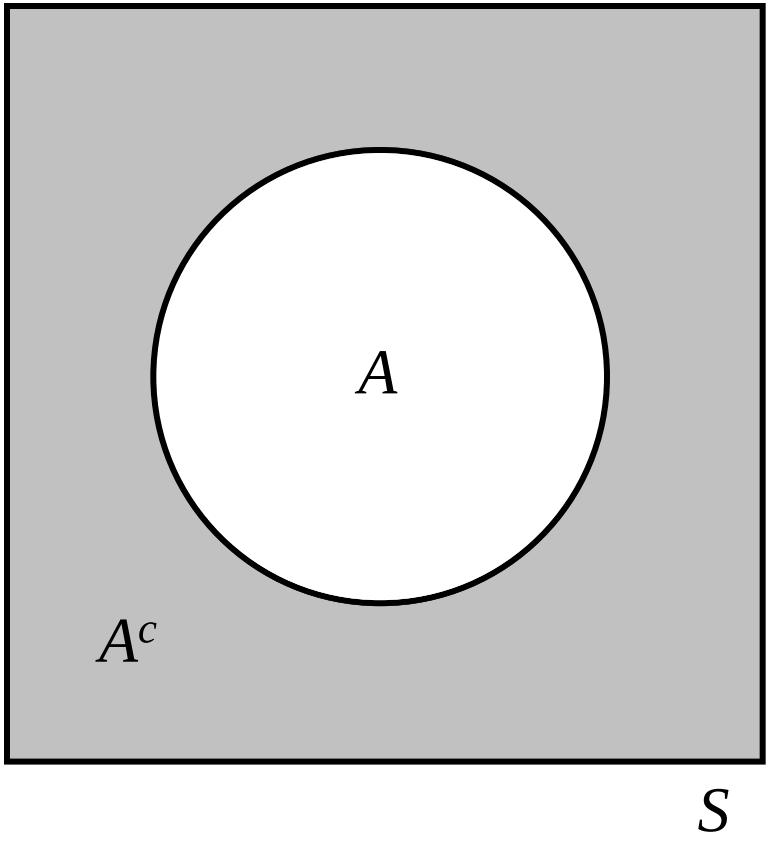
  tapahtuman A **komplementtitapahtuma**.



Komplementtitapahtuman todennäköisyys 2/2

- Olkoon tapahtuman A todennäköisyys $\Pr(A)$.
- Tällöin tapahtuman A **komplementtitapahtuman** A^c **todennäköisyys** on

$$\boxed{\Pr(A^c) = 1 - \Pr(A)}$$



Komplementtitapahtuman todennäköisyys: Perustelu

- Tapahtuma A ja sen *komplementti-tapahtuma* A^c ovat toisensa pois-sulkevia:

  $A \cap A^c = \varnothing$

- Lisäksi otosavaruudelle S pätee aina:

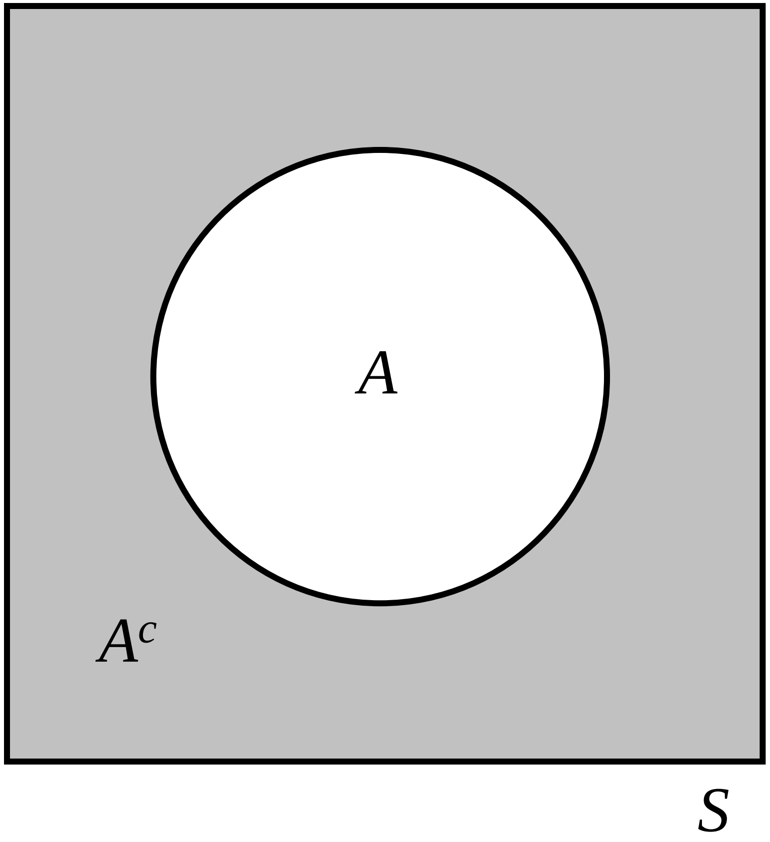
  $S = A \cup A^c$

- Siten *toisensa poissulkevien tapahtumien yhteenlaskusäännön mukaan*

  $\Pr(S) = 1 = \Pr(A) + \Pr(A^c)$

- Siten *komplementtitapahtuman* A^c *todennäköisyydeksi* saadaan:

  $\Pr(A^c) = 1 - \Pr(A)$



Komplementtitapahtuman todennäköisyys: Esimerkki

Heitetään 2 kolikkoa yhtä aikaa. Mikä on tapahtuman A = saadaan vähintään yhden Kruunan todennäköisyys ? (Oletetaan että kolikko on virheetön).

$n(S) = 4$

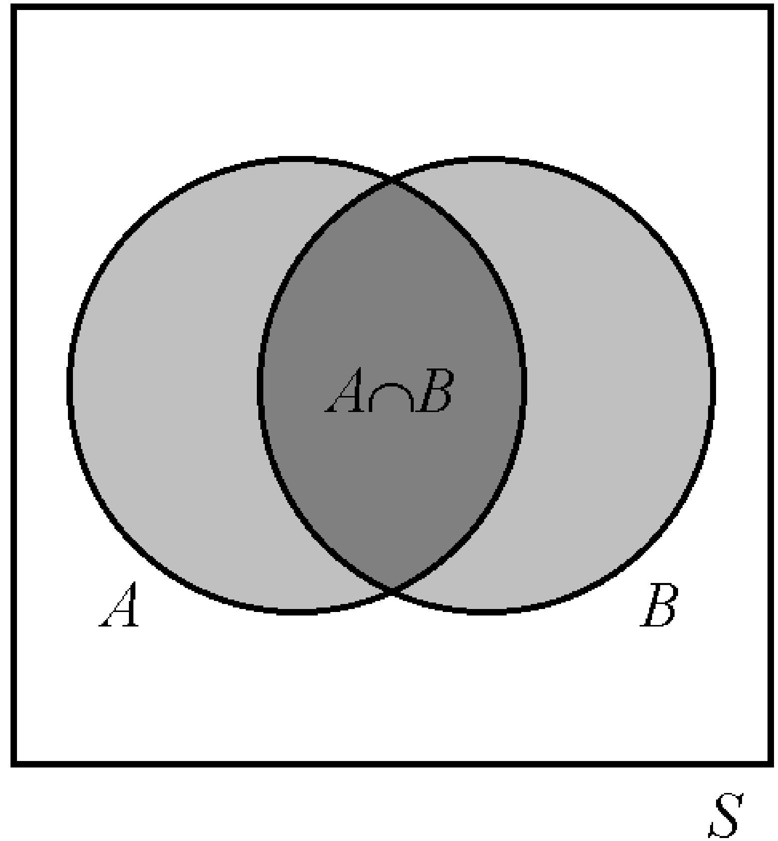
$\Pr(A^c)$ = todennäköisyys ettei saada yhtään Kruunaa = ¼

$\Pr(A) = 1 - \Pr(A^c) = ¾$

Yleinen yhteenlaskusääntö

- Olkoot tapahtumien A, B, $A \cap B$ todennäköisyydet $\Pr(A)$, $\Pr(B)$, $\Pr(A \cap B)$
- Tällöin **yhdisteen** $A \cup B$ = ”*A tai B sattuu*” **todennäköisyys** on

$$\boxed{\begin{aligned}&\Pr(A \cup B)\\ &= \Pr(A) + \Pr(B) - \Pr(A \cap B)\end{aligned}}$$



Yleinen yhteenlaskusääntö: Esimerkki levikkitutkimuksesta

- *Levikkitutkimuksessa* saatiin selville, että erään kunnan asukkaat lukevat Seuraa ja Apua seuraavasti:

Seura	20 %
Apu	16 %
Seura ja Apu	1 %

- Tällöin

$$\begin{aligned}
\Pr(\text{Seura } \textit{tai} \text{ Apu}) &= \Pr(\text{Seura}) + \Pr(\text{Apu}) - \Pr(\text{Seura } \textit{ja} \text{ Apu}) \\
&= \frac{20}{100} + \frac{16}{100} - \frac{1}{100} \\
&= \frac{35}{100} \\
&= 0.35
\end{aligned}$$

Erotustapahtuman todennäköisyys 1/2

- Olkoot $A \subset S$ ja $B \subset S$ otos-avaruuden S tapahtumia.
- Olkoon

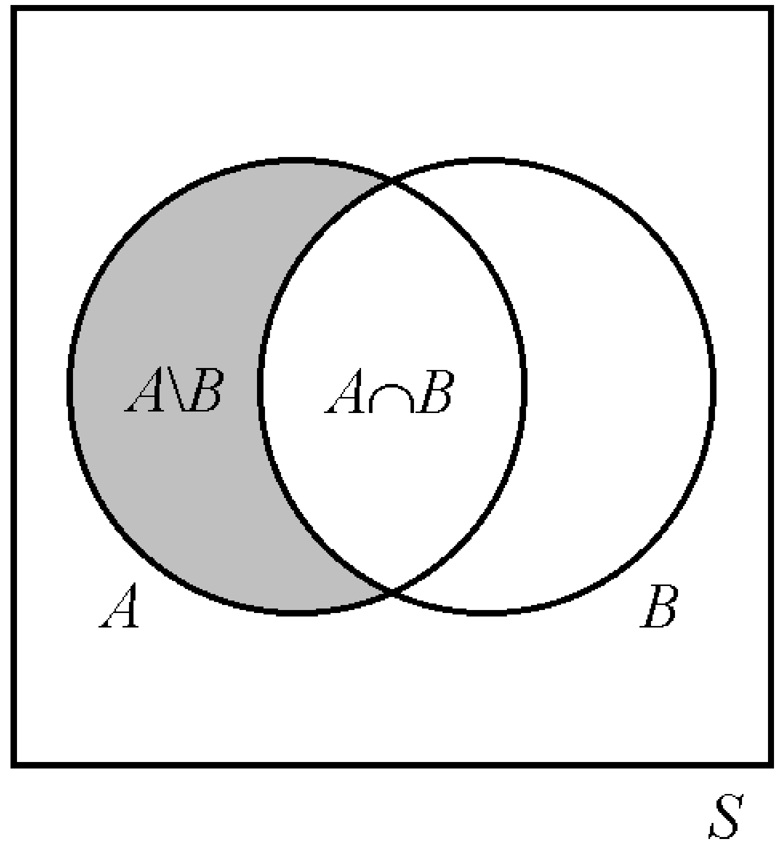
$A \backslash B$

$=$ *"A sattuu, mutta B ei satu"*

$= \{s \in S \mid s \in A \text{ ja } s \notin B\}$

$= A \cap B^c$

tapahtumien A ja B **erotus**.



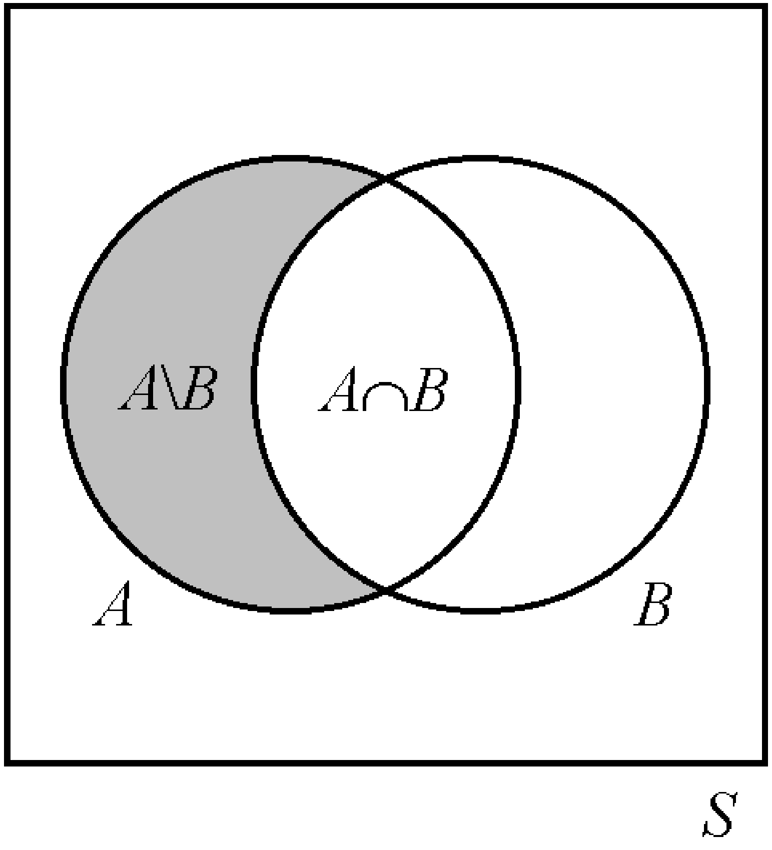
Erotustapahtuman todennäköisyys 2/2

- Olkoon tapahtuman A todennäköisyys $\Pr(A)$.
- Olkoon tapahtuman $A \cap B$ todennäköisyys $\Pr(A \cap B)$.
- Tällöin **erotustapahtuman**

  $A \setminus B$ = ”*A sattuu, mutta B ei satu*”

  todennäköisyys on

$$\boxed{\begin{aligned}&\Pr(A \setminus B)\\ &= \Pr(A) - \Pr(A \cap B)\end{aligned}}$$



Erotustapahtuman todennäköisyys: Esimerkki korttipakasta 1/2

- 52:n kortin pakassa on 16 kuvakorttia ja 13 patakorttia, joista 4 on kuvakortteja.
- Olkoon

  A = ”Satunnaisesti valittu kortti on pata”
  B = ”Satunnaisesti valittu kortti on kuva”

- Tällöin

  $A \backslash B$ = ”Satunnaisesti valittu kortti on pata, *mutta* ei ole kuvakortti”

Erotustapahtuman todennäköisyys: Esimerkki korttipakasta 2/2

- Tällöin

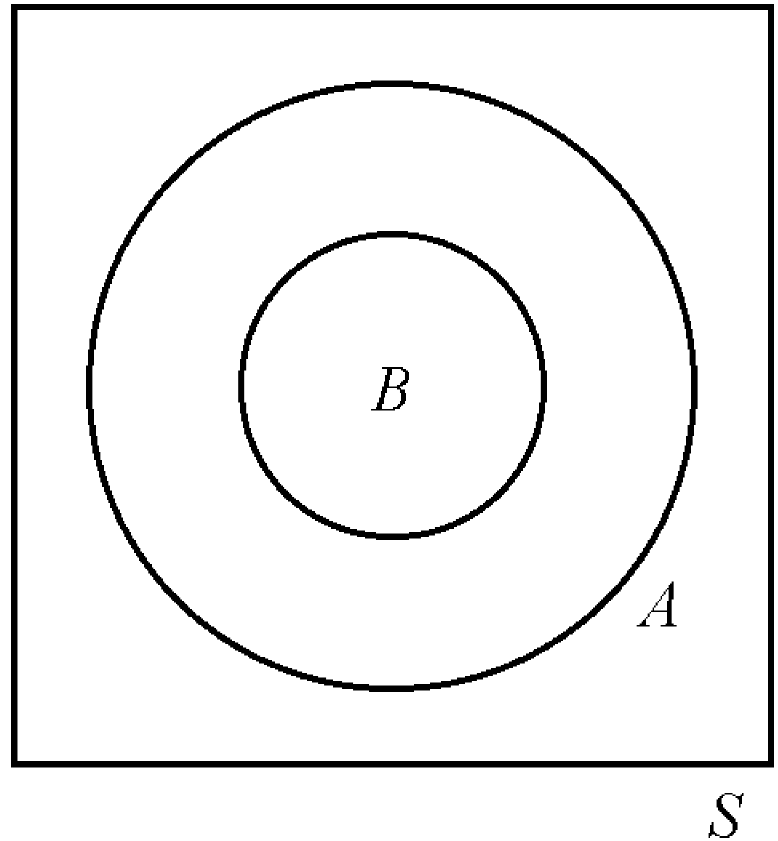
$$\begin{aligned}\Pr(A \setminus B) &= \Pr(A) - \Pr(A \cap B)\\ &= \frac{13}{52} - \frac{4}{52}\\ &= \frac{9}{52}\end{aligned}$$

- Tulos on tietysti selvä muutenkin, koska patoja, jotka eivät ole kuvia, on 9 kpl eli kortit 10, 9, 8, 7, 6, 5, 4, 3, 2.

Tapahtuman *B* sattumisesta seuraa tapahtuman *A* sattuminen

- Olkoot $A \subset S$ ja $B \subset S$ otos-avaruuden S tapahtumia.
- Oletetaan, että *jos B sattuu*, *niin A sattuu*.
- Tällöin $B \subset A$.
- Olkoot tapahtumien A ja B todennäköisyydet $\Pr(A)$ ja $\Pr(B)$.
- Tällöin:

$\Pr(A) \geq \Pr(B)$



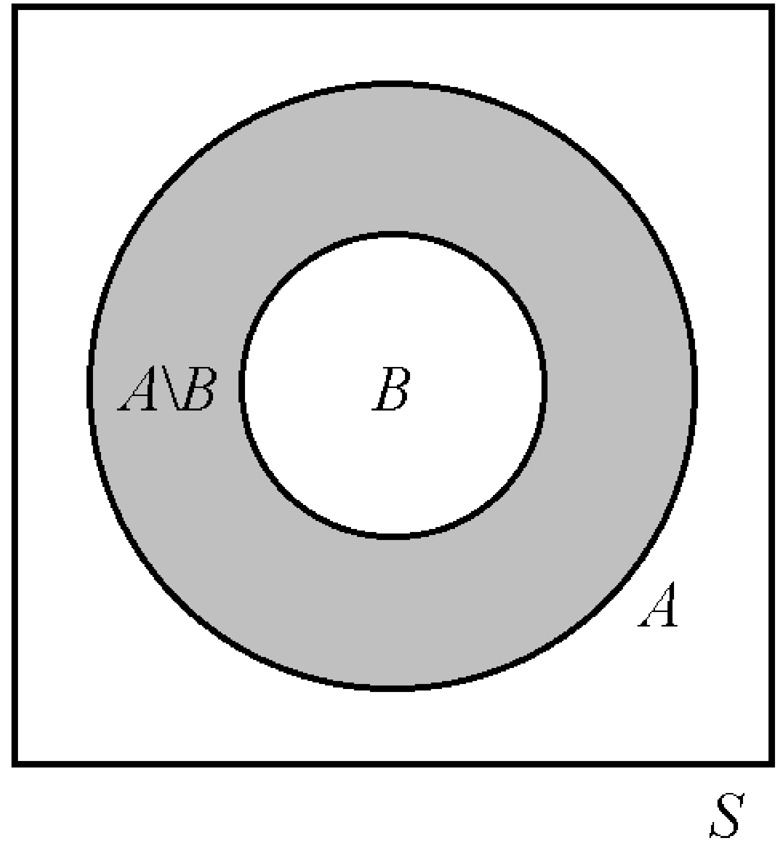
Erotustapahtuman todennäköisyys, kun *B*:n sattumisesta seuraa *A*:n sattuminen

- Olkoot $A \subset S$ ja $B \subset S$ otos-avaruuden S tapahtumia.
- Olkoon $B \subset A$.
- Olkoot tapahtumien A ja B todennäköisyydet $\Pr(A)$ ja $\Pr(B)$.
- Tällöin **erotustapahtuman**

  $A \backslash B$ = ”*A sattuu,*
  mutta B ei satu”

  todennäköisyys on

  $\Pr(A \backslash B) = \Pr(A) - \Pr(B)$



Todennäköisyyslaskenta

Osa 1: Todennäköisyys ja sen laskusäännöt

-Peruslaskusäännöt todennäköisyydelle

- Ehdollinen todennäköisyys ja riippumattomuus

Ehdollinen todennäköisyys

- Joskus haluamme tietää tapahtuman *A* todennäköisyys ehdolla, että tapahtuma *B* on sattunut. Tämä todennäköisyys on *tapahtuman A* **ehdollinen todennäköisyys** *ehdolla, että tapahtuma B on sattunut* ja se kirjoitetaan $\Pr(A|B)$.

- Tässä tilanteessa *B* on uusi (pienempi) otosavaruus, ja tämä ehdollinen todennäköisyys on se murto-osa $\Pr(B)$:sta joka vastaa tapahutmaa $A \cap B$.

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

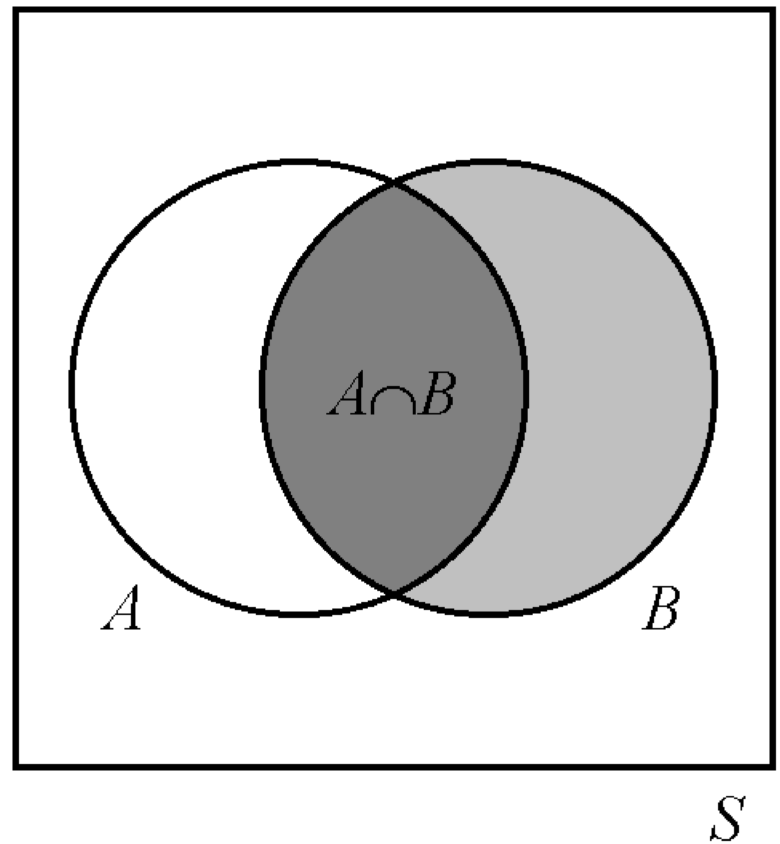
Ehdollinen todennäköisyys

- Olkoon tapahtuman

  ”*A ja B sattuvat*”

  todennäköisyys $\Pr(A \cap B)$.
- Olkoon tapahtuman B todennäköisyys $\Pr(B) \neq 0$.
- Tällöin *tapahtuman A* **ehdollinen todennäköisyys** *ehdolla, että tapahtuma B on sattunut* on

$$\boxed{\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}}$$



Ehdollinen todennäköisyys: Esimerkki eduskunnasta 1/2

- Vuoden 2011 eduskunnan 200 kansanedustajasta 86 on naisia.
- Mikä on todennäköisyys, että satunnaisesti valittu edustaja oli nainen?
- Merkitään

  A = ”Edustaja oli nainen”
- Tällöin

$$\begin{aligned}\Pr(A) &= \frac{86}{200} \\ &= 0.43\end{aligned}$$

Ehdollinen todennäköisyys: Esimerkki eduskunnasta 2/2

- SDP:lla on 42 kansanedustajaa, 15 miestä, 27 naista.
- Mikä on todennäköisyys, että satunnaisesti valittu edustaja, joka kuului SDP:n ryhmään, oli nainen?
- Tällöin *ehdollinen todennäköisyys*

$$\Pr(\text{Nainen} \mid \text{SDP}) = \frac{\Pr(\text{Nainen } \textit{ja} \text{ SDP})}{\Pr(\text{SDP})}$$

$$= \frac{27/200}{42/200} = \frac{27}{42} = 0.643 > 0.43 = \Pr(\text{Nainen})$$

- *Tässä tapauksessa* tieto siitä, että satunnaisesti valittu kansanedustaja oli SDP:stä muuttaa todennäköisyyttä, että hän oli nainen.

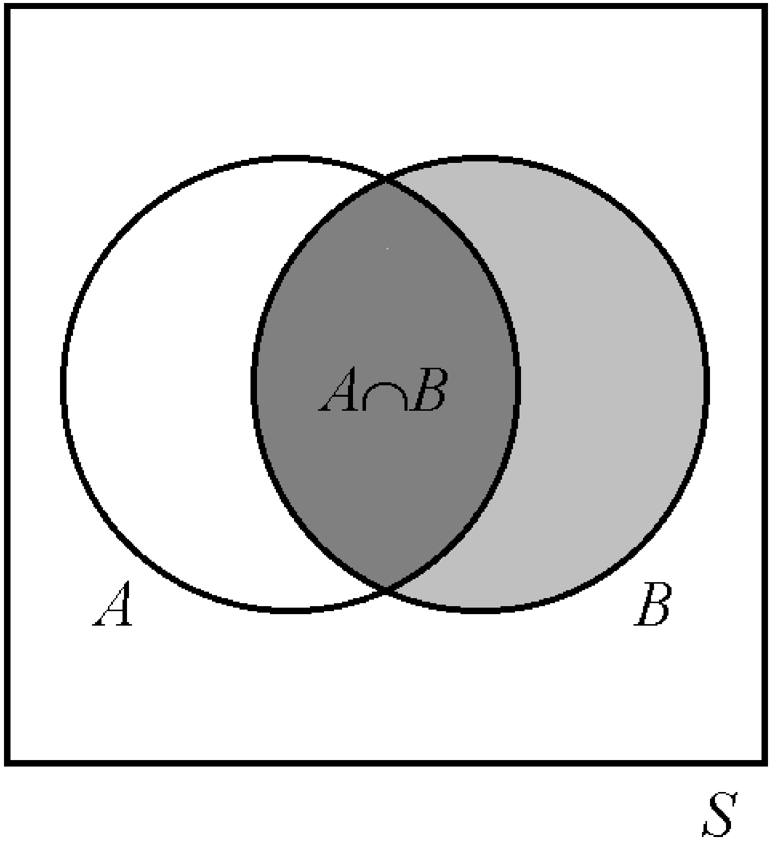
Yleinen tulosääntö

- Olkoon tapahtuman A *ehdollinen todennäköisyys ehdolla, että tapahtuma B on sattunut* $\Pr(A|B)$.
- Olkoon tapahtuman B todennäköisyys $\Pr(B) \neq 0$.
- Tällöin **leikkauksen**

  $A \cap B$ = ”*A ja B sattuvat*”

  todennäköisyys on

$$\boxed{\Pr(A \cap B) = \Pr(B)\Pr(A|B)}$$



Yleistetty yleinen tulosääntö

- Tarkastellaan tapahtumia $A_1, A_2, \ldots, A_k$.
- Tällöin **leikkauksen**

$$A_1 \cap A_2 \cap \cdots \cap A_k = \text{"}A_1 \textit{ ja } A_2 \textit{ ja } \ldots \textit{ ja } A_k \textit{ sattuvat"}$$

todennäköisyys on

$$\begin{aligned}
&\Pr(A_1 \cap A_2 \cap \cdots \cap A_k) \\
&\quad = \Pr(A_1) \times \Pr(A_2 \mid A_1) \times \Pr(A_3 \mid A_1 \cap A_2) \times \\
&\quad\quad \cdots \times \Pr(A_k \mid A_1 \cap A_2 \cap \cdots \cap A_{k-1})
\end{aligned}$$

Yleistetty yleinen tulosääntö: Perustelu

- Perustellaan *yleistetty yleinen tulosääntö* tapauksessa $k = 3$.
- Leikkauksen

  $A_1 \cap A_2 \cap A_3$ = ”A_1 *ja* A_2 *ja* A_3 *sattuvat*”

  todennäköisyys on

$$\begin{aligned}\Pr(A_1 \cap A_2 \cap A_3) &= \Pr((A_1 \cap A_2) \cap A_3) \\ &= \Pr(A_1 \cap A_2)\Pr(A_3 \mid A_1 \cap A_2) \\ &= \Pr(A_1)\Pr(A_2 \mid A_1)\Pr(A_3 \mid A_1 \cap A_2)\end{aligned}$$

- *Yleinen tapaus* voidaan todistaa *induktiolla*.

Yleistetty yleinen tulosääntö: Esimerkki korttipakasta

- Nostetaan korttipakasta *peräkkäin* 3 korttia.
- Mikä on todennäköisyys, että ne ovat kaikki patoja?
- Olkoon tapahtuma A_i = ”*i*. kortti on pata”, $i = 1, 2, 3$.
- Koska korttipakassa on 52 korttia, joista 13 on patoja, leikkauksen ”A_1 *ja* A_2 *ja* A_3 *sattuvat*” todennäköisyys on

$$\begin{aligned}\Pr(A_1 \cap A_2 \cap A_3) &= \Pr(A_1)\Pr(A_2 \mid A_1)\Pr(A_3 \mid A_1 \cap A_2) \\ &= \frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} = \frac{33}{2550} \\ &= 0.0129\end{aligned}$$

- Huomautus:

  Korttien nosto toteutettiin **ilman takaisinpanoa**.

Riippumattomuus

- Olkoot A ja B kaksi samaan satunnaisokeeseen liittyvää tapahtumaa. Eräs keino tutkia tapahtumien A ja B syy-seuraussuhdetta on verrata esimerkiksi A:n todennäköisyyttä Pr(A) ehdolliseen todennäköisyyteen $\Pr(A|B)$.

- Jos

$$\Pr(A|B) \neq \Pr(A)$$

tieto siitä, että tapahtuma B on sattunut, sisältää informaatiota, jota voidaan käyttää hyväksi tapahtuman A todennäköisyyttä määrättäessä.

Riippumattomuus

- Tapahtuma *A* on **riippumaton** tapahtumasta *B*, jos *B*:n tapahtuminen (tai tapahtumatta jääminen) *ei vaikuta A*:n tapahtumisen todennäköisyyteen.

- Riippumattomuus on *symmetrinen* ominaisuus -

  jos *A* on riippumaton *B*:stä, niin *B* on riippumaton *A*:sta – ja sanomme tavallisesti, että *tapahtumat A* ja *B ovat riippumattomia*.

Riippumattomuus: Esimerkit rahanheitosta ja arvonnasta

- **Rahanheitto:**

  Heitetään toistuvasti rahaa.

  Tällöin on järkevää olettaa, että heittojen tulokset eivät riipu aikaisemmin tehtyjen heittojen tuloksista.

- **Arvonta:**

  Nostetaan uurnasta toistuvasti arpalippuja niin, että jokaisen noston jälkeen nostettu lippu palautetaan uurnaan ja uurnan sisältö sekoitetaan huolellisesti **(otanta takaisinpanolla).**

  Tällöin on järkevää olettaa, että nostojen tulokset eivät riipu aikaisemmin tehtyjen nostojen tuloksista.

Tulosääntö riippumattomille tapahtumille

Olkoon tapahtuman A todennäköisyys $\Pr(A)$ ja tapahtuman B todennäköisyys $\Pr(B)$. Matemaattisesti määritellään, että tapahtumat

A ja B ovat **riippumattomia** jos ja vain jos

$$\Pr(A \cap B) = \Pr(A)\Pr(B) \qquad (1)$$

Huom. Tämä ehto on yhtäpitävä ehdon $\Pr(A|B) = \Pr(A)$ kanssa – nähdään sijoittamalla (1) kaavaan

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

Riippumattomuuden yhtäpitävät ehdot

- Tapahtumat A ja B ovat **riippumattomia**, jos ja vain jos mikä tahansa seuraavista *riippumattomuuden yhtäpitävistä ehdoista* pätee:

  (i) $\Pr(A \cap B) = \Pr(A)\Pr(B)$

  (ii) $\Pr(A \mid B) = \Pr(A)$

  (iii) $\Pr(B \mid A) = \Pr(B)$

Riippumattomuuden seurauksia

- Oletetaan, että tapahtumat A ja B ovat riippumattomia.
- Tällöin:

  (i) Tapahtumat A ja B^c ovat riippumattomia.

  (ii) Tapahtumat A^c ja B ovat riippumattomia.

  (iii) Tapahtumat A^c ja B^c ovat riippumattomia.

- Tapahtumien riippumattomuus siis ”*leviää*” myös niiden *komplementtitapahtumiin*.

Tulosääntö riippumattomille tapahtumille: Esimerkki rahanheitosta

- Heitetään rahaa kaksi kertaa.

  Pr(Kruuna) = Pr(Klaava) =1/2
  A = ”Saadaan kruuna 1. heitolla”
  B = ”Saadaan kruuna 2. heitolla”

- Voidaan olettaa, että *A* ja *B* ovat riippumattomia.
- Tällöin

$$\Pr(A \ \ ja \ \ B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0.25$$

Tulosääntö riippumattomille tapahtumille: Esimerkki korttipakasta

- A = ”Satunnaisesti korttipakasta valittu kortti on pata”

  $\Pr(A) = 13/52 = 1/4$
- B = ”Satunnaisesti korttipakasta valittu kortti on ässä”

  $\Pr(B) = 4/52 = 1/13$
- $A \cap B$ = ”Satunnaisesti korttipakasta valittu kortti on pataässä”
- Tällöin

$$\begin{aligned}\Pr(A \cap B) &= \Pr(A)\Pr(B) \\ &= \frac{1}{4} \times \frac{1}{13} \\ &= \frac{1}{52}\end{aligned}$$

Yleistetty tulosääntö riippumattomille tapahtumille

- Olkoon tapahtuman A_i todennäköisyys

$$\Pr(A_i)\ ,\ i = 1, 2, \ldots, k$$

- Tapahtumat A_1, A_2, … , A_k ovat **riippumattomia**, jos ja vain jos *kaikille* **leikkauksille**

$$A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}$$

joissa

$$\{i_1, i_2, \ldots, i_m\} \subset \{1, 2, \ldots, k\}$$

pätee:

$$\Pr(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}) = \Pr(A_{i_1}) \times \Pr(A_{i_2}) \times \cdots \times \Pr(A_{i_m})$$

- Merkitään tapahtumien A_1, A_2, … , A_k riippumattomuutta:

$$A_1, A_2, \ldots, A_k \perp$$

Yleistetty tulosääntö riippumattomille tapahtumille: Esimerkki rahanheitosta

- Heitetään rahaa 10 kertaa.
- Pr(Kruuna) = 1/2.
- Tällöin

$$
\begin{aligned}
\Pr(10 \text{ kruunaa}) &= \Pr(\text{Kruuna 1. heitolla}) \times \cdots \times \Pr(\text{Kruuna 10. heitolla}) \\
&= \underbrace{\frac{1}{2} \times \cdots \times \frac{1}{2}}_{\text{10 kappaletta}} = \frac{1}{2^{10}} = \frac{1}{1024} \\
&= 0.000977
\end{aligned}
$$

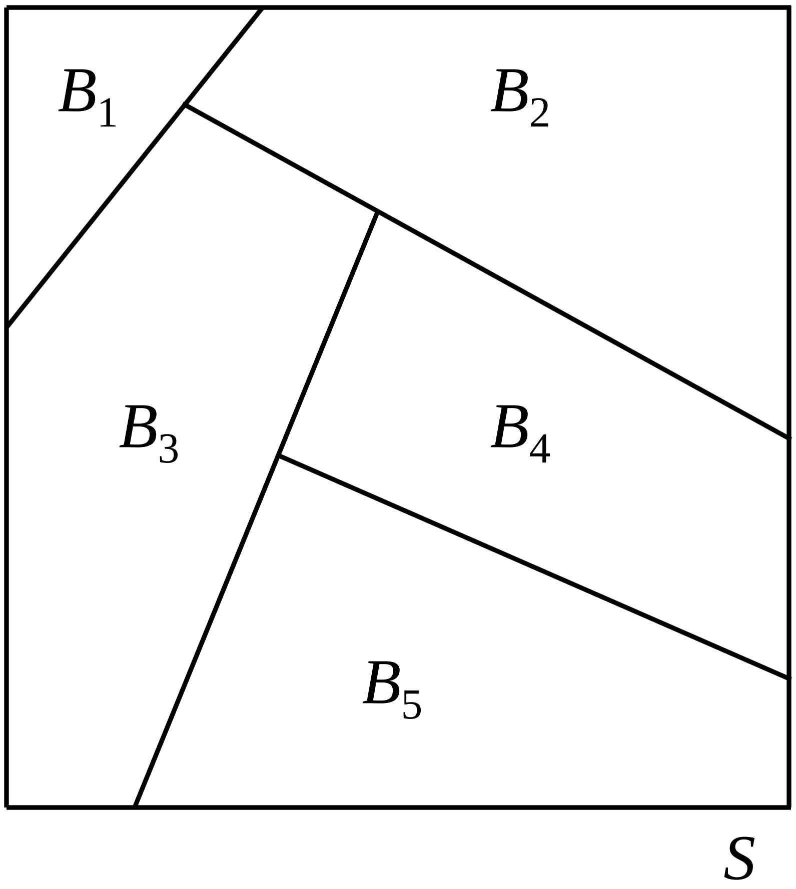
Todennäköisyyslaskenta

Osa 1: Todennäköisyys ja sen laskusäännöt

- **Kokonaistodennäköisyys ja Bayesin kaava**

Otosavaruuden ositus

- *Otosavaruuden S osajoukot*
  $B_1, B_2, \ldots, B_n$
  muodostavat otosavaruuden
  S **osituksen**, jos
  (i) $B_i \neq \emptyset, i = 1, 2, \ldots, n$
  (ii) $B_i \cap B_j = \emptyset, i \neq j$
  (iii) $S = B_1 \cup B_2 \cup \ldots \cup B_n$



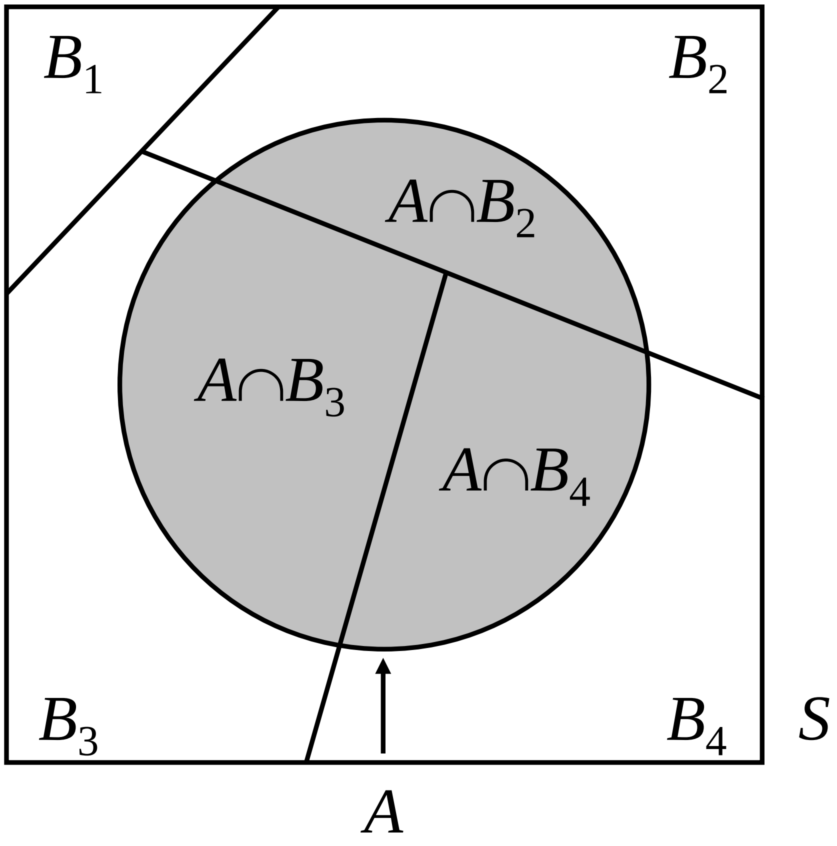
Otosavaruuden osituksen indusoima ositus

- Olkoon $A \subset S$, $A \neq \emptyset$ otosavaruuden S *osajoukko*.
- Olkoon $B_1, B_2, \ldots , B_n$ otosavaruuden S *ositus*.
- Ositus $B_1, B_2, \ldots , B_n$ **indusoi osituksen** joukkoon A:

  $(A \cap B_i) \cap (A \cap B_j) = \emptyset, \ i \neq j$

  ja

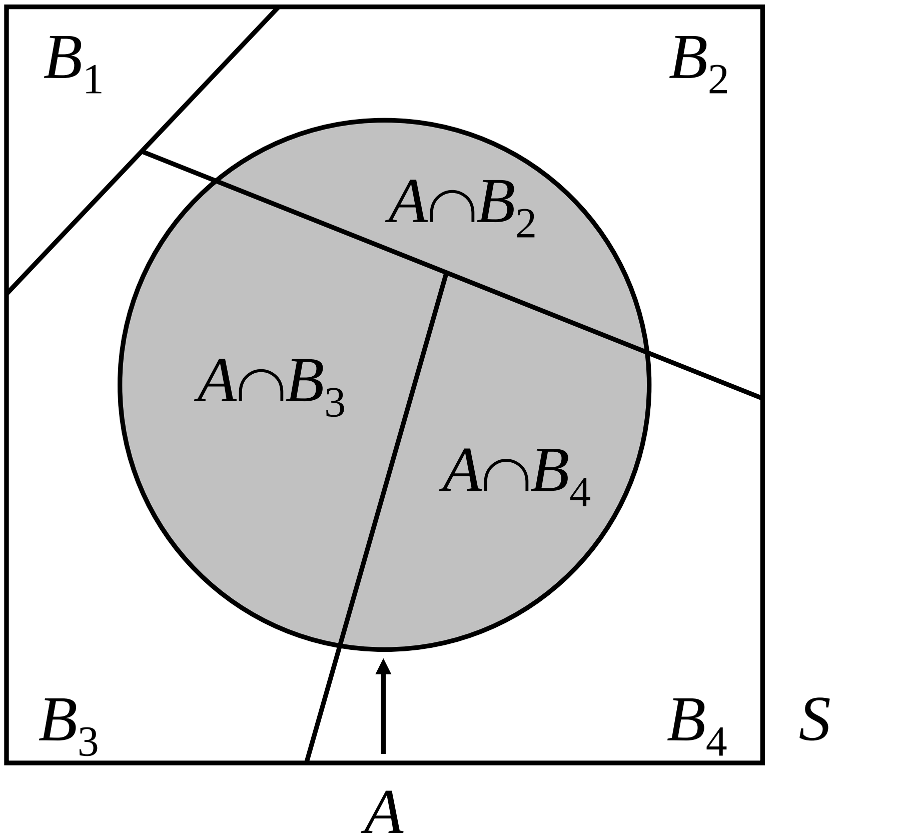
  $A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)$



Kokonaistodennäköisyyden kaava: Määritelmä 1/2

- Olkoon $A \subset S$.
- Olkoon $B_1, B_2, \ldots, B_n$ otosavaruuden S *ositus*.
- *Yhteenlaskusäännön (toisensa poissulkeville tapahtumille)* perusteella

$$\Pr(A) = \sum_{i=1}^{n} \Pr(A \cap B_i) \quad (1)$$



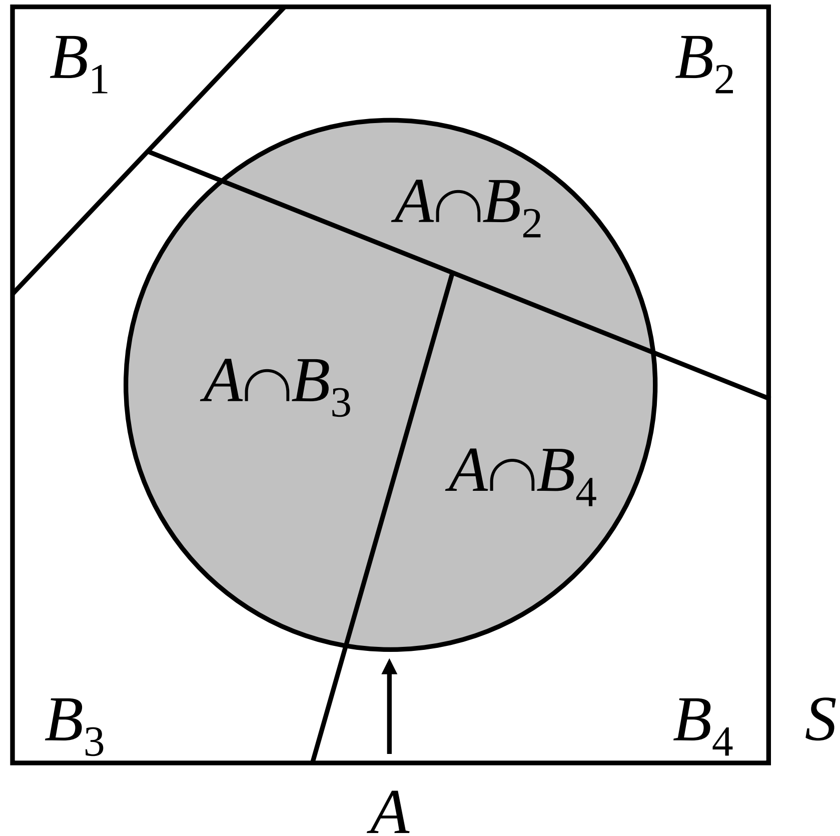
Kokonaistodennäköisyyden kaava: Määritelmä 2/2

- *Yleisen tulosäännön* perusteella

$$\Pr(A \cap B_i) = \Pr(B_i)\Pr(A|B_i)$$

- Sijoittamalla nämä lausekkeet kaavaan (1), saadaan **kokonaistodennäköisyyden** kaava

$$\boxed{\Pr(A) = \sum_{i=1}^{n} \Pr(B_i)\Pr(A|B_i)}$$



Kokonaistodennäköisyyden kaava: Kommentteja

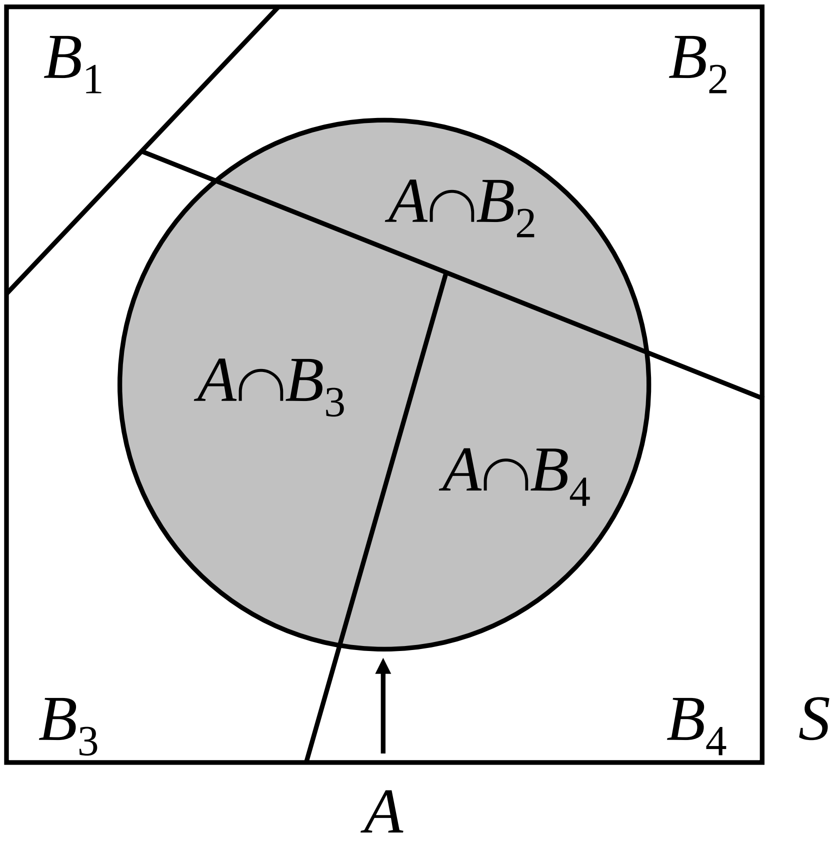
- Kokonaistodennäköisyyden kaava on käyttökelpoinen sellaisissa tilanteissa, joissa todennäköisyydet $\Pr(B_i)$ ja ehdolliset todennäköisyydet $\Pr(A|B_i)$ ovat *tunnettuja*.

- Jos tapahtuma A on *riippumaton* tapahtumista $B_1, B_2, \ldots, B_n$, kokonaistodennäköisyyden kaavasta *ei ole hyötyä* tapahtuman A todennäköisyyttä määrättäessä.

Bayesin kaava: Määritelmä 1/2

- Olkoon $A \subset S$, $A \neq \emptyset$ otos-avaruuden S *osajoukko*.
- Olkoon $B_1, B_2, \ldots, B_n$ otosavaruuden S *ositus*.
- *Ehdollisen todennäköisyyden määritelmän* mukaan

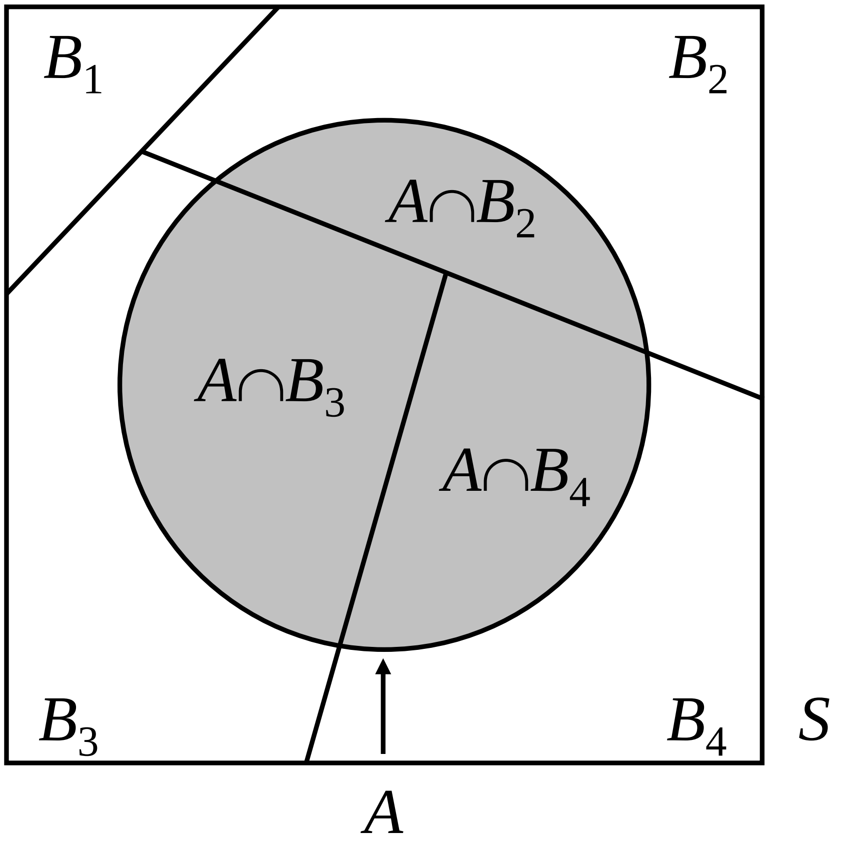
$$\Pr(B_i | A) = \frac{\Pr(A \cap B_i)}{\Pr(A)} = \frac{\Pr(B_i)\Pr(A | B_i)}{\Pr(A)}$$



Bayesin kaava: Määritelmä 2/2

- Soveltamalla nimittäjään *kokonaistodennäköisyyden* kaavaa saadaan **Bayesin kaava**:

$$\Pr(B_i \mid A) = \frac{\Pr(B_i)\Pr(A \mid B_i)}{\sum_{i=1}^{n} \Pr(B_i)\Pr(A \mid B_i)}$$



Bayesin kaava: Kommentteja 1/2

- $\Pr(B_i)$ kutsutaan **priori-todennäköisyydeksi**.

  prior (*lat*.), edeltävä, aikaisempi

  (käsityksemme tapahtuman B_i todennäköisyydestä ennen kuin saamme tietää onko tapahtuma A sattunut vai ei).

- $\Pr(B_i|A)$ kutsutaan **posteriori-todennäköisyydeksi**.

  posterior (*lat*.), jälkeen tuleva, myöhempi

  (miten ennakkokäsitystämme tapahtuman B_i todennäköisyydestä kannattaa korjata, kun saamme tietää, että tapahtuma A on sattunut).

Bayesin kaava: Kommentteja 2/2

- Bayesin kaava on käyttökelpoinen sellaisissa tilanteissa, joissa todennäköisyydet $\Pr(B_i)$ ja ehdolliset todennäköisyydet $\Pr(A|B_i)$ ovat *tunnettuja*.

- Jos tapahtuma A on *riippumaton* tapahtumista $B_1, B_2, \ldots, B_n$, tieto tapahtuman A sattumisesta *ei muuta* prioritodennäköisyyksiä $\Pr(B_i)$.

Esimerkki laadunvalvonnasta 1/4

- Ruuvitehtaalla on kaksi konetta A ja B, joilla tehdään samanlaisia ruuveja. A- ja B-koneen valmistamat ruuvit sekoitetaan ja pakataan laatikoihin *suhteessa* 3:5.
- Osa kummankin koneen valmistamista ruuveista on *viallisia*:

  5 % A-koneen valmistamista ruuveista on viallisia.

  8 % B-koneen valmistamista ruuveista on viallisia.
- Valitaan *satunnaisesti* 1 ruuvi tutkittavaksi satunnaisesti valitusta laatikosta.

  (i) Mikä on todennäköisyys, että poimittu *ruuvi on viallinen*?

  (ii) Mikä on todennäköisyys, että ruuvin on valmistanut A-kone, *jos ruuvi osoittautuu vialliseksi*?

Esimerkki laadunvalvonnasta 2/4

- Merkintöjä:

  Otosavaruus S = Laatikollinen ruuveja

  Tapahtuma A = ”Ruuvin on valmistanut A-kone”

  Tapahtuma B = ”Ruuvin on valmistanut B-kone”

  Tapahtuma V = ”Ruuvi on viallinen”

- **Seuraavat todennäköisyydet *tunnetaan***:

  $\Pr(A) = 3/8$ $\Pr(V|A) = 0.05$

  $\Pr(B) = 5/8$ $\Pr(V|B) = 0.08$

- **Seuraavia todennäköisyyksiä *kysytään***:

  (i) $\Pr(V)$

  (ii) $\Pr(A|V)$

Esimerkki laadunvalvonnasta 3/4

- Tapahtumat A ja B muodostavat *otosavaruuden* S **osituksen**.
- Ositus $S = A \cup B$ **indusoi osituksen** tapahtumaan V

(i) **Kokonaistodennäköisyyden kaavasta** saadaan todennäköisyydeksi, että satunnaisesti poimittu ruuvi on viallinen:

$$
\begin{aligned}
\Pr(V) &= \Pr(A)\Pr(V|A) + \Pr(B)\Pr(V|B) \\
&= (3/8)\times 0.05 + (5/8)\times 0.08 \\
&= 0.06875 \\
&= 6.875\ \%
\end{aligned}
$$

Esimerkki laadunvalvonnasta 4/4

(ii) **Bayesin kaavasta** saadaan **ehdolliseksi todennäköisyydeksi** $\Pr(A|V)$ että ruuvin on valmistanut A-kone, *jos ruuvi osoittautuu vialliseksi* :

$$\Pr(A|V) = \frac{\Pr(A)\Pr(V|A)}{\Pr(A)\Pr(V|A) + \Pr(B)\Pr(V|B)}$$

$$= \frac{\frac{3}{8}\times 0.05}{\frac{3}{8}\times 0.05 + \frac{5}{8}\times 0.08} = \frac{3}{11} = 0.27$$