

Todennäköisyyslaskenta: Todennäköisyys ja sen laskusäännöt

1. Johdanto
2. Joukko-opin peruskäsitteet
3. Todennäköisyyslaskennan peruskäsitteet
4. Todennäköisyyslaskennan peruslaskusäännöt
5. Klassinen todennäköisyys ja kombinatoriikka
6. Todennäköisyyden aksioomat
7. Kokonaistodennäköisyyden ja Bayesin kaavat
8. Verkot ja todennäköisyyslaskenta

Sisällys

1. JOHDANTO	8
1.1. DETERMINISTISYYS JA SATUNNAISUUS	9
DETERMINISTISYYS	9
SATUNNAISUUS	9
1.2. SATUNNAISKOKEET JA KOETOISTOT	10
1.3. SATUNNAISUUS JA TILASTOLLINEN STABILITEETTI	10
TILASTOTIEDE, TILASTOLLINEN STABILITEETTI JA REILUT PELIT	11
1.4. TILASTOLLINEN TUTKIMUS PELINÄ LUONTOA VASTAAN	12
1.5. TODENNÄKÖISYYDEN MÄÄRITTELEMINEN	13
KLASSINEN TODENNÄKÖISYYS	14
ONGELMAT KLASSISEN TODENNÄKÖISYYDEN MÄÄRITELMÄSSÄ	14
KLASSINEN TODENNÄKÖISYYS JA TODENNÄKÖISYYDEN FREKVENSSITULKINTA	14
EMPIIRINEN TODENNÄKÖISYYS	15
ONGELMAT EMPIIRISEN TODENNÄKÖISYYDEN MÄÄRITELMÄSSÄ	16
EMPIIRINEN TODENNÄKÖISYYS JA TODENNÄKÖISYYDEN FREKVENSSITULKINTA	16
TODENNÄKÖISYYS MITTANA	18
ONGELMAT TODENNÄKÖISYYDEN NAIIVISSA MÄÄRITELMÄSSÄ MITTANA	18
KOLMOGOROVIN AKSIOOMAT TODENNÄKÖISYYDELLE	18
1.6. SATUNNAISILMIÖT JA NIIDEN TILASTOLLISET MALLIT	18
2. JOUKKO-OPIN PERUSKÄSITTEET	20
2.1. JOUKKO JA SEN ALKIOT	21
2.2. VENN-DIAGRAMMIT	22
2.3. OSAJOUKKO	22
2.4. TYHJÄ JOUKKO	23
2.5. JOUKKO-OPIN PERUSOPERAATIOIT	23
KOMPLEMENTTIJOUKKO	23
YHDISTE	23
LEIKKAUS	24
PISTEVIERAAT JOUKOT	24
EROTUS	25
3. TODENNÄKÖISYYSLASKENNAN PERUSKÄSITTEET	26
3.1. SATUNNAISILMIÖT	27
3.2. TODENNÄKÖISYYSLASKENNAN PERUSKÄSITTEET	27
3.3. TODENNÄKÖISYYS JA SEN PERUSOMINAISUUDET	29
TODENNÄKÖISYYDEN PERUSOMINAISUUDET	29
TODENNÄKÖISYYKSIEN VERTAILU	30
TODENNÄKÖISYYKSIEN VERTAILU JA TODENNÄKÖISYYDEN FREKVENSSITULKINTA	30
3.4. KLASSINEN TODENNÄKÖISYYS	30
SYMMETRISET ALKEISTAPAHTUMAT	30
ALKEISTAPAHTUMIEN SYMMETRIA JA UHKAPELIT	30
LUKUMÄÄRÄFUNKTIO	31
KLASSISEN TODENNÄKÖISYYDEN MÄÄRITELMÄ	31

3.5. EMPIIRINEN TODENNÄKÖISYYS	32
3.6. TODENNÄKÖISYYS MITTANA	33
3.7. TODENNÄKÖISYYDEN FREKVENSSITULKINTA	33

4. TODENNÄKÖISYYSLASKENNAN PERUSLASKUSÄÄNNÖT **34**

4.1. TODENNÄKÖISYYSLASKENNAN PERUSOPERAATIOI JA -LASKUSÄÄNNÖT	35
4.2. TODENNÄKÖISYYSLASKENNAN PERUSOPERAATIOI	35
VENN-DIAGRAMMIT	35
TAPAHTUMAN A KOMPLEMENTTI	36
TAPAHTUMIEN A JA B YHDISTE	36
TAPAHTUMIEN A JA B LEIKKAUS	37
TAPAHTUMIEN A JA B EROTUS	37
4.3. TODENNÄKÖISYYSLASKENNAN PERUSLASKUSÄÄNNÖT	37
KOMPLEMENTTITAPAHTUMAN TODENNÄKÖISYYS	39
TOISENSA POISSULKEVAT TAPAHTUMAT	39
YHTEENLASKUSÄÄNTÖ TOISENSA POISSULKEVILLE TAPAHTUMILLE	40
YLEISTETTY YHTEENLASKUSÄÄNTÖ TOISENSA POISSULKEVILLE TAPAHTUMILLE	40
TAPAHTUMIEN RIIPPUMATTOMUUS	40
RIIPPUMATTOMUUS VS RIIPPUVUUS	41
TULOSÄÄNTÖ RIIPPUMATTOMILLE TAPAHTUMILLE	41
YLEISTETTY TULOSÄÄNTÖ RIIPPUMATTOMILLE TAPAHTUMILLE	43
YLEINEN YHTEENLASKUSÄÄNTÖ	44
EROTUSTAPAHTUMAN TODENNÄKÖISYYS	45
TAPAHTUMASTA B SEURAA TAPAHTUMA A	45
EROTUSTAPAHTUMAN TODENNÄKÖISYYS, JOS TAPAHTUMASTA B SEURAA TAPAHTUMA A	46
YHDISTEEN TODENNÄKÖISYYS	46
4.4. EHDOLLINEN TODENNÄKÖISYYS JA RIIPPUMATTOMUUS	46
EHDOLLINEN TODENNÄKÖISYYS	46
RIIPPUMATTOMUUS	49
YLEINEN TULOSÄÄNTÖ	51
YKSINKERTAINEN SATUNNAISOTANTA JA TULOSÄÄNNÖT	52

5. KLASSINEN TODENNÄKÖISYYS JA KOMBINATORIIKKA **53**

5.1. KLASSINEN TODENNÄKÖISYYS	54
5.2. KOMBINATORIIKAN PERUSPERIAATTEET	54
5.3. KOMBINATORIIKAN PERUSONGELMAT	56
JOUKKO	56
JONO	56
5.4. KOMBINATORIIKAN PERUSONGELMIEN RATKAISEMINEN	57
PERMUTAATIO	57
N -KERTOMA	59
VARIAATIO	59
KOMBINAATIO	60
PERMUTAATIOI, VARIAATIOI, KOMBINAATIOI	61
PASCALIN KOLMIO	63
PASCALIN KOLMIO JA BINOMIKERTOIMET	63
BINOMIKAAVA	64
ÄÄRELLISEN JOUKON OSAJOUKKOJEN LUKUMÄÄRÄ	66
ESIMERKKEJÄ LUKUMÄÄRIEN LASKEMISESTA	67

5.5. MULTINOMIKERROIN	68
6. TODENNÄKÖISYYDEN AKSIOOMAT	70
6.1. TODENNÄKÖISYYS ÄÄRELLISISSÄ OTOSAVARUUKSISSA	71
BOOLEN ALGEBRAT	71
TODENNÄKÖISYYDEN AKSIOOMAT ÄÄRELLISISSÄ OTOSAVARUUKSISSA	74
ALKEISTODENNÄKÖISYYSLASKENNAN LASKUSÄÄNTÖJEN TODISTAMINEN	74
ÄÄRELLINEN TODENNÄKÖISYYSKENTTÄ	77
RIIPPUMATTOMUUS JA RIIPPUMATTOMIEN TAPAHTUMIEN TULOSÄÄNTÖ	77
6.2. KLASSINEN TODENNÄKÖISYYS, SUHTEELLINEN FREKVENSSI JA EHDOLLINEN TODENNÄKÖISYYS	77
TODENNÄKÖISYYKSINÄ	77
KLASSINEN TODENNÄKÖISYYS TODENNÄKÖISYYTENÄ	77
SUHTEELLINEN FREKVENSSI TODENNÄKÖISYYTENÄ	79
EMPIIRINEN TODENNÄKÖISYYS	80
TODENNÄKÖISYYDEN FREKVENSSITULKINTA	81
EHDOLLINEN TODENNÄKÖISYYS TODENNÄKÖISYYTENÄ	81
6.3. TODENNÄKÖISYYS MIELIVALTAISISSA OTOSAVARUUKSISSA	82
σ -ALGEBRAT	82
KOLMOGOROVIN AKSIOOMAT	84
TODENNÄKÖISYYSKENTTÄ	86
MITALLISET JA EPÄMITALLISET JOUKOT	86
KOLMOGOROVIN AKSIOOMIEN SEURAUKSIA	87
7. KOKONAISTODENNÄKÖISYYDEN JA BAYESIN KAAVAT	92
7.1. JOHDATTELEVA ESIMERKKI	93
7.2. OSITUKSET	95
7.3. KOKONAISTODENNÄKÖISYYDEN KAAVA	95
7.4. BAYESIN KAAVA	97
7.5. KOKONAISTODENNÄKÖISYYDEN KAAVAN JA BAYESIN KAAVAN SYSTEEMITEOREETTINEN TULKINTA	98
8. VERKOT JA TODENNÄKÖISYYSLASKENTA	101
8.1. JOHDATTELEVA ESIMERKKI	102
8.2. PUUDIAGRAMMIT JA TODENNÄKÖISYYSLASKENTA	103
VERKOT	103
PUUT	104
PUIDEN KONSTRUOINTI	104
PUUTODENNÄKÖISYYDET	105
TULOSÄÄNTÖ PUUTODENNÄKÖISYYKSILLE	105
YHTEENLASKUSÄÄNTÖ PUUTODENNÄKÖISYYKSILLE	106
PÄÄTÖSPUUT: ESIMERKKI	107
8.3. TOIMINTAVERKOT JA NIIDEN TOIMINTATODENNÄKÖISYYDET	109
TOIMINTAVERKOT	109
SARJAAN KYTKENNÄN TOIMINTATODENNÄKÖISYYS	110
RINNANKYTKENNÄN TOIMINTATODENNÄKÖISYYS	110
SOVELLUS	111

1. Johdanto

1.1. Deterministisyys ja satunnaisuus

1.2. Satunnaiskokeet ja koetoistot

1.3. Satunnaisuus ja tilastollinen stabiliteetti

1.4. Tilastollinen tutkimus pelinä luontoa vastaan

1.5. Todennäköisyyden määritteleminen

1.6. Satunnaisilmiöt ja niiden tilastolliset mallit

Todennäköisyyslaskenta kehittää *matemaattisia malleja satunnaisilmiöille*. Mallien avulla yritetään kuvata ja erottaa satunnaisilmiöiden systemaattiset ja satunnaiset piirteet.

Tarkastelemme tässä luvussa ensin sitä, mitkä tekijät tekevät reaalimaailman ilmiöistä **deterministisiä** ja mitkä **satunnaisia** sekä sitä, mitä satunnaisuudella tarkoitetaan. Samalla yritämme perustella myös sen, että satunnaisilmiön tuloksissa havaittava **satunnainen vaihtelu ei saa olla mielivaltaista**, vaan satunnaisvaihteluun on liitettävä sellaista *säännönmukaista käyttäytymistä*, jota kutsutaan **tilastolliseksi stabiliteetiksi**.

Toiseksi pyrimme määrittelemään **todennäköisyyden** käsitteen. Esitämme tässä luvussa kolme *naiivia* määritelmää todennäköisyydelle:

- Tapahtuman **klassinen todennäköisyys** on *tapahtumalle suotuisten tulostulovaihtoehtojen lukumäärän suhde kaikkien mahdollisten tulostulovaihtoehtojen lukumäärään*.
- Tapahtuman **empiirinen todennäköisyys** on *tapahtuman tilastollisesti stabiili suhteellinen osuus satunnaisilmiön esiintymiskertojen joukossa*.
- **Todennäköisyys** on *tapahtuman sattumisen mahdollisuuden mitta*.

Kaikki kolme tässä luvussa esitettävää määritelmää todennäköisyydelle ovat *naiiveja* siinä mielessä, että ne eivät ole *matemaattisesti tarkkoja*. Määritelmillä on kuitenkin ollut suuri merkitys *todennäköisyyslaskennan historiassa* ja niihin liittyy *tulkinnallisesti tärkeitä näkökulmia* todennäköisyyden käsitteeseen. Todennäköisyyden matemaattisesti tarkka määritelmä esitetään luvussa **Todennäköisyyden aksioomat**.

Avainsanat:

Deterministinen ilmiö, Deterministisyys, Empiirinen todennäköisyys, Epäreilu peli, Frekvenssi, Frekvenssitulkinta, Klassinen todennäköisyys, Koetoisto, Kolmogorovin aksioomat, Luonto, Mitta, Peli luontoa vastaan, Reaalimaailman ilmiö, Reilu peli, Satunnaisilmiö, Satunnaiskoe, Satunnaisuus, Stokastinen malli, Stokastisuus, Suhteellinen frekvenssi, Suhteellinen osuus, Suotuisa tulostulovaihtoehto, Tilastollinen malli, Tilastollinen stabiliteetti, Tilastollinen tutkimus, Tilastotiede, Todennäköisyyden aksioomat, Todennäköisyyden frekvenssitulkinta, Todennäköisyyden määritteleminen, Todennäköisyys, Todennäköisyyslaskenta, Todennäköisyysmalli, Tulos, Tulostulovaihtoehto, Venn-diagrammi

1.1. Deterministisyys ja satunnaisuus

Deterministisyys

Reaalimaailman ilmiö on **deterministinen**, jos *ilmiön alkutilan perusteella voidaan tarkasti ennustaa ilmiön lopputila eli tulos*. Deterministisen ilmiön *alkutila määrää ilmiön lopputilan*. Deterministisiä ilmiöitä kutsutaan usein **eksakteiksi** tai **kausaalisiksi**.

Monia *fysiikan*, kuten esimerkiksi *klassisen mekaniikan* ilmiöitä voidaan kuvata deterministisinä ilmiöinä.

Esimerkki 1. Ammuksen radan ennustaminen.

Ammuksen rata voidaan *ennustaa* hyvin tarkasti, jos tunnetaan ammuksen paino, lähtönopeus, lähtökulma, lähtösuunta, ilmanvastus jne.

Huomautus:

- Deterministisiä ilmiöitä koskeviin *havaintoihin* liittyy hyvin usein luonteeltaan *satunnaista havaintovirhettä*. Lisäksi deterministisiin ilmiöihin saattaa liittyä *ennustamattomuutta*, jota kutsutaan *kaaokseksi*.

Satunnaisuus

Reaalimaailman ilmiö on **stokastinen** eli **satunnainen**, jos sillä on seuraavat ominaisuudet:

- Ilmiö voi päätyä *alkutilastaan* useisiin erilaisiin *lopputiloihin* eli ilmiöllä on useita erilaisia *vaihtoehtoisia tuloksia*.
- Ilmiön alkutilan perusteella *ei voida* tarkasti *ennustaa* ilmiön lopputilaa eli sitä, mikä ilmiön mahdollisista *tulosvaihtoehdoista realisoituu* eli *toteutuu*.
- Vaikka ilmiön lopputilaa ei voida ennustaa tarkasti, *tulosvaihtoehtojen suhteellisten frekvenssien* eli *osuuksien* voidaan *nähdä* ilmiön toistuessa *käyttäytyvän säännönmukaisesti*.

Esimerkki 2. Erilaisia satunnaisilmiöitä.

Satunnaisilmiöitä tai ilmiöitä, joissa satunnaisuus on mukana ainakin osatekijänä kohdataan mitä erilaisimmilla alueilla.

- Sellaiset perustuvaa laatua olevat *biologiset ilmiöt* kuten sukupuolen määräytyminen tai ominaisuuksien periytyminen vanhemmilta jälkeläisille noudattavat sattuman lakeja.
- *Kvanttimekaniikan ilmiöt* kuten alkeishiukkasten käyttäytyminen tai radioaktiivinen hajoaminen ovat luonteeltaan satunnaisilmiöitä.
- Useimpia *sosiologian* tai *taloustieteen* tutkimia *yhteiskunnallisia ilmiöitä* ei voida ymmärtää, jollei satunnaisten tekijöiden roolia oteta huomioon.
- Tavallisimmissa *uhkapeleissä* kuten napanheitossa, korttipeleissä, lotossa, ruletissa tai arpajaisissa sattumalla on aivan keskeinen rooli.
- Empiirisen tutkimuksen mittauksiin välttämättä liittyvät *havaintovirheet* käyttäytyvät satunnaisesti.
- Tilastollisten tutkimusaineistojen kerääminen kuten *otoksien poiminta* perustuu arvontaan ja siten sattuman hyväksikäyttöön.

Satunnaisilmiöiden tulosta ei siis voida ennustaa tarkasti, mutta ilmiön toistuessa tulostavaihtoehtojen suhteellisten frekvenssien eli osuuksien havaitaan käyttäytyvän säännönmukaisesti.

Esimerkki 3. Satunnaisilmiöt ja niiden säännönmukaiset piirteet.

Esimerkkejä satunnaisilmiöiden säännönmukaisista piirteistä:

- Lapsen sukupuolta ei voida ennustaa etukäteen, mutta noin *puolet* syntyvistä lapsista on tyttöjä ja noin *puolet* on poikia.
- Satunnaisesti valitun ihmisen älykkyysosamäärää ei tiedetä, mutta *älykkyysosamäärät jakautuvat suurissa ihmisjoukoissa ns. normaalijakauman mukaan*. Käsitlemme normaalijakaumaa luvussa **Jatkuvia jakaumia**.
- Radioaktiivisen aineen yksittäisen atomin hajoamishetkeä ei voida ennustaa, mutta ns. *puoliintumisaika on jokaiselle radioaktiiviselle aineelle ominainen vakio*.
- Havaintovirheen suuruutta ja suuntaa ei voida ennustaa yksittäiselle havainnolle, mutta *suurissa havaintojoukoissa havaintovirheet jakautuvat hyvin usein ns. normaalijakauman mukaan*. Käsitlemme normaalijakaumaa luvussa **Jatkuvia jakaumia**.

1.2. Satunnaiskokeet ja koetoistot

Satunnaisilmiötä kutsutaan todennäköisyyslaskennassa usein **satunnaiskokeeksi** ja satunnais-ilmiön esiintymiskertaa kutsutaan usein **koetoistoksi**.

Esimerkki 1. Sukupuolen määräytyminen.

Voimme kutsua *sukupuolen määräytymismekanismeja* munasolun hedelmöityessä satunnaiskokeeksi ja tietyn yksilön sukupuolen määräytymistä koetoistoksi.

Esimerkki 2. Nopanheitto.

Nopanheitto on satunnaiskoe ja yksittäinen nopanheitto on koetoisto.

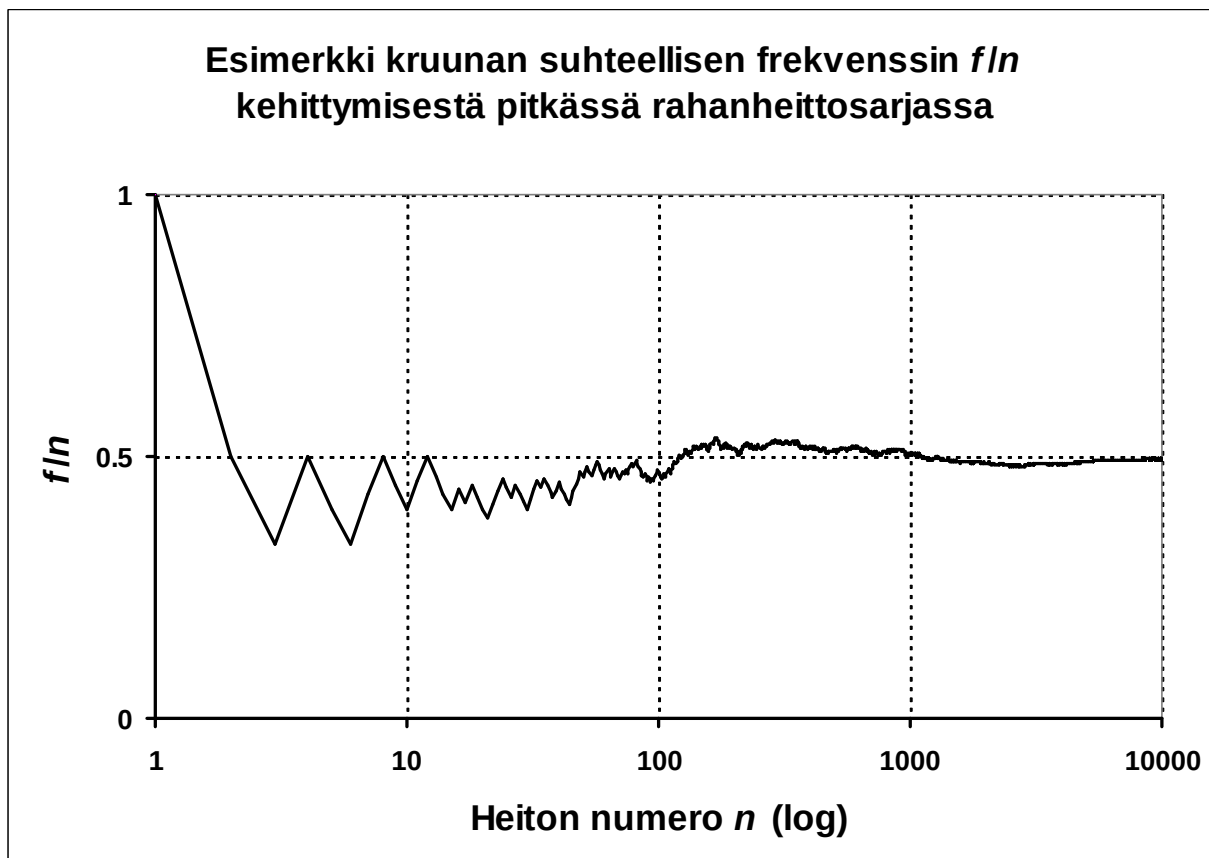
1.3. Satunnaisuus ja tilastollinen stabiliteetti

Todennäköisyyslaskennan *satunnaisuus ei ole mielivaltaisuutta*. Vaikka ilmiön tulos vaihtelee ilmiön toistuessa tavalla, jota ei voida ennustaa tarkasti, niin ilmiön säännönmukaisten piirteiden on tultava esille ilmiön toistuessa. **Todennäköisyyslaskennan** tehtävänä on kehittää *matemaattisia malleja* tällaisille säännönmukaisille piirteille.

Esimerkki 1. Rahanheitto satunnaisilmiönä.

Heitetään *virheetöntä* rahaa toistuvasti ja pidetään kirjaa kruunien *frekvenssistä* eli lukumäärästä f sekä *suhteellisesta frekvenssistä* eli osuudesta f/n , jossa n on tehtyjen heittojen lukumäärä. *Yksittäisen heiton tulosta* ei voida ennustaa etukäteen ja kruunien suhteellinen frekvenssi f/n *vaihtelee* heittoja toistettaessa, mutta *lähestyy* virheettömän rahan tapauksessa kuitenkin lukua $1/2$ niin, että *suuret poikkeamat luvusta $1/2$ tulevat yhä harvinaisemmiksi*.

Alla oleva kuva esittää kruunan suhteellisen frekvenssin f/n kehittymistä *eräässä* realisoituneessa heittosarjassa. Kuvasta nähdään, kuinka kruunan suhteellisen frekvenssin f/n suuret poikkeamat luvusta $1/2$ tulevat yhä harvinaisemmiksi, kun heittojen lukumäärä n kasvaa.



On syytä huomata, että luku $1/2$ ei ole kruunien suhteellisen frekvenssin f/n raja-arvo tavanomaisen lukujonokonvergenssin mielessä. Tapa, jolla kruunien suhteellinen frekvenssi lähestyy lukua $1/2$, on esimerkki ns. *stokastisesta konvergenssista*; ks. luvun **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet** kappaletta **Suurten lukujen lait**.

Kutsumme satunnaisilmiön toistuessa ilmenevää säännönmukaisuutta **tilastolliseksi stabiliteetiksi**. Jos satunnaisilmiö ei ole tilastollisesti stabiili, sitä ei voida mallintaa todennäköisyyslaskennan matemaattisilla malleilla.

Huomautus:

- Tilastollisen stabiliteetin idea saa matemaattisesti täsmällisen muotoilun ns. *suurten lukujen laeissa*; ks. luvun **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet** kappaletta **Suurten lukujen lait**.

Tilastotiede, tilastollinen stabiliteetti ja reilut pelit

Tilastotiede on yleinen menetelmätiede, jonka tehtävänä on kehittää ja soveltaa matemaattisia malleja, joiden avulla voidaan kuvata ja selittää mekanismeja, jotka tuottavat reaalimaailman ilmiöistä numeerisia tai kvantitatiivisia **tietoja**, sellaisissa tilanteissa, joissa tietoihin liittyy epävarmuutta tai satunnaisuutta. Tietoihin liittyvä epävarmuus tai satunnaisuus voi olla seurausta ilmiöstä itsestään tai siitä tavasta, jolla tiedot on kerätty.

Koska **tilastollisissa tutkimusasetelmissä** tutkimuksen kohteena olevaa reaalimaailman ilmiötä koskeviin tietoihin liittyy epävarmuutta tai satunnaisuutta, tilastotieteen malleja eli **tilastollisia malleja** rakennettaessa sovelletaan todennäköisyyslaskentaa.

Satunnaisilmiölle voidaan rakentaa tilastollisia malleja vain, jos *ilmiön tulokset* (tai niitä koskevat tiedot) *eivät vaihtelee mielivaltaisella tavalla*. *Ei-mielivaltaisuuksella* tarkoitetaan sitä, että ilmiön toistuksessa tulosvaihtoehtojen *suhteelliset frekvenssit* eli *osuudet* käyttäytyvät *tilastollisesti stabiilisti*.

Vaatusatunnaisilmiön tulosten käyttäytymisen tilastollisesta stabiiliteetista voidaan tulkita vaatimukseksi **reilusta pelistä**.

Esimerkki 2. Reilu vs epäreilu peli.

Pelaat Mr. Ebenezer Scroogea vastaan *peliiä*, jolla on seuraavat säännöt:

- (i) Mr. Scroogella on hallussaan useita *erilaisia* noppia, joiden silmälukuja *et tiedä*.
- (ii) Ensin Mr. Scrooge valitsee nopistaan *yhden*.
- (iii) Et saa ottaa Mr. Scroogen valitsemaa noppaa käteesi, mutta Mr. Scrooge heittää valitsemaansa noppaa *niin monta kertaa kuin haluat*.
- (iv) Jokaisen heiton jälkeen Mr. Scrooge näyttää sinulle heiton *tuloksen* eli sen silmäluvun, joka on nopan ylösjääneellä tahkolla.
- (v) Voitat ennalta sovitun rahasumman, jos saat selville Mr. Scroogen valitseman nopan silmäluvut.

Vaikka et saakaan ottaa noppaa käteesi, saat suurella varmuudella Mr. Scroogen valitseman nopan silmäluvut selville, jos *toistatat nopanheittoa* riittävän monta kertaa ja *tarkkailet* heittojen tuloksena esiintyvien silmälukujen *suhteellisia frekvenssejä*.

Oletetaan esimerkiksi, että Mr. Scroogen valitseman nopan silmäluvut ovat

1, 1, 1, 2, 2, 3

On ilmeistä, että silmälukujen 1, 2 ja 3 suhteellisten frekvenssien on *jakauduttava* pitkässä heittosarjassa suunnilleen suhteessa

3 : 2 : 1

Siten saat Mr. Scroogen valitseman nopan silmäluvut selville toistattamalla nopanheittoja riittävän monta kertaa.

Oletetaan nyt, että Mr. Scrooge *vaihtaa noppaansa koko ajan salaten vaihdot*. Tällöin *et voi voittaa pelissä*, koska Mr. Scrooge *rikkoo tietämättäsi pelin sääntöä* (ii) vastaan. *Vaatusatunnaisesta pelistä* tarkoittaa sitä, että tällaista sääntöjen rikkomista *ei saa tapahtua*.

1.4. Tilastollinen tutkimus pelinä luontoa vastaan

Tilastollista tutkimusta voidaan kuvata **pelinä luontoa vastaan**. Tilastollisessa tutkimuksessa pyritään tekemään *luonnon* (eli *reaalimaailman*) *tilaa koskevia johtopäätöksiä luonnon* (eli *reaalimaailman*) *tilasta kerättyjen havaintojen perusteella*.

Voimme ajatella, että luonnolla on kädessään joukko ”kortteja” ja tutkija pyrkii *ottamaan selville* luonnon kädessä olevat ”kortit”, kun taas luonto pyrkii *salaamaan* ne. Peli koostuu eristä, joissa tutkija voi katsoa yhden *satunnaisesti* luonnon kädestä valitsemansa ”kortin” – *tämä on havaintojen keräämistä*.

Tutkija voi saada selville luonnon tilan eli luonnon kädessä olevat ”kortit” pelaamalla *riittävän monta erää* eli keräämällä *riittävästi havaintoja* – mutta vain sillä edellytyksellä, että luonto pelaa *reilusti* eli ei muuta pelin aikana jatkuvasti kätensä sisältöä.

Tilastollisessa tutkimuksessa pyritään havaintojen perusteella päättämään, millainen on *havainnot tuottanut mekanismi*. Päättely ei onnistu, jos havainnot tuottanut mekanismi ei ole jossakin mielessä pysyvä eli **tilastollisesti stabiili**. Oletus havainnot tuottaneen mekanismin pysyvyydestä voidaan tulkita oletukseksi siitä, että luonto pelaa *reilusti* eli ei riko pelin sääntöjä vastaan vaihtamalla jatkuvasti kädessään olevia ”kortteja”. Olosuhteiden säilyminen vakiona takaa tavallisesti sen, että satunnaiskokeen tulokset käyttäytyvät tilastollisesti stabiilisti.

On syytä olla tietoinen siitä, että tilastotiede on kehittänyt myös sellaisia menetelmiä, joilla voidaan pyrkiä paljastamaan *muutokset* havainnot tuottaneessa mekanismissa. Tilastollisen tutkimuksen kohteena ovatkin usein seuraavat kysymykset:

- (i) *Onko* havainnot tuottaneessa mekanismissa *tapahtunut muutoksia*?
- (ii) *Mitkä ovat muutosten syyt*?

Jotta tilastotiede pystyisi *mallintamaan* havainnot tuottaneessa mekanismissa tapahtuneet muutokset, muutokset eivät kuitenkaan saa tapahtua *mielivaltaisella tavalla*, vaan niissä on oltava jokin *systemaattinen piirre*.

1.5. Todennäköisyyden määritteleminen

Todennäköisyyslaskennan tehtävänä on rakentaa *matemaattisia malleja* satunnaisilmiöiden *säännönmukaisille* piirteille.

Satunnaisilmiö koostuu **tulosvaihtoehdoista**, joista jokin sattuu, kun satunnaisilmiö esiintyy. Mikä tulosvaihtoehdoista sattuu, ei voida ennustaa tarkasti etukäteen, mutta ilmiön toistuessa ilmiön säännönmukaiset piirteet tulevat esiin tulosvaihtoehtojen suhteellisissa frekvensseissä eli osuuksissa.

Satunnaisilmiön tulosvaihtoehtojen yhdistelmiä eli tulosvaihtoehtojen muodostamia *joukkoja* kutsutaan todennäköisyyslaskennassa **tapahtumiksi**. Todennäköisyyslaskennan eräänä tehtävänä on määrätä satunnaisilmiön tapahtumien **todennäköisyydet**.

Mitä todennäköisyys on? Hyvän määritelmän antaminen todennäköisyydelle on osoittautunut vaikeaksi. Todennäköisyydelle on todennäköisyyslaskennan historiassa esitetty mm. seuraavat *naiivit määritelmät*:

- (i) Tapahtuman **klassinen todennäköisyys** on *ko. tapahtumalle suotuisien tulosvaihtoehtojen osuus kaikista mahdollisista tulosvaihtoehdoista*.
- (ii) Tapahtuman **empiirinen todennäköisyys** on *ko. tapahtuman* (tilastollisesti stabiilisti käyttäytyvä) *suhteellinen frekvenssi*.
- (iii) Tapahtuman **todennäköisyys** on *ko. tapahtuman sattumisen mahdollisuuden mitta*.

Em. määritelmiä on tapana kutsua **naiiveiksi**, koska *ne eivät määrittele todennäköisyyttä matematiikan kannalta tyydyttävällä tavalla*.

On osoittautunut, että todennäköisyyden täsmällinen määrittelemineen vaatii *aksiomaattista lähestymistapaa*. Todennäköisyyden aksiomat esitti venäläinen matemaatikko A. N. Kolmogorov 1930-luvun alussa. **Kolmogorovin aksiomien** mukaan *todennäköisyys on mitta matemaattisen mittateorian tarkoittamassa mielessä* ja siten todennäköisyydellä on samantapaiset matemaattiset ominaisuudet kuin *pituus-, pinta-ala- ja tilavuusmitoilla*; ks. lukua **Todennäköisyyden aksiomat**. Todennäköisyyden naiivit määritelmät voidaan sisällyttää (sopivasti muotoiltuina) Kolmogorovin aksiomien muodostamaan kehikkoon todennäköisyyden *erikoistapauksina* tai *tulkintoina*.

Tarkastelemme seuraavassa todennäköisyyden naiiveja määritelmiä lähemmin.

Klassinen todennäköisyys

Olkoon *satunnaisilmiöllä* n yhtä *todennäköistä tulosvaihtoehtoa* ja tarkastellaan ilmiön *tapahtumaa*, johon liittyy *tulosvaihtoehtoja* k kpl, joita sanotaan ko. tapahtumalle *suotuisiksi*. Tällöin ko. tapahtuman **klassinen todennäköisyys** p on ko. tapahtumalle suotuisien tulosvaihtoehtojen *osuus* satunnais-ilmiön *kaikista mahdollisista tulosvaihtoehtoja*:

$$p = \frac{k}{n}$$

Klassisen todennäköisyyden käsite sopii erityisesti *uhkapelien* analysointiin. Uhkapeleissä pelitapahtumien todennäköisyydet voidaan tavallisesti määrätä *päättelemällä* ne pelin säännöistä. *Historiallisesti* todennäköisyysslaskenta sai alkunsa 1600-luvulla eräisiin noppapeleihin liittyvien ongelmien ratkaisuyrityksistä.

On syytä huomata, että tulosvaihtoehtojen *lukumäärien laskeminen* on usein epätriviaali tehtävä ja apuna tarvitaan tavallisesti *kombinatoriikaksi* kutsuttua matematiikan osa-aluetta; ks. lukua **Klassinen todennäköisyys ja kombinatoriikka**.

Ongelmat klassisen todennäköisyyden määritelmässä

Klassisen todennäköisyyden määritelmä ei anna mahdollisuutta puhua sellaisten tapahtumien todennäköisyyksistä, joihin liittyvät tulosvaihtoehdot *eivät ole yhtä todennäköisiä*. Määritelmä ei myöskään anna mahdollisuutta puhua sellaisten tapahtumien todennäköisyyksistä, joihin liittyy *äärettömän monta tulosvaihtoehtoa*.

Klassinen todennäköisyys ja todennäköisyyden frekvenssitulkinta

Oletetaan, että *toistamme* jotakin satunnaiskoetta yhä uudelleen ja tarkkailemme jonkin *tapahtuman suhteellisen frekvenssin* käyttäytymistä koetoistojen aikana. **Todennäköisyyden frekvenssitulkinnan** mukaan *tapahtuman suhteellinen frekvenssi vaihtelee satunnaisesti koetoistosta toiseen, mutta saa keskimäärin ko. tapahtuman todennäköisyyttä lähellä olevia arvoja*. *Vahvistavatko havainnot tämän, on empiirinen kysymys*.

Esimerkki 1. Nopanheitto.

Heitetään *virheetöntä* noppaa. Tällöin tulosvaihtoehtoja on 6 kpl:

Silmäluvut 1, 2, 3, 4, 5, 6

Tarkastellaan tapahtumia

$A = \text{"Silmäluku on parillinen"}$

$B = \text{"Silmäluku} < 3\text{"}$

Tapahtumalle A *suotuisia* tulosvaihtoehtoja on 3 kpl: Silmäluvut 2, 4, 6

Tapahtumalle B *suotuisia* tulosvaihtoehtoja on 2 kpl: Silmäluvut 1, 2

Tapahtuman A klassinen todennäköisyys on

$$\frac{3}{6} = \frac{1}{2} = 0.5$$

Tapahtuman B klassinen todennäköisyys on

$$\frac{2}{6} = \frac{1}{3} \approx 0.333$$

Siten tapahtuma A on *todennäköisempi* kuin tapahtuma B .

Oletetaan, että heität noppaa toistuvasti. *Todennäköisyyden frekvenssitulkinnan* mukaan on odotettavissa, että tapahtuma A esiintyy *useammin kuin* tapahtuma B niin, että keskimäärin puolet heitoista antaa tulokseksi tapahtuman A ja keskimäärin $1/3$ heitoista antaa tulokseksi tapahtuman B .

Empiirinen todennäköisyys

Tarkastellaan *satunnaiskoetta*, jota voidaan *toistaa* niin, että seuraavat ehdot pätevät:

- (i) Kokeen olosuhteet *säilyvät muuttumattomina* koetoistosta toiseen.
- (ii) Koetoistot *ovat riippumattomia* siinä mielessä, että yhdenkään koetoiston tulos ei riipu siitä mitä tuloksia muissa koetoistoissa saadaan.

Tarkkaillaan jonkin *tapahtuman* esiintymistä koetoistojen aikana. Jos ko. tapahtuman *suhteellinen frekvenssi* eli *osuus koetoistoista* lähestyy jotakin *kiinteätä lukua*, kun koetoistojen lukumäärää kasvatetaan, kutsumme lukua ko. tapahtuman **empiiriseksi todennäköisyydeksi**.

Oletetaan siis, että satunnaiskoetta *toistetaan* n kertaa. Olkoon f ko. tapahtuman **frekvenssi** eli *lukumäärä* koetoistojen joukossa. Tällöin

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

on ko. tapahtuman **suhteellinen frekvenssi** eli *suhteellinen osuus* koetoistoista. Annetaan nyt koetoistojen lukumäärän n kasvaa rajatta. Jos suhteellinen frekvenssi f/n lähestyy tällöin (jossakin mielessä) lukua p eli

$$\frac{f}{n} \rightarrow p$$

niin sanomme, että luku p on ko. tapahtuman **empiirinen todennäköisyys**.

Huomautuksia:

- Suhteellisen frekvenssin f/n rajakäyttäytyminen koetoistojen lukumäärän n kasvaessa *ei ole* tavanomaista lukujonokonvergenssia. Todennäköisyysslaskennan konvergenssi-käsitteitä tarkastellaan luvussa **Stokastikan konvergenssikäsitteet ja raja-arvolauseet**.
- Voimme sanoa (huonolla suomenkielellä), että tapahtuman empiirinen todennäköisyys on ko. tapahtuman suhteellinen frekvenssi *"pitkässä juoksussa"*.
- Empiirisen todennäköisyyden määritelmä edellyttää tapahtuman suhteellisilta frekvensseiltä *tilastollista stabiliteettia*.
- *Matemaattista todennäköisyyden käsitettä* voidaan pitää empiirisen todennäköisyyden käsitteen *idealisointina*.

Ongelmat empiirisen todennäköisyyden määritelmässä

Empiirinen todennäköisyys on *empiirinen käsite* siinä mielessä, että tapahtuman suhteellisen frekvenssin f/n määrittäminen vaatii satunnaiskokeen *toistamista* ja *havaintojen keräämistä*.

Todellisuudessa tapahtuman empiiristä todennäköisyyttä *ei voida* – nimestään huolimatta – *määrittää kokeellisesti*, koska määritelmä vaatii satunnaiskokeen toistamista *äärettömän monta kertaa*. Empiirisen todennäköisyyden käsite ei myöskään anna mahdollisuutta puhua sellaisten tapahtumien todennäköisyyksistä, joista *ei ole havaintoja*. Lisäksi on syytä huomata, että empiirisen todennäköisyyden määritelmässä esiintyvä raja-arvo *ei ole hyvin määritelty*, koska mikään ei takaa, että määritelmässä esiintyvä raja-arvo *on olemassa*.

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

Empiirinen todennäköisyys ja todennäköisyyden frekvenssitulkinta

Oletetaan, että *toistamme* jotakin satunnaiskoetta yhä uudelleen ja tarkkailemme jonkin *tapahtuman suhteellisen frekvenssin* käyttäytymistä koetoistojen aikana. **Todennäköisyyden frekvenssitulkinnan** mukaan *tapahtuman suhteellinen frekvenssi vaihtelee satunnaisesti koetoistosta toiseen, mutta saa keskimäärin ko. tapahtuman todennäköisyyttä lähellä olevia arvoja*. Vahvistavtko havainnot tämän, on empiirinen kysymys.

Esimerkki 2. Tilastollinen laadunvalvonta.

Tehdas valmistaa erästä sähkölaitetta 300 kpl päivässä. Laitteille on asetettu niin ankarat laatukriteerit, että osa tehtaan valmistamista laitteista ei täytä niitä. Tehdään oletus, että vialliset laitteet syntyvät tuotannossa *täysin sattumanvaraisesti*.

Merkitään:

K = Laite on kelvollinen

V = Laite on viallinen

Oletetaan, että eräänä päivänä löydetään 6 viallista laitetta.

Satunnaisilmiö:	Laitteen laatu
Koetoisto:	Valmistetaan 1 laite
Koetoistojen lkm n :	300
Tulosvaihtoehto V:	Laite on viallinen
Viallisten laitteiden <i>frekvenssi</i> :	$f = 6$

Viallisten laitteiden *suhteellinen frekvenssi*:

$$\frac{f}{n} = \frac{6}{300} = 0.02$$

Oletetaan nyt, että vaikka viallisten laitteiden suhteellinen osuus vaihteleeikin päivästä toiseen, se pysyy kuitenkin *suunnilleen samana* eli, että viallisten laitteiden suhteellinen osuus käyttäytyy *tilastollisesti stabiilisti*.

Tällöin suhteellista frekvenssiä

$$\frac{f}{n} = \frac{6}{300} = 0.02$$

voidaan kutsua *todennäköisyydeksi* saada viallinen laite, jos tehtaan valmistamien laitteiden joukosta *poimitaan satunnaisesti* 1 laite tarkastettavaksi.

Esimerkki 3. Satunnaisotanta.

Väestötilaston mukaan Suomen väestö jakautui vuoden 1998 lopussa miehiin ja naisiin seuraavasti:

Miehet	2 516 000
Naiset	2 643 600
Yhteensä	5 159 600

Tyypillisessä *kyselytutkimuksessa* kyselyn kohteiksi poimitaan 300-2000 suomalaista kaikkien suomalaisten joukosta käyttäen (yksinkertaista) *satunnaisotantaa*; lisätietoja otannasta: ks. monistetta **Tilastolliset menetelmät**.

Satunnaisotoksen poimimista voidaan kuvata *arvontana*, jossa jokaista suomalaista vastaa hänen nimellään varustettu arpalippu, arvat on asetettu (hyvin sekoitettuna) suureen urnaan ja urnasta poimitaan *satunnaisesti* haluttu määrä arpalippuja, jolloin *otokseen* tulevat poimituksi ne suomalaiset, joiden nimet esiintyvät poimituissa lipuissa.

Todennäköisyys poimia tietty henkilö otokseen on

$$\frac{1}{5159600}$$

Miesten *suhteellinen osuus* eli *suhteellinen frekvenssi* kaikkien suomalaisten joukosta:

$$\frac{2516000}{5159600} = 0.488$$

Naisten *suhteellinen osuus* eli *suhteellinen frekvenssi* kaikkien suomalaisten joukosta:

$$\frac{2643600}{5159600} = 0.512$$

Koska suomalaisia on näinkin paljon, miesten ja naisten suhteelliset frekvenssit voidaan *tulkita* empiirisen todennäköisyyden määritelmän mukaisiksi todennäköisyyksiksi.

Siten *todennäköisyys*, että satunnaisesti suomalaisten joukosta poimittu henkilö on *mies*, on 0.488 ja *todennäköisyys*, että satunnaisesti suomalaisten joukosta poimittu henkilö on *nainen*, on 0.512. Todennäköisyys poimia suomalaisten joukosta nainen on *suurempi kuin* todennäköisyys poimia mies, koska naisia on suomalaisten joukossa *enemmän kuin* miehiä.

Oletetaan, että suomalaisten joukosta poimitaan *arpomalla* yhä uusia 1000 henkilön *satunnaisotoksia*. Tällöin otokseen poimitujen miesten ja naisten suhteelliset osuudet *vaihtelevat otoksesta toiseen*, mutta otokseen poimituista henkilöistä *keskimäärin* 48.8 % on miehiä ja *keskimäärin* 51.2 % on naisia. *Tilastollinen stabiliteetti* on sitä, että nämä suhdeluvut pysyvät otoksesta toiseen *suunnilleen samoina*.

Todennäköisyys mittana

Hyödyllisen *mielikuvan* todennäköisyyden luonteesta antaa seuraava *naiivi määritelmä*: **Todennäköisyys** on **mitta**, joka mittaa satunnaisilmiön tapahtumien *sattumisen mahdollisuutta*.

Ongelmat todennäköisyyden naiivissa määritelmässä mittana

Määritelmä *ei täytä hyvän määritelmän tunnusmerkkejä*, koska se on *kehämääritelmä*: Sattumisen mahdollisuus ja todennäköisyys *tarkoittavat* suunnilleen *samaa*.

Kolmogorovin aksioomat todennäköisyydelle

Mielikuva todennäköisyydestä satunnaisilmiöiden tapahtumien sattumisen mahdollisuuden mittana voidaan täsmentää matemaattisesti ns. **Kolmogorovin aksioomiksi**; ks. lukua **Todennäköisyyden aksioomat**.

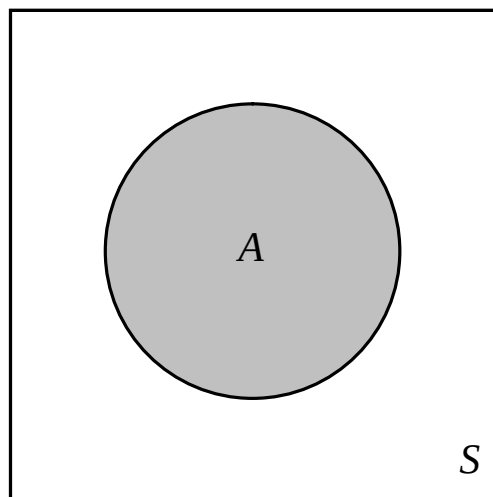
Kolmogorovin aksioomien mukaan todennäköisyys on *mitta matemaattisen mittateorian tarkoittamassa mielessä*. Aksioomien mukaan *todennäköisyysmitta* käyttäytyy samalla tavalla kuin esimerkiksi *pituus*-, *pinta-ala*- tai *tilavuusmitta*, lukuun ottamatta sitä, että todennäköisyysmitalla on ylärajana *varman tapahtuman todennäköisyys*, joksi sovitaan luku 1.

Todennäköisyyden tulkinnasta mittana seuraa se, että todennäköisyyden laskusääntöjä voidaan havainnollistaa joukko-opin operaatioiden havainnollistamisessa käytettävien **Venn-diagrammien** avulla.

Venn-diagrammi konstruoidaan seuraavalla tavalla:

- (i) Kuvataan tapahtumia *tasoalueilla*.
- (ii) Kuvataan tapahtumien todennäköisyyksiä *tasoalueiden pinta-aloilla*.

Viereisessä diagrammissa satunnaisilmiön kaikkien mahdollisten tulosvaihtoehtojen muodostamaa *otosavaruutta* S on kuvattu suorakaiteella ja ko. satunnaisilmiön *tapahtumaa* A on kuvattu varjostetulla ympyrällä.



1.6. Satunnaisilmiöt ja niiden tilastolliset mallit

Tilastotieteen tehtävänä on kehittää ja soveltaa **tilastollisiksi malleiksi** kutsuttuja *matemaattisia malleja* satunnaisilmiöille.

Mallien avulla pyritään tekemään ilmiöitä koskevia *johtopäätöksiä* ilmiöstä kerättyjen **havaintojen** perusteella. Malleja kutsutaan tilastotieteessä ja sen sovelluksissa **tilastollisiksi malleiksi** ja niiden avulla pyritään erottamaan ja kuvamaan satunnaisilmiöiden *systemaattiset* ja *satunnaiset piirteet*.

Koska tilastolliset mallit perustuvat **todennäköisyyslaskentaan**, niitä kutsutaan usein myös **stokastisiksi malleiksi** tai **todennäköisyysmalleiksi**.

Satunnaisilmiön **tilastollinen malli** koostuu seuraavista kahdesta osasta:

- (i) Satunnaisilmiön kaikkien mahdollisten **tulosvaihtoehtojen** kuvaus.
- (ii) Tulosvaihtoehtojen **todennäköisyyksien** kuvaus.

Satunnaisilmiön tilastollinen malli esitetään satunnaisilmiön tulosvaihtoja numeerisessa muodossa kuvaavaan *satunnaismuuttujan* ja sen *todennäköisyysjakauman* avulla; ks. lukua **Satunnaismuuttujat ja todennäköisyysjakaumat**.

Vaikka tilastollisten mallien rakentamista käsitellään vasta monisteessa **Tilastolliset menetelmät**, on hyvä painaa mieleensä jo nyt, että tilastollisen mallin rakentamiseen on aina syytä liittää seuraavat työvaiheet:

1. **Mallin muodostaminen** tutkimuksen kohteena olevalle satunnaisilmiölle.
2. Ilmiötä koskevien **havaintojen kerääminen**.
3. Mallin **parametrien estimointi** eli parametrien arvojen määrittäminen kerättyjen havaintojen perusteella.
4. Mallin ja kerättyjen havaintojen **yhteensopivuuden testaaminen**.

Jos mallissa havaitaan vaiheessa 4 puutteita, on mallin rakentamisessa palattava vaiheeseen 1.

2. Joukko-opin peruskäsitteet

2.1. Joukko ja sen alkiot

2.2. Venn-diagrammit

2.3. Osajoukko

2.4. Tyhjä joukko

2.5. Joukko-opin perusoperaatiot

Todennäköisyyslaskennan historian tärkeimpiä teoreettisia oivalluksia on ollut se, että satunnais-ilmiöiden tapahtumia voidaan käsitellä **joukkoina** ja todennäköisyydellä on matemaattisesti samantyyppiset ominaisuudet kuin pituus-, pinta-ala- tai tilavuusmitalla. Tästä seuraa se, että **joukko-oppi** muodostaa matemaattisen todennäköisyysteorian perustan.

Tässä luvussa esitellään *naiivin* **joukko-opin peruskäsitteet** ja **-määritelmät** siinä laajuudessa kuin niitä tarvitaan todennäköisyyslaskennassa; joukko-oppia tarkastellaan lähes koko perus-matematiikan vaatimassa laajuudessa liitteessä **Joukko-oppi**.

Avainsanat:

Alkio, Erotus, Joukko, Komplementti, Kuulua joukkoon, Leikkaus, Olla osajoukkona, Osajoukko, Perusjoukko, Pistevieraus, Tyhjä joukko, Unioni, Venn-diagrammi, Yhdiste

2.1. Joukko ja sen alkiot

Joukko on joukon **alkioiksi** kutsuttujen *olioiden kokoelma*. Sanomme, että joukko on **hyvin määritelty**, jos *jokaisesta oliosta voidaan sanoa onko se joukon alkio vai ei*. Joukko on siis hyvin määritelty, jos *sen alkiot tunnetaan*. Joukkoja määriteltäessä on tärkeää spesifioida **perusjoukko**, jonka alkioita kaikkien tarkasteltavien olioiden on oltava.

Merkitsemme joukkoja isoilla kirjaimilla

$$A, B, C, \dots, X, \dots$$

ja joukon alkioita pienillä kirjaimilla

$$a, b, c, \dots, x, \dots$$

Jos joukko on **äärellinen**, se voidaan määritellä luettelemalla sen alkiot. Jos äärellisen joukon X alkiot ovat

$$x_1, x_2, \dots, x_n$$

niin merkitsemme

$$X = \{x_1, x_2, \dots, x_n\}$$

Esimerkki 1. Nopan silmälukujen muodostama joukko.

Olkoon B tavallisen nopan silmälukujen muodostama joukko. Voimme tällöin kirjoittaa:

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

Joukon ja sen alkioiden suhde on joukko-opin perusrelaatio. Jos x on joukon A alkio eli x **kuuluu** joukkoon A , niin merkitsemme

$$x \in A$$

Jos x *ei ole* joukon A alkio eli x **ei kuulu** joukkoon A , niin merkitsemme

$$x \notin A$$

Esimerkki 2. Nopan silmälukujen muodostama joukko.

Olkoon B nopan silmälukujen muodostama joukko, jolloin

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

Tällöin

$$2 \in B$$

ja

$$7 \notin B$$

Matematiikassa joukko määritellään tavallisesti antamalla *looginen ehto*, jonka joukon alkioiden on toteutettava. Jos A on niiden perusjoukon S alkioiden x joukko, jotka *toteuttavat ehdon* $P(x)$ eli joille *lause* $P(x)$ *on tosi*, niin merkitsemme

$$A = \{x \in S \mid P(x)\}$$

Esimerkki 3. Parittomien positiivisten kokonaislukujen muodostama joukko.

Olkoon luonnollisten lukujen $0, 1, 2, 3, \dots$ joukko. Tällöin joukko

$$C = \{k \in \mathbb{N} \mid k = 2n+1, n \in \mathbb{N}\}$$

on parittomien positiivisten kokonaislukujen 1, 3, 5, ... muodostama joukko.

Esimerkiksi $7 \in C$, koska luku 7 voidaan esittää muodossa

$$7 = 2 \times 3 + 1$$

jossa $3 \in \mathbb{N}$. Sen sijaan esimerkiksi $8 \notin C$.

2.2. Venn-diagrammit

Joukko-opin käsitteitä ja operaatioita voidaan havainnollistaa ns. **Venn-diagrammien** avulla.

Venn-diagrammi konstruoidaan seuraavalla tavalla:

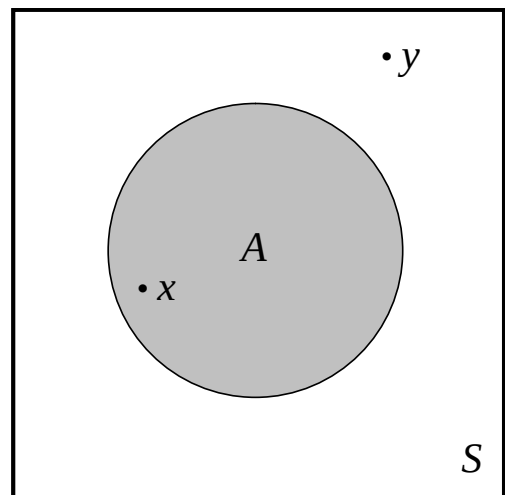
- (i) Kuvataan perusjoukkoa *suorakaiteella*.
- (ii) Kuvataan perusjoukossa määriteltyjä joukkoja *suorakaiteen osa-alueilla*.

Venn-diagrammi oikealla havainnollistaa sitä, että joukko A on perusjoukossa S määritelty joukko. Lisäksi kuva havainnollistaa sitä, että perusjoukon S alkio x on joukon A alkio:

$$x \in A$$

ja perusjoukon S alkio y ei ole joukon A alkio:

$$y \notin A$$



2.3. Osajoukko

Olkoot joukot A ja B perusjoukossa S määriteltyjä joukkoja.

Jos jokainen joukon B alkio on myös joukon A alkio eli

$$x \in B \Rightarrow x \in A$$

niin joukko B on joukon A **osajoukko** ja merkitsemme

$$B \subset A$$

Havainnollistus: ks. Venn-diagrammia oikealla.

Esimerkki 1. Parittomien positiivisten kokonaislukujen muodostama joukko.

Joukko

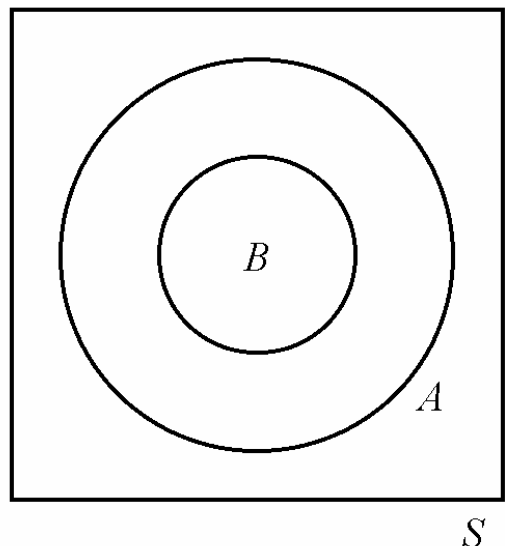
$$C = \{k \in \mathbb{N} \mid k = 2n+1, n \in \mathbb{N}\}$$

on parittomien positiivisten kokonaislukujen muodostama joukko. Olkoon

$$D = \{3, 5, 7, 9, 11\}$$

Tällöin

$$D \subset C$$



2.4. Tyhjä joukko

Joukko on **tyhjä**, jos siinä ei ole yhtään alkioita. Merkitsemme tyhjää joukkoa symbolilla

$$\emptyset$$

Tyhjä joukko on kaikkien joukkojen osajoukko. Jos siis A on perusjoukon S mielivaltainen osajoukko, niin

$$\emptyset \subset A$$

2.5. Joukko-opin perusoperaatiot

Olkoot joukot A ja B perusjoukon S osajoukkoja. Tarkastelemme seuraavien joukoista A ja B johdettujen perusjoukon S osajoukkojen määrittelemistä:

- (i) Joukon A **komplementtijoukko**.
- (ii) Joukkojen A ja B **unioni** eli **yhdiste**.
- (iii) Joukkojen A ja B **leikkaus**.
- (iv) Joukkojen A ja B **erotus**.

Komplementtijoukko

Olkoon joukko A perusjoukon S osajoukko.

Joukon A **komplementti** A^c on niiden perusjoukon S alkioiden joukko, jotka *eivät kuulu* joukkoon A :

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

Esimerkki 1.

Olkoon

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

ja

$$A = \{2, 4, 6\}$$

Tällöin

$$A^c = \{1, 3, 5\}$$

Yhdiste

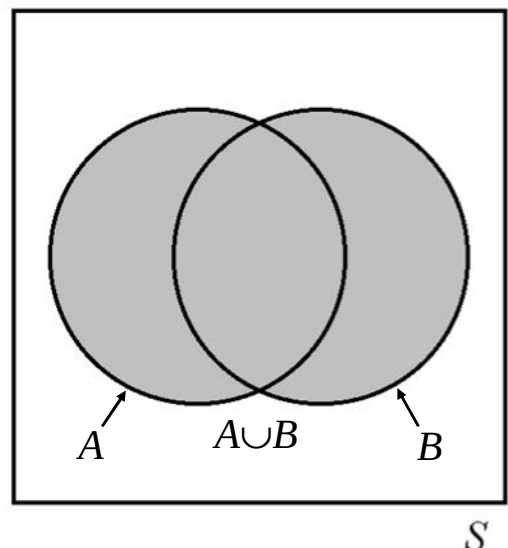
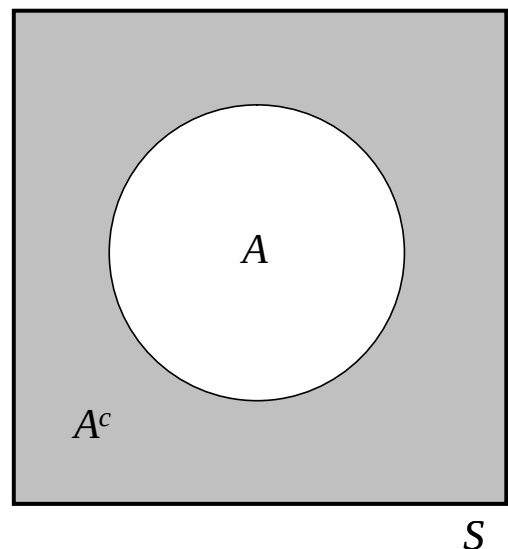
Olkoot joukot A ja B perusjoukon S osajoukkoja.

Joukkojen A ja B **unioni** eli **yhdiste** $A \cup B$ on niiden perusjoukon S alkioiden joukko, jotka kuuluvat joukkoon A tai joukkoon B :

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

Huomaa, että lause

$$"x \in A \text{ tai } x \in B"$$



tarkoittaa tässä sitä, että alkio x saa kuulua joukkoon A tai joukkoon B tai molempiin.

Esimerkki 2.

Olkoon

$$A = \{1, 2\}$$

ja

$$B = \{2, 3, 4\}$$

Tällöin

$$A \cup B = \{1, 2, 3, 4\}$$

Leikkaus

Olkoot joukot A ja B perusjoukon S osajoukkoja.

Joukkojen A ja B **leikkaus** $A \cap B$ on niiden perusjoukon S alkioiden joukko, jotka kuuluvat joukkoon A ja joukkoon B :

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

Esimerkki 3.

Olkoon

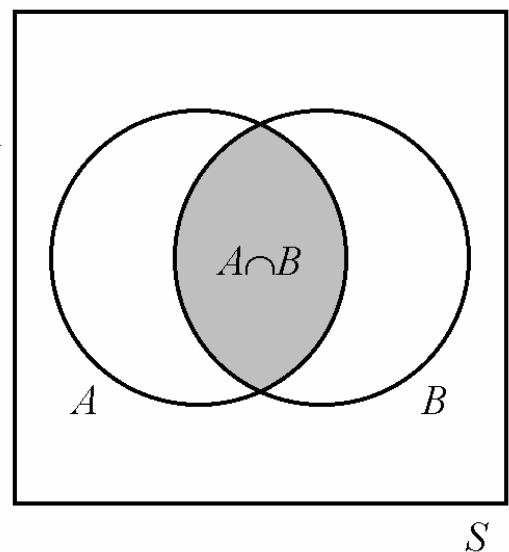
$$A = \{1, 2\}$$

ja

$$B = \{2, 3, 4\}$$

Tällöin

$$A \cap B = \{2\}$$



Pistevieraat joukot

Jos

$$A \cap B = \emptyset$$

niin sanomme, että joukot A ja B ovat **pistevieraita**. Siten pistevierailta joukoilla *ei ole* yhteisiä alkioita.

Esimerkki 4.

Olkoon

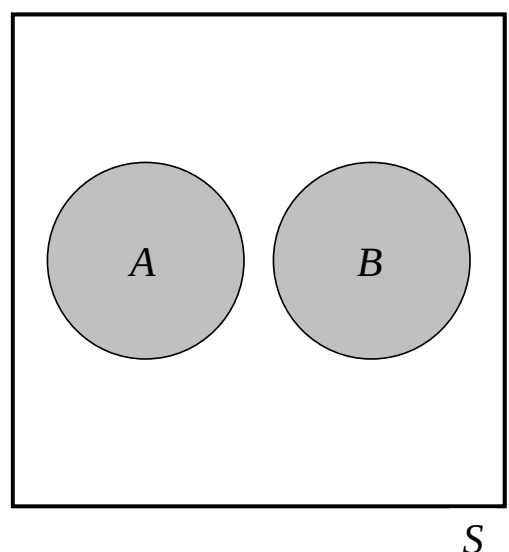
$$A = \{1, 2\}$$

ja

$$B = \{3, 4\}$$

Tällöin

$$A \cap B = \emptyset$$



Erotus

Olkoot joukot A ja B perusjoukon S osajoukkoja.

Joukkojen A ja B **erotus** $A \setminus B$ on niiden perusjoukon S alkuiden joukko, jotka kuuluvat joukkoon A , mutta eivät kuulu joukkoon B :

$$A \setminus B = \{x \mid x \in A \text{ ja } x \notin B\}$$

Selvästi

$$A \setminus B = A \cap B^c$$

Esimerkki 5.

Olkoon

$$A = \{1, 2\}$$

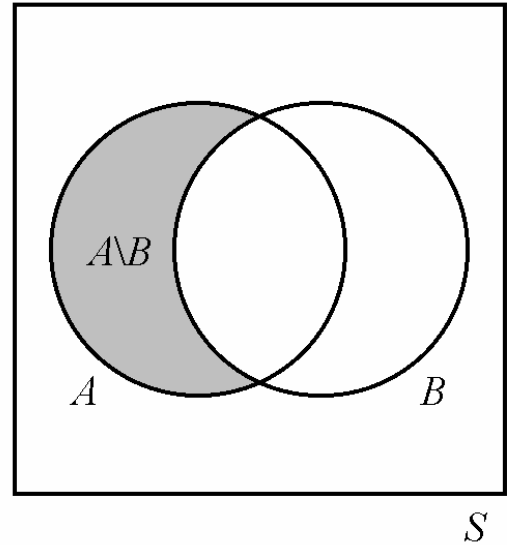
ja

$$B = \{2, 3, 4\}$$

Tällöin

$$A \setminus B = \{1\}$$

$$B \setminus A = \{3, 4\}$$



3. Todennäköisyyslaskennan peruskäsitteet

- 3.1. Satunnaisilmiöt
- 3.2. Todennäköisyyslaskennan peruskäsitteet
- 3.3. Todennäköisyys ja sen perusominaisuudet
- 3.4. Klassinen todennäköisyys
- 3.5. Empiirinen todennäköisyys
- 3.6. Todennäköisyys mittana
- 3.7. Todennäköisyyden frekvenssitulkinta

Tässä luvussa tarkastellaan **todennäköisyyslaskennan peruskäsitteitä**: **satunnaisilmiö**, **otosavaruus**, **tapahtuma** ja **alkeistapahtuma**. Lisäksi tarkastelemme **todennäköisyyden perusominaisuuksia** sekä liitämme ne luvun **Johdanto** kappaleessa **Todennäköisyyden määrittelemiseen** esitettyihin todennäköisyyden naiiveihin määritelmiin.

Avainsanat:

Alkeistapahtuma, Alkio, Empiirinen todennäköisyys, Frekvenssi, Frekvenssitulkinta, Joukko, Klassinen todennäköisyys, Koetoisto, Komplementtitapahtuma, Lukumääräfunktio, Mahdoton tapahtuma, Mitta, Osajoukko, Otosavaruus, Perusjoukko, Sattuma, Satunnaisilmiö, Satunnaiskoe, Suhteellinen frekvenssi, Suhteellinen osuus, Suotuisa alkeistapahtuma, Symmetrisyys, Tapahtuma, Todennäköisyys, Tulostavaihtoehto, Tyhjä joukko, Varma tapahtuma

3.1. Satunnaisilmiöt

Luvussa **Johdanto** esitimme satunnaisilmiölle seuraavan määritelmän: Reaalimaailman ilmiö on **stokastinen ilmiö** eli **satunnaisilmiö**, jos sillä on seuraavat ominaisuudet:

- (i) Ilmiö voi päätyä *alkutilastaan* useisiin erilaisiin *lopputiloihin* eli ilmiöllä on useita erilaisia *vaihtoehtoisia tuloksia*.
- (ii) Ilmiön alkutilan perusteella *ei voida ennustaa* tarkasti ilmiön lopputilaa eli sitä, mikä ilmiön mahdollisista *tulosvaihtoehdoista realisoituu* eli *toteutuu*.
- (iii) Vaikka ilmiön lopputilaa ei voida ennustaa tarkasti, tulosvaihtoehtojen *suhteellisten frekvenssien* eli *osuuksien voidaan nähdä* ilmiön toistuessa *käyttäytyvän säännönmukaisesti*.

Kutsumme *satunnaisilmiöitä* **satunnaiskokeiksi** ja *satunnaisilmiön esiintymisiä* **koetoistoiksi**.

3.2. Todennäköisyyslaskennan peruskäsitteet

Todennäköisyyslaskennan *peruskäsitteet* ovat *alkeistapahtuma*, *otosavaruus* ja *tapahtuma*:

- (i) Sanomme, että satunnaisilmiön tulosvaihtoehto on **alkeistapahtuma**, jos satunnaisilmiötä ei voida ”purkaa” sitä alkeellisempiin tulosvaihtoehtoihin
- (ii) Kutsumme satunnaisilmiön tai satunnaiskokeen *kaikkien* alkeistapahtumien muodostamaa joukkoa **otosavaruudeksi**.
- (iii) Satunnaisilmiön **tapahtumalla** tarkoitetaan jotakin otosavaruuden alkioden muodostamaa joukkoa.

Merkitsemme otosavaruutta tavallisesti isolla kirjaimella S (otosavaruus = *engl. sample space*) ja sen alkioita pienellä kirjaimella s . Jos alkeistapahtuma eli alkio s *kuuluu* otosavaruuteen S , niin merkitsemme

$$s \in S$$

Jos alkeistapahtuma s *ei kuulu* otosavaruuteen S merkitsemme

$$s \notin S$$

Otosavaruus S muodostaa siis sen *perusjoukon*, jossa satunnaisilmiön tulosvaihtoehtoja tarkastellaan. Koska satunnaisilmiötä ei voida purkaa alkeistapahtumia alkeellisempiin tulos-vaihtoehtoihin, niin *alkeistapahtumat* ovat otosavaruuden *alkioita*. Tarkasteltavan satunnaisilmiön *tapahtumat* ovat otosavaruuden S alkioden eli alkeistapahtumien muodostamia otosavaruuden S *osajoukkoja*.

Siten voimme asettaa *todennäköisyyslaskennan* ja *joukko-opin peruskäsitteet* vastaamaan seuraavalla tavalla toisiaan:

Otosavaruus	$\leftrightarrow$	Perusjoukko
Tapahtuma	$\leftrightarrow$	Perusjoukon osajoukko
Mahdoton tapahtuma	$\leftrightarrow$	Tyhjä joukko
Alkeistapahtuma	$\leftrightarrow$	Alkio

Kun sanomme, että jokin *tapahtuma sattuu*, tarkoitamme tällä aina sitä, että *jokin tapahtumaan liittyvistä alkeistapahtumista sattuu*.

Esimerkki 1. Sukupuolen määräytyminen satunnaisilmiönä.

Satunnaisilmiö: Lapsen sukupuolen määräytyminen

Otosavaruus:

$$S = \{\text{Tyttö, Poika}\}$$

Alkeistapahtumat:

Tyttö, Poika

Esimerkki 2. Nopanheitto satunnaisilmiönä.

Satunnaisilmiö: Nopanheitto

Otosavaruus: Nopan silmälukujen joukko

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

Alkeistapahtumat: Silmäluvut

1, 2, 3, 4, 5, 6

Esimerkki tapahtumasta:

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

Esimerkki 3. Kahden nopan heitto satunnaisilmiönä.

Satunnaisilmiö: Kahden nopan heitto

Otosavaruus S : Silmälukuparien (i, j) (36 kpl) joukko, jossa

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

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

Siten

$$S = \{(i, j) \mid i = 1, 2, 3, 4, 5, 6; j = 1, 2, 3, 4, 5, 6\}$$

Otosavaruuden S alkiot on kätevintä esittää taulukkomuodossa seuraavalla tavalla:

(i, j)	$j = \text{tulos 2. nopan heitosta}$					
$i = \text{tulos 1. nopan-heitosta}$	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)

Olkoon

$$A = \text{"Kummallakin nopanheitolla saadaan sama silmäluku"}$$

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

$$= \{(i, j) \in S \mid i = j, j = 1, 2, 3, 4, 5, 6\}$$

Tällöin

$$A \subset S$$

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

$$(6, 5) \notin A$$

Sanomme, että tapahtuma on **varma**, jos se *esiintyy aina*, kun satunnaisilmiö toistuu. *Otosavaruus* S itse on selvästi varma tapahtuma. Sanomme, että tapahtuma on **mahdoton**, jos se *ei voi esiintyä koskaan*, kun satunnaisilmiö toistuu. *Tyhjä joukko* $\emptyset$ on selvästi mahdoton tapahtuma.

Esimerkki 4. Rahanheitto.

Satunnaisilmiö: Rahanheitto.

Otosavaruus:

$$S = \{\text{Kruuna, Klaava}\}$$

Koska rahaa heitettäessä tuloksena on aina kruuna tai klaava, niin S on varma tapahtuma.

Esimerkki 5. Nopanheitto.

Satunnaisilmiö: Nopan heitto.

Otosavaruus: Silmälukujen joukko

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

Koska tavallista noppaa heitettäessä tuloksena ei voi olla silmäluku 7, niin $\{7\}$ on noppaa heitettäessä mahdoton tapahtuma.

3.3. Todennäköisyys ja sen perusominaisuudet

Olkoon A jokin otosavaruuden S tapahtuma eli olkoon

$$A \subset S$$

Todennäköisyys

$$\Pr(\cdot)$$

on *joukkofunktio*, joka liittää jokaiseen otosavaruuden S tapahtumaan A reaaliluvun:

$$\Pr(A) \in$$

Todennäköisyyden perusominaisuudet

(i) Olkoon A jokin otosavaruuden S tapahtuma. Tällöin

$$0 \leq \Pr(A) \leq 1$$

(ii) Samaistetaan tyhjä joukko $\emptyset$ ja **mahdoton tapahtuma** ja sovitaan, että

$$\Pr(\emptyset) = 0$$

(iii) Samaistetaan otosavaruus S ja **varma tapahtuma** ja sovitaan, että

$$\Pr(S) = 1$$

Todennäköisyyksien vertailu

Jos

$$\Pr(A) > \Pr(B)$$

niin sanomme

”Tapahtuma A on **todennäköisempi** kuin tapahtuma B ”

tai

”Tapahtuma B on **epätodennäköisempi** kuin tapahtuma A ”

Todennäköisyyksien vertailu ja todennäköisyyden frekvenssitulkinta

Mitä *todennäköisempi* tapahtuma on, sitä *useammin* tapahtumalla on taipumus esiintyä satunnais-
ilmiön toistuesssa eli sitä *suurempi* on (yleensä) tapahtuman *havaittu* suhteellinen frekvenssi ja mitä
epätodennäköisempi tapahtuma on, sitä *harvemmin* tapahtumalla on taipumus esiintyä satunnais-
ilmiön toistuesssa eli sitä *pienempi* on (yleensä) tapahtuman *havaittu* suhteellinen frekvenssi.

3.4. Klassinen todennäköisyys

Symmetriset alkeistapahtumat

Oletetaan, että äärellisen otosavaruuden

$$S = \{s_1, s_2, \dots, s_n\}$$

alkeistapahtumat

$$s_i, i = 1, 2, \dots, n$$

ovat *yhtä todennäköisiä* eli

$$\Pr(s_i) = \frac{1}{n}, i = 1, 2, \dots, n$$

Tällöin sanomme, että alkeistapahtumat $s_i, i = 1, 2, \dots, n$ ovat **symmetrisiä**. *Klassisen todennäköisyyden määritelmä* edellyttää sitä, että otosavaruus on äärellinen ja sen alkeistapahtumat ovat symmetrisiä.

Alkeistapahtumien symmetria ja uhkapelit

Useimmissa uhkapeleissä *pelin säännöt edellyttävät, että peliin liittyvät alkeistapahtumat ovat symmetrisiä*.

Tyypillisiä uhkapelien säännöissä esitettyjä *symmetriavaatimuksia* ovat seuraavat:

- (i) Käytettävien pelivälineiden (esim. nopan, rahan tai rulettypöyrän) on oltava *fysikaalisesti symmetrisiä*.
- (ii) Käytettävillä pelivälineillä (esim. arpalipuilla tai korteilla) on oltava *sama todennäköisyys tulla valituiksi tai jaetuiksi*. Huomaa, että tämä vaatimus edellyttää pelivälineiden (esim. arpalippujen tai korttien) huolellista *sekoittamista*.

Otosavaruuden alkeistapahtumien symmetrisyyttä voidaan vain harvoin *perustella* uhkapelien ulkopuolella. Oletus alkeistapahtumien symmetrisyydestä on *oletus*, jota voidaan *testata* tilastollisesti, jos ko. satunnaisilmiöstä kerätään *havaintoja*.

Lukumääräfunktio

Olkoon

$$n_A = n(A)$$

funktio, joka kertoo *joukon A alkioden lukumäärän*. Jos siis

$$A = \{a_1, a_2, \dots, a_k\}$$

on äärellinen joukko, jonka alkioden lukumäärä on k , niin

$$n_A = n(A) = k$$

Kutsumme funktiota $n(\cdot)$ **lukumääräfunktioksi**.

Klassisen todennäköisyyden määritelmä

Olkoon tapahtuma A otosavaruuden S osajoukko. Tällöin tapahtuman A **klassinen todennäköisyys** $\Pr_c(A)$ saadaan määrittämällä *tapahtumalle A suotuisien alkeistapahtumien osuus kaikista mahdollisista alkeistapahtumista* eli

$$\Pr_c(A) = \frac{n_A}{n_S} = \frac{n(A)}{n(S)}$$

jossa

$$\begin{aligned} n_A = n(A) &= \text{tapahtumalle } A \text{ suotuisien alkeistapahtumien lukumäärä} \\ &= \text{joukkoon } A \text{ kuuluvien alkeistapahtumien lukumäärä} \end{aligned}$$

ja

$$\begin{aligned} n_S = n(S) &= \text{kaikkien mahdollisten alkeistapahtumien lukumäärä} \\ &= \text{otosavaruuteen } S \text{ kuuluvien alkeistapahtumien lukumäärä} \end{aligned}$$

On helppo nähdä, että klassisella todennäköisyydellä on edellä esitetyt todennäköisyyden perusominaisuudet.

Perustelu:

Oletetaan, että

$$n(S) = n$$

(i) Olkoon A jokin otosavaruuden S tapahtuma ja

$$n(A) = k$$

Koska

$$A \subset S$$

niin

$$n(A) = k \leq n = n(S)$$

Siten

$$0 \leq \Pr_c(A) = \frac{n(A)}{n(S)} = \frac{k}{n} \leq 1$$

(ii) Koska

$$n(\emptyset) = 0$$

niin

$$\Pr_c(\emptyset) = \frac{n(\emptyset)}{n(S)} = \frac{0}{n} = 0$$

(iii) Koska

$$n(S) = n$$

niin

$$\Pr_c(S) = \frac{n(S)}{n(S)} = \frac{n}{n} = 1$$

■

3.5. Empiirinen todennäköisyys

Tarkastellaan *satunnaiskoetta*, jota voidaan *toistaa* siten, että seuraavat ehdot pätevät:

- (i) Kokeen olosuhteet *säilyvät muuttumattomina* koetoistosta toiseen.
- (ii) Koetoistot *ovat riippumattomia* siinä mielessä, että yhdenkään koetoiston tulos ei riipu siitä mitä tuloksia muissa koetoistoissa saadaan.

Tarkkaillaan *tapahtuman A* esiintymistä koetoistojen aikana. Jos tapahtuman *A* *suhteellinen frekvenssi* eli *osuus koetoistoista* lähestyy jotakin kiinteätä lukua koetoistojen lukumäärän kasvaessa rajatta, lukua kutsutaan tapahtuman *A* **empiiriseksi todennäköisyydeksi**.

Oletetaan, että satunnaiskoetta *toistetaan* n kertaa. Olkoon f_A tapahtuman *A* **frekvenssi** eli *lukumäärä* koetoistojen joukossa. Tällöin

$$\frac{f_A}{n}$$

on tapahtuman *A* **suhteellinen frekvenssi** eli *suhteellinen osuus* koetoistojen joukossa.

On helppo nähdä, että tapahtuman *A* suhteellisella frekvenssillä koetoistojen joukossa on edellä esitetyt todennäköisyyden perusominaisuudet.

Perustelu:

Oletetaan, että satunnaiskoetta toistetaan n kertaa.

- (i) Olkoon *A* jokin *otosavaruuden S tapahtuma*, joka esiintyy f_A kertaa toistojen aikana.

Koska

$$0 \leq f_A \leq n$$

niin

$$0 \leq \frac{f_A}{n} \leq 1$$

- (ii) Jos tapahtuma A ei satu yhtään kertaa, niin

$$f_A = 0$$

ja

$$\frac{f_A}{n} = 0$$

- (iii) Jos tapahtuma A sattuu joka kerran, niin

$$f_A = n$$

ja

$$\frac{f_A}{n} = 1$$

■

Jos koetoistojen lukumäärän n annetaan kasvaa rajatta ja tällöin (jossakin mielessä)

$$\frac{f_A}{n} \rightarrow p$$

niin luku p on tapahtuman A **empiirinen todennäköisyys**.

3.6. Todennäköisyys mittana

Todennäköisyys on **mitta**, joka mittaa satunnaisilmiön tapahtumien *sattumisen mahdollisuutta*. On syytä huomata, että **Kolmogorovin aksioomien** mukaan *todennäköisyys on mitta matemaattisen mittateorian mielessä*; ks. lukua **Todennäköisyyden aksioomat**

3.7. Todennäköisyyden frekvenssitulkinta

Oletetaan, että *toistamme* jotakin satunnaiskoetta ja tarkkailemme jonkin *tapahtuman suhteellisen frekvenssin* käyttäytymistä koetoistojen aikana. **Todennäköisyyden frekvenssitulkinnan** mukaan *ko. tapahtuman suhteellinen frekvenssi vaihtelee satunnaisesti koetoistosta toiseen, mutta saa keskimäärin ko. tapahtuman todennäköisyyttä lähellä olevia arvoja. Vahvistavatko havainnot tämän on empiirinen kysymys.*

Olkoon tapahtuman A todennäköisyys

$$\Pr(A) = p$$

Oletetaan, että sitä satunnaiskoetta, jonka tulosvaihtoehtona tapahtuma A on, toistetaan n kertaa. Tällöin todennäköisyyden frekvenssitulkinnasta seuraa, että *on odotettavissa, että tapahtuman A frekvenssi f_A on lähellä lukua*

$$np$$

4. Todennäköisyyslaskennan peruslaskusäännöt

- 4.1. Todennäköisyyslaskennan perusoperaatiot ja -laskusäännöt
- 4.2. Todennäköisyyslaskennan perusoperaatiot
- 4.3. Todennäköisyyslaskennan peruslaskusäännöt
- 4.4. Ehdollinen todennäköisyys ja riippumattomuus

Tarkastelemme tässä luvussa todennäköisyyslaskennan **perusoperaatioita** ja **-laskusääntöjä**.

Teemme operaatiot ja laskusäännöt ilmeisiksi **Venn-diagrammien** ja esimerkkien avulla.

Aksiomaattista lähestymistapaa todennäköisyyslaskennan laskusääntöjen todistamisessa tarkastellaan luvussa **Toden-näköisyyden aksioomat**.

Avainsanat:

Alkeistapahtuma, Ehdollinen todennäköisyys, Ehtotapahtuma, Erotus, Joukko, Komplementti, Leikkaus, Mahdoton tapahtuma, Otanta, Otanta palauttaen, Otanta palauttamatta, Orosavaruus, Pistevieraus, Riippumattomuus, Satunnaisotanta, Tapahtuma, Todennäköisyys, Toisensa poissulkevuus, Tulosääntö, Unioni, Varma tapahtuma, Venn-diagrammi, Yhdiste, Yhdistetty tapahtuma, Yhteenlaskusääntö, Yleinen tulosääntä, Yleinen yhteenlaskusääntö

4.1. Todennäköisyyslaskennan perusoperaatiot ja -laskusäännöt

Todennäköisyyslaskennan **perusoperaatioilla** tarkoitetaan *joukko-opin operaatioita*, joilla jonkin satunnaisilmiön tapahtumista johdetaan uusia tapahtumia.

Todennäköisyyslaskennan **peruslaskusäännöillä** tarkoitetaan *laskusääntöjä*, joilla uusien, todennäköisyyslaskennan perusoperaatiolla johdettujen tapahtumien todennäköisyydet määrätään niiden tapahtumien todennäköisyyksien avulla, joista uudet tapahtumat on johdettu.

Olkoon S **otosavaruus** eli tarkasteltavan satunnaisilmiön alkeistapahtumien muodostama joukko. Jos A on tarkasteltavan satunnaisilmiön **tapahtuma**, niin se on alkio, jota ovat otosavaruuden S alkeistapahtumia. Siten tapahtumat ovat tarkasteltavaan satunnaisilmiöön liittyvän otosavaruuden S *osajoukkoja*.

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 sitä, että jokin tapahtumaan A kuuluvista alkeistapahtumista s sattuu.

Olko A ja B otosavaruuden S *tapahtumia*. Jokaista operaatiota, jolla *tapahtumista* A ja B *johdetaan uusia tapahtumia*, vastaa jokin **joukko-opin operaatio**. Todennäköisyyslaskennan ja joukko-opin *perusoperaatiot* vastaavat seuraavalla tavalla toisiaan:

- (i) Tapahtuma A *ei* satu eli tapahtuman A *komplementtitapahtuma* $ei-A$ sattuu

$$\Leftrightarrow \text{Joukon } A \text{ komplementti } A^c$$

- (ii) Tapahtuma A sattuu *tai* tapahtuma B sattuu *tai* *molemmat* sattuvat

$$\Leftrightarrow \text{Joukkojen } A \text{ ja } B \text{ unioni eli yhdiste } A \cup B$$

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

$$\Leftrightarrow \text{Joukkojen } A \text{ ja } B \text{ leikkaus } A \cap B$$

- (iv) Tapahtuma A sattuu *ja* tapahtuma B *ei* satu

$$\Leftrightarrow \text{Joukkojen } A \text{ ja } B \text{ erotus } A \setminus B$$

4.2. Todennäköisyyslaskennan perusoperaatiot

Venn-diagrammit

Joukko-opin operaatioita ja laskusääntöjä voidaan havainnollistaa **Venn-diagrammien** avulla:

- (i) Kuvataan *otosavaruutta* S *suorakaiteella*, jonka pinta-ala vastaa koko otosavaruuden todennäköisyyttä

$$\Pr(S) = 1$$

- (ii) Kuvataan *tapahtumaa* $A \subset S$ *suorakaiteen osa-alueella*, jonka pinta-ala vastaa tapahtuman A todennäköisyyttä

$$\Pr(A)$$

Kuva oikealla havainnollistaa Venn-diagrammia, jossa $A \subset S$.

Venn-diagrammien käyttö todennäköisyyslaskennan operaatioiden ja laskusääntöjen *havainnollituksissa* perustuu mielikuvaan todennäköisyydestä tapahtumien sattumisen mahdollisuuden *mittana*.

Todennäköisyys käyttäytyy mittana samalla tavalla kuin *pinta-alamitta* paitsi, että todennäköisyysmitalla on ylärajana varman tapahtuman todennäköisyys 1.

On syytä huomata, että todennäköisyyslaskennan operaatioiden ja laskusääntöjen *havainnollitus Venn-diagrammin avulla ei ole* sama asia kuin säännön *todistus*.

Todennäköisyyslaskennan laskusääntöjen todistamista *Kolmogorovin aksioomajärjestelmässä* tarkastellaan luvussa **Todennäköisyyden aksioomat**.

Huomautus:

- Todennäköisyyslaskennan operaatioita ja laskusääntöjä voidaan havainnollistaa myös ns. *puudiagrammeilla*; ks. liitettä **Todennäköisyyslaskenta ja puudiagrammit**.

Tapahtuman A komplementti

Olkoon

$$A \subset S$$

otosavaruuden S tapahtuma.

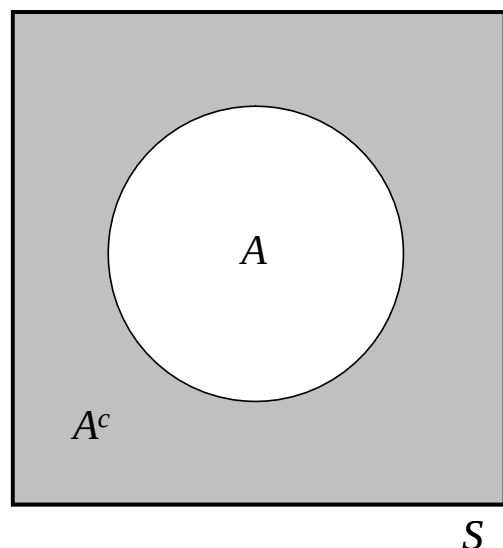
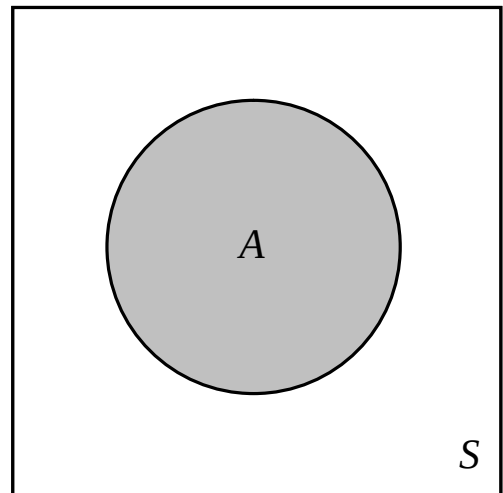
Tapahtuman A **komplementtitapahtuma**

$$A^c = \text{"}A \text{ ei satu"} = \text{ei-}A$$

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

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

Ks. kuvaa oikealla.



Tapahtumien A ja B yhdiste

Olkoot

$$A \subset S, B \subset S$$

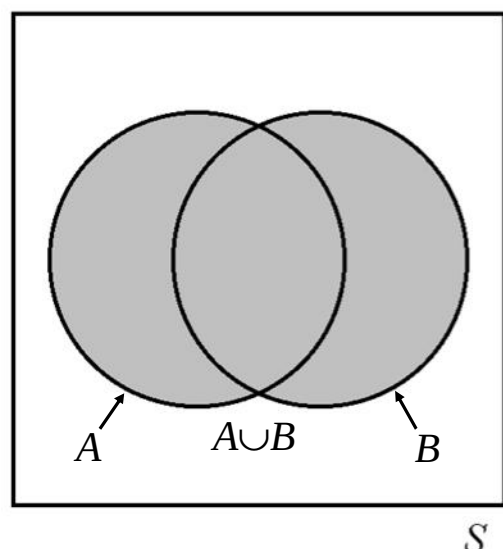
otosavaruuden S tapahtumia.

Tapahtumien A ja B **unioni** eli **yhdiste**

$$A \cup B = \text{"}A \text{ sattuu tai } B \text{ sattuu"}$$

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

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



Ks. kuvaa oikealla.

Tapahtumien A ja B leikkaus

Olkoot

$$A \subset S, B \subset S$$

otosavaruuden S tapahtumia.

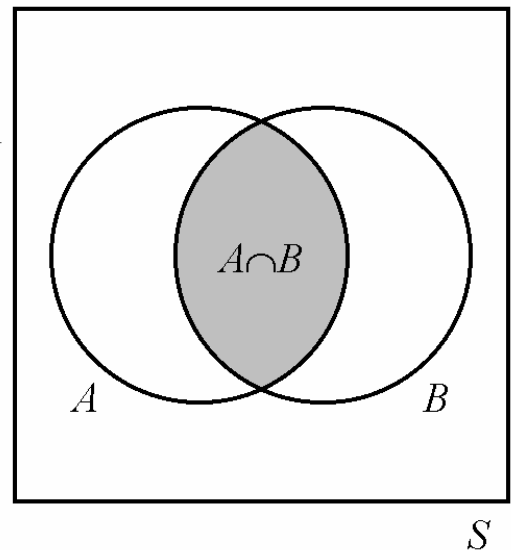
Tapahtumien A ja B **leikkaus**

$$A \cap B = \text{"}A \text{ sattuu ja } B \text{ sattuu"}$$

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

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

Ks. kuvaa oikealla.



Tapahtumien A ja B erotus

Olkoot

$$A \subset S, B \subset S$$

otosavaruuden S tapahtumia.

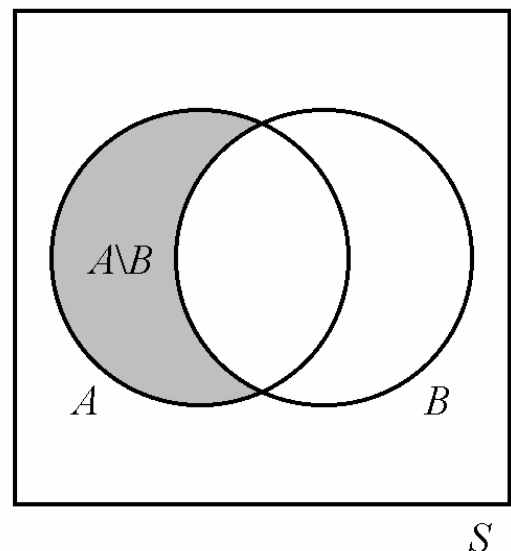
Tapahtumien A ja B **erotus**

$$A \setminus B = \text{"}A \text{ sattuu, mutta } B \text{ ei satu"}$$

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

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

Ks. kuvaa oikealla.



4.3. Todennäköisyyslaskennan peruslaskusäännöt

Oletetaan, että *tuntemme* tapahtumien A ja B todennäköisyydet. Tehtävänä on *määrätä* seuraavat todennäköisyydet:

- (i) Todennäköisyys sille, että **A ei satu**.
- (ii) Todennäköisyys sille, että **A sattuu tai B sattuu tai molemmat sattuvat**.
- (iii) Todennäköisyys sille, että **A sattuu ja B sattuu**.
- (iv) Todennäköisyys sille, että **A sattuu, mutta B ei satu**.

Esimerkkisarja 1: Eduskunta.

Todennäköisyyslaskennan laskusääntöjä havainnollistetaan tässä luvussa useilla esimerkeillä, joissa valitaan *satunnaisesti* yksi tai useampia kansanedustajia kaikkien kansanedustajien joukosta. Esimerkeissä eduskunnan kokoonpano on vuoden 1999 eduskuntavaalien tuloksen mukainen.

Edustajan valintaa *satunnaisesti* voidaan kuvata *arvontana*, joka toteutetaan esim. seuraavalla tavalla:

- (1) Asetetaan jokaista edustajaa vastaamaan *arpalippu*.
- (2) Pannaan arpaliput *urna*an.
- (3) *Sekoitetaan* urnan sisältö huolellisesti.
- (4) *Nostetaan* urnasta arpalippu, jota vastaava edustaja valitaan.

Arvontaa voidaan *toistaa* kahdella erilaisella tavalla:

- (i) Nostettu arpalippu *palautetaan* noston jälkeen urnaan ja urnan sisältö *sekoitetaan* huolellisesti. Tällöin sama edustaja *voi* tulla valituksi *uudelleen*.
- (ii) Nostettua arpalippua *ei palauteta* noston jälkeen urnaan. Tällöin sama edustaja *ei voi* tulla valituksi *kuin kerran*.

Arvontamenetelmää (i) on tapana kutsua **otoksen poimimiseksi palauttaen** eli **takaisinpanolla**, kun taas arvontamenetelmää (ii) on tapana kutsua **otoksen poimimiseksi palauttamatta** eli **ilman takaisinpanoa**. Esimerkiksi *lottoarvonnassa* sovelletaan otantaa ilman takaisinpanoa.

Edustajapaikkojen jakautuminen vuoden 1999 vaaleissa:

Puolue	Paikat	Miehet	Naiset
SDP	51	29	22
Kesk	48	35	13
Kok	46	29	17
Vas	20	14	6
RKP	12	9	3
Vihr	11	2	9
SKL	10	7	3
PS	1	1	0
Rem	1	1	0
Yhteensä	200	127	73

Esimerkkisarja 2: Korttipakka.

Todennäköisyyslaskennan laskusääntöjä havainnollistetaan tässä kappaleessa useilla esimerkeillä, joissa poimitaan *satunnaisesti* pelikortteja hyvin sekoitetusta korttipakasta.

Oletamme, että korttipakassa on 52 korttia (pakassa ei siis ole jokereita).

Korttipakka jakautuu 4:ään *maahan*:

Pata, Hertta, Ruutu, Risti

Jokaisessa maassa on 13 korttia:

A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2

Sanomme, että seuraavat kortit ovat *kuvakortteja*:

A = Ässä (kortti, jonka numerona on 1)
 K = Kuningas
 Q = Kuningatar
 J = Sotilas

Komplementtitapahtuman todennäköisyys

Olkoon A otosavaruuden S tapahtuma ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$.

Tällöin tapahtuman A **komplementtitapahtuman**

$$A^c = \text{"}A \text{ ei satu"} = \{s \in S \mid s \notin A\}$$

todennäköisyys on

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

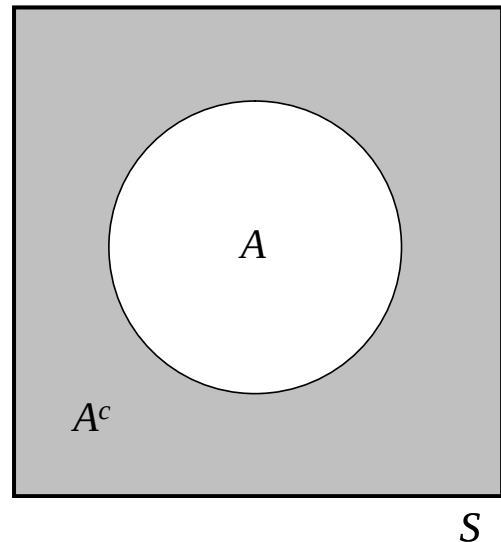
Esimerkki 1. Eduskunta.

Vuoden 1999 eduskunnassa:

$$\begin{aligned}
 n_{\text{Sosiaalistit}} &= n_{SDP} + n_{Vas} \\
 &= 51 + 20 = 71 \\
 n_{\text{Ei-sosiaalistit}} &= 200 - n_{\text{Sosiaalistit}} \\
 &= 200 - 71 = 129
 \end{aligned}$$

Valitaan satunnaisesti yksi edustaja. Todennäköisyys, että valittu edustaja on ei-sosialisti on

$$\Pr(\text{Ei-sosialisti}) = 1 - \Pr(\text{Sosialisti}) = 1 - \frac{71}{200} = \frac{129}{200} = 0.645$$



Toisensa poissulkevat tapahtumat

Olkoot A ja B otosavaruuden S tapahtumia.

Tapahtumat A ja B ovat **toisensa poissulkevia**, jos tapahtumat A ja B eivät voi sattua samanaikaisesti. Tällöin tapahtumat A ja B ovat otosavaruuden S osajoukkoina pistevieraita eli niiden leikkaus on tyhjä:

$$A \cap B = \emptyset$$

Esimerkki 2. Eduskunta.

Yksikään kansanedustaja ei voi olla samaan aikaan kahden eri puolueen jäsen.

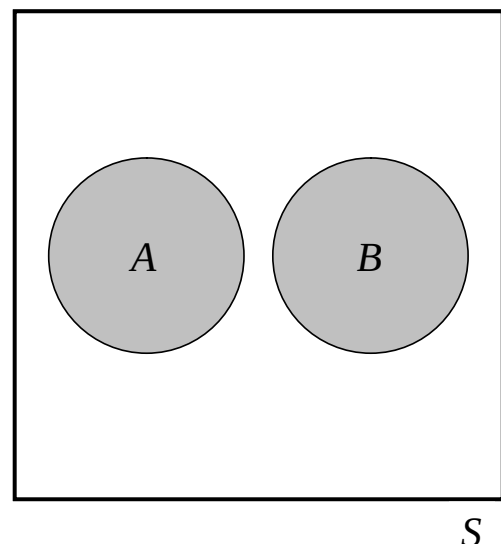
Valitaan satunnaisesti yksi edustaja ja olkoon

$$A = \text{"Edustaja kuuluu vasemmistoliittoon"}$$

ja

$$B = \text{"Edustaja kuuluu kokoomukseen"}$$

Tällöin tapahtumat A ja B ovat **toisensa poissulkevia**, millä siis tarkoitetaan sitä, että A ja B ovat pistevieraita joukkoina:



$$A \cap B = \emptyset$$

Yhteenlaskusääntö toisensa poissulkeville tapahtumille

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$, tapahtuman B todennäköisyys on $\Pr(B)$ ja oletetaan lisäksi, että tapahtumat A ja B ovat *toisensa poissulkevia* eli

$$A \cap B = \emptyset$$

Tällöin tapahtumien A ja B **unionin** eli **yhdisteen**

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

todennäköisyys on

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

Esimerkki 3. Eduskunta.

Vuoden 1999 eduskunnassa:

$$n_{\text{Sosialistit}} = n_{\text{SDP}} + n_{\text{Vas}} = 51 + 20 = 71$$

Valitaan satunnaisesti yksi edustaja. Todennäköisyys, että valittu edustaja on sosialisti on

$$\Pr(\text{Sosialisti}) = \Pr(\text{SDP}) + \Pr(\text{Vas}) = \frac{51}{200} + \frac{20}{200} = \frac{71}{200} = 0.355.$$

Yleistetty yhteenlaskusääntö toisensa poissulkeville tapahtumille

Olkoot $A_1, A_2, \dots, A_k$ otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A_i todennäköisyys on

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

ja oletetaan lisäksi, että tapahtumat $A_1, A_2, \dots, A_k$ ovat *pareittain toisensa poissulkevia*, jolloin

$$A_i \cap A_j = \emptyset, i \neq j$$

Tällöin tapahtumien $A_1, A_2, \dots, A_k$ **yhdisteen**

$$\begin{aligned} A_1 \cup A_2 \cup \dots \cup A_k &= \text{"}A_1 \text{ sattuu tai } A_2 \text{ sattuu tai } \dots \text{ tai } A_k \text{ sattuu"} \\ &= \{s \in S \mid s \in A_1 \text{ tai } s \in A_2 \text{ tai } \dots \text{ tai } s \in A_k\} \end{aligned}$$

todennäköisyys on

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

Tapahtumien riippumattomuus

Tapahtuma A on **riippumaton** tapahtumasta B , jos tapahtuman B :n sattuminen (tai sattumatta jääminen) *ei vaikuta* tapahtuman A sattumisen todennäköisyyteen.

Riippumattomuus on *symmetrinen* ominaisuus: Jos tapahtuma A on riippumaton tapahtumasta B , niin tapahtuma B on riippumaton tapahtumasta A . Riippumattomuuden symmetrisyyden takia sanomme yksinkertaisesti, että

$$\text{"}Tapahtumat A \text{ ja } B \text{ ovat riippumattomia"}$$

Merkitsemme tapahtumien A ja B riippumattomuutta usein seuraavalla tavalla:

$$A \perp B$$

Riippumattomuuden käsite saa lisävalaistusta *ehdollisen todennäköisyyden* käsitteestä; ks. kappaletta **Ehdollinen todennäköisyys**.

Esimerkki 4. Rahanheitto.

Jos heitämme toistuvasti rahaa, on järkevää olettaa, että yhdenkään heiton tulos ei riipu aikaisemmin tehtyjen heittojen tuloksista.

Esimerkki 5. Arvonta.

Nostetaan uurnasta toistuvasti arpalippuja niin, että jokaisen noston jälkeen nostettu lippu palautetaan uurna ja uurnan sisältö sekoitetaan huolellisesti (eli arpalippujen nostossa sovelletaan *otantaa takaisinpanolla*). Tällöin on järkevää olettaa, että nostojen tulokset eivät riipu aikaisemmin tehtyjen nostojen tuloksista.

Riippumattomuus vs riippuvuus

Alkeistilastotieteen esityksissä oletetaan tavallisesti, että satunnaisilmiötä koskevat *havainnot* ovat *riippumattomia*. Monissa tilastotieteen tutkimusasetelmissä oletus havaintojen riippumattomuudesta on kuitenkin liian rajoittava. Esimerkiksi *ennustaminen aikasarja-analyysissä* perustuu siihen, että *tulevaisuus riippuu aikasarjan historiasta*.

On syytä huomata, että *riippumattomuus* on oletus, jota voidaan *testata* tilastollisesti.

Tulosääntö riippumattomille tapahtumille

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$, tapahtuman B todennäköisyys on $\Pr(B)$ ja lisäksi, että tapahtumat A ja B **riippumattomia**.

Tällöin tapahtumien A ja B **leikkauksen**

$$\begin{aligned} A \cap B &= \text{"}A \text{ sattuu ja } B \text{ sattuu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \in B\} \end{aligned}$$

todennäköisyys on

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

On osoittautunut järkeväksi pitää tapahtumia A ja B *riippumattomina*, jos ja vain jos **tulosääntö**

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

pätee.

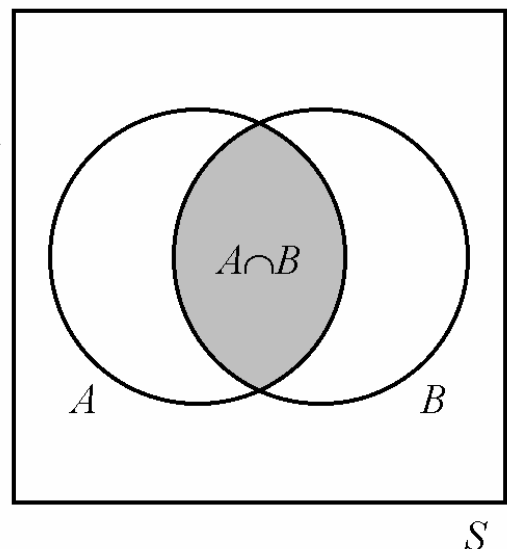
Esimerkki 6. Tulosäännön intuitiivinen perustelu.

Tarkastellaan TV-vastaanottimien laadunvalvontaa. Oletetaan, että vastaanottimissa voi esiintyä kaksi erilaista vikaa:

A = "Kuvaputki ei toimi"

B = "Vahvistin ei toimi"

Oletetaan lisäksi, että viat syntyvät *toisistaan riippumatta*.



Laadunvalvonnassa on todettu, että vika A esiintyy 5 %:ssa valmistettuja TV-vastaanottimia ja vika B esiintyy 3 %:ssa valmistettuja TV-vastaanottimia.

Siitä, että viat A ja B syntyvät *toisistaan riippumattomatta* seuraa:

- (i) 3 %:ssa niistä TV-vastaanottimia, joissa on vika A , on *myös* vika B .
- (ii) 5 %:ssa niistä TV-vastaanottimia, joissa on vika B , on *myös* vika A .

Siten niiden TV-vastaanottimien osuus, joissa on *sekä* vika A *että* vika B on

$$0.03 \times 0.05 = 0.0015 = 0.15 \%$$

Saatu tulos vastaa riippumattomien tapahtumien tulosääntöä: Olkoon

$$A \cap B = \text{"Kuvaputki ei toimi ja vahvistin ei toimi"}$$

Tällöin

$$\Pr(A \cap B) = \Pr(A)\Pr(B) = 0.03 \times 0.05 = 0.0015$$

Esimerkki 7. Rahanheitto.

Heitetään rahaa kaksi kertaa ja tarkastellaan tapahtumia

$$A = \text{"Kruuna 1. heitolla"}$$

$$B = \text{"Kruuna 2. heitolla"}$$

Voimme olettaa, että tapahtumat A ja B ovat *riippumattomia*.

Oletetaan, että

$$\Pr(\text{Kruuna}) = \Pr(\text{Klaava}) = 1/2$$

Tällöin

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

Esimerkki 8. Korttipakka.

Nostetaan korttipakasta satunnaisesti 1 kortti ja tarkastellaan seuraavia tapahtumia:

$$A = \text{"Kortti on pata"}$$

$$B = \text{"Kortti on ässä"}$$

Koska korttipakassa 13 pataa ja 4 ässää, niin

$$\Pr(A) = \frac{13}{52} = \frac{1}{4}$$

$$\Pr(B) = \frac{4}{52} = \frac{1}{13}$$

Tarkastellaan yhdistettyä tapahtumaa

$$A \cap B = \text{"Kortti on pataässä"}$$

Koska pataässiä on täsmälleen 1 kappale, niin

$$\Pr(A \cap B) = \frac{1}{52}$$

Koska

$$\Pr(A \cap B) = \frac{1}{52} = \frac{1}{13} \cdot \frac{1}{4} = \Pr(A) \Pr(B)$$

voimme pitää tapahtumia A ja B riippumattomina.

Intuitiivinen perustelu tulosäännön soveltamiselle: Korttipakan korteista $1/4$ on patoja ja $1/13$ on ässiä. Myös niistä korteista, jotka ovat patoja $1/13$ on ässiä. Siten pataässien osuus korttipakan korteista on

$$\frac{1}{13} \cdot \frac{1}{4} = \frac{1}{52}$$

Yleistetty tulosääntö riippumattomille tapahtumille

Olkoot $A_1, A_2, \dots, A_k$ otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A_i todennäköisyys on

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

Sanomme, että tapahtumat $A_1, A_2, \dots, A_k$ ovat **riippumattomia**, jos ja vain jos *kaikille* leikkauksille

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

joissa

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

pätee:

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

Merkitsemme tapahtumien $A_1, A_2, \dots, A_k$ riippumattomuutta usein seuraavalla tavalla:

$$A_1, A_2, \dots, A_k \perp$$

Esimerkki 9. Rahanheitto.

Heitetään rahaa 10 kertaa ja tarkastellaan tapahtumia

$$A_i = \text{"Kruuna } i. \text{ heitolla"} , i = 1, 2, \dots, k$$

Voimme olettaa, että tapahtumat A_i ovat *riippumattomia*.

Oletetaan, että

$$\Pr(A_i) = \Pr(\text{Kruuna } i. \text{ heitolla}) = 1/2, i = 1, 2, \dots, k$$

Tällöin

$$\Pr(10 \text{ kruunaa } 10:\text{llä heitolla})$$

$$= \Pr(\text{Kruuna } 1. \text{ heitolla ja Kruuna } 2. \text{ heitolla ja } \dots \text{ ja Kruuna } 10. \text{ heitolla})$$

$$= \Pr(\text{Kruuna } 1. \text{ heitolla}) \times \Pr(\text{Kruuna } 2. \text{ heitolla}) \times \dots \times \Pr(\text{Kruuna } 10. \text{ heitolla})$$

$$= \frac{1}{2} \times \frac{1}{2} \times \dots \times \frac{1}{2} = \frac{1}{2^{10}} = \frac{1}{1024} = 0.000977$$

10 kappaletta

Todennäköisyyden frekvenssitulkinnasta seuraa: Jos suuresta joukosta ihmisiä jokainen pannaan heittämään rahaa 10 kertaa, on odotettavissa, että suunnilleen yksi tuhannesta saa tulokseksi 10 kruunaa!

Yleinen yhteenlaskusääntö

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$, tapahtuman B todennäköisyys on $\Pr(B)$ ja leikkaustapahtuman

$$\begin{aligned} A \cap B &= \text{"}A \text{ sattuu ja } B \text{ sattuu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \in B\} \end{aligned}$$

todennäköisyys on $\Pr(A \cap B)$.

Tällöin tapahtumien A ja B **unionin** eli **yhdisteen**

$$\begin{aligned} A \cup B &= \text{"}A \text{ sattuu tai } B \text{ sattuu"} \\ &= \{s \in S \mid s \in A \text{ tai } s \in B\} \end{aligned}$$

todennäköisyys on

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

Esimerkki 10. Yleisen yhteenlaskusäännön motivointi.

Tilastojen mukaan 80 % japanilaisista on shintolaisia ja 80 % japanilaisista on buddhalaisia.

Kysymys: Onko 160 % japanilaisista shintolaisia *tai* buddhalaisia?

Vastaus: *Ei*, koska suuri osa japanilaisista noudattaa *kummankin* uskonnon menoja: Häät järjestetään tavallisesti shintolaisten menojen mukaan, kun taas hautajaiset järjestetään tavallisesti buddhalaisten menojen mukaan.

Annetuista tiedoista voidaan päätellä:

- 80 % – 100 % japanilaisista on *joko* shintolaisia *tai* buddhalaisia
- 60 % – 80 % japanilaisista on *sekä* shintolaisia *että* buddhalaisia

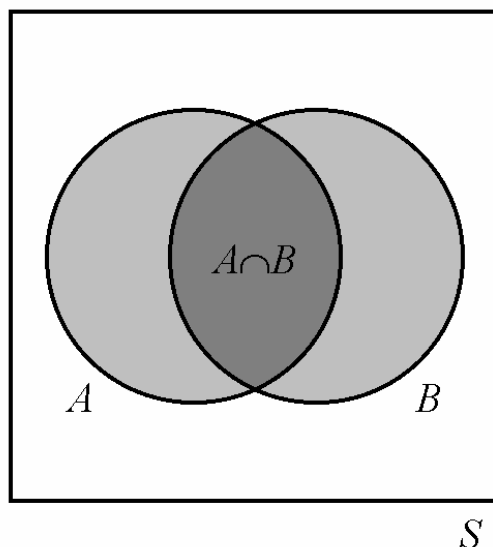
Esimerkki 11. Levikkitutkimus.

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

Seura	20 %
Apu	16 %
Seura ja Apu	1 %

Tällöin

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



Erotustapahtuman todennäköisyys

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$ ja leikkaustapahtuman

$$\begin{aligned} A \cap B &= \text{"}A \text{ sattuu ja } B \text{ sattuu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \in B\} \end{aligned}$$

todennäköisyys on $\Pr(A \cap B)$.

Tällöin **erotustapahtuman**

$$\begin{aligned} A \setminus B &= \text{"}A \text{ sattuu, mutta } B \text{ ei satu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \notin B\} \\ &= A \cap B^c \end{aligned}$$

todennäköisyys on

$$\Pr(A \setminus B) = \Pr(A \cap B^c) = \Pr(A) - \Pr(A \cap B)$$

Esimerkki 12. Korttipakka.

Nostetaan korttipakasta satunnaisesti 1 kortti ja tarkastellaan seuraavia tapahtumia:

A = "Kortti on pata"

B = "Kortti on kuva"

Tällöin

$A \setminus B$ = "Kortti on pata, mutta ei ole kuvakortti"

Koska korttipakassa on 13 patakorttia, joista 4 on kuvia, niin

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

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

Tapahtumasta B seuraa tapahtuma A

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$ ja tapahtuman B todennäköisyys on $\Pr(B)$.

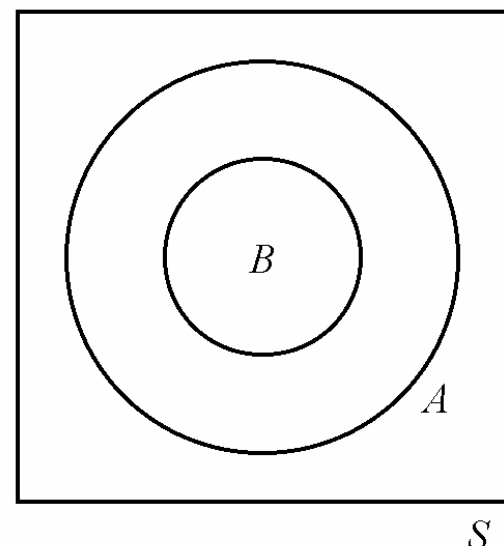
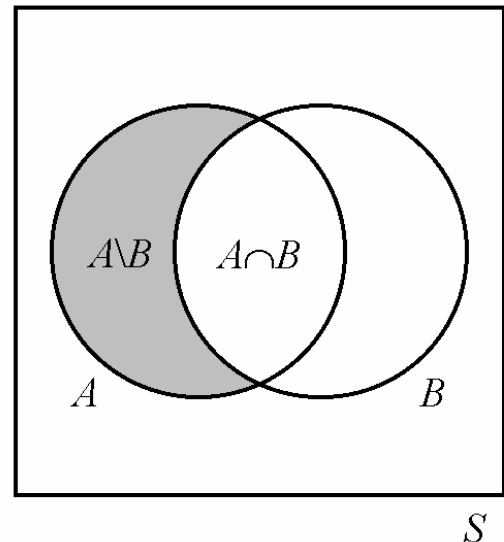
Oletetaan, että jos tapahtuma B sattuu, niin tapahtuma A sattuu.

Tällöin

$$B \subset A$$

ja siis

$$\Pr(B) \leq \Pr(A)$$



Erotustapahtuman todennäköisyys, jos tapahtumasta B seuraa tapahtuma A

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$ ja tapahtuman B todennäköisyys on $\Pr(B)$.

Oletetaan, että jos tapahtuma B sattuu, niin tapahtuma A sattuu, mikä siis merkitsee sitä, että

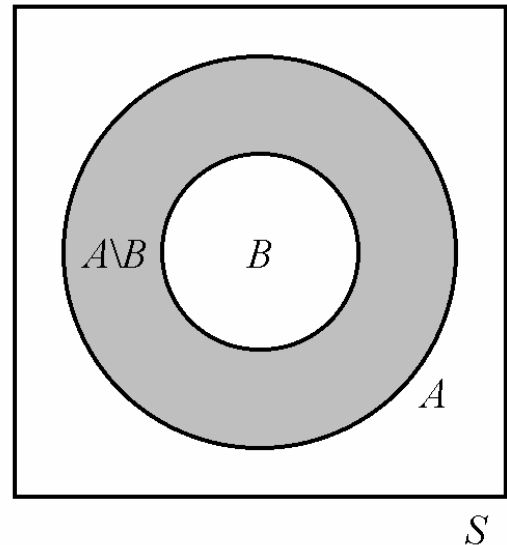
$$B \subset A$$

Tällöin erotustapahtuman

$$\begin{aligned} A \setminus B &= \text{"}A \text{ sattuu, mutta } B \text{ ei satu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \notin B\} \\ &= A \cap B^c \end{aligned}$$

todennäköisyys on

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



Yhdisteen todennäköisyys

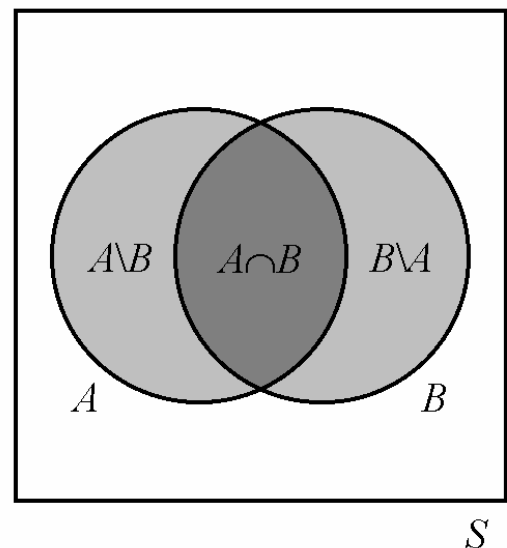
Olkoot A ja B otosavaruuden S tapahtumia.

Tällöin tapahtumien A ja B **unionin** eli **yhdisteen**

$$\begin{aligned} A \cup B &= \text{"}A \text{ sattuu tai } B \text{ sattuu"} \\ &= \{s \in S \mid s \in A \text{ tai } s \in B\} \end{aligned}$$

todennäköisyys on

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



4.4. Ehdollinen todennäköisyys ja riippumattomuus

Ehdollinen todennäköisyys

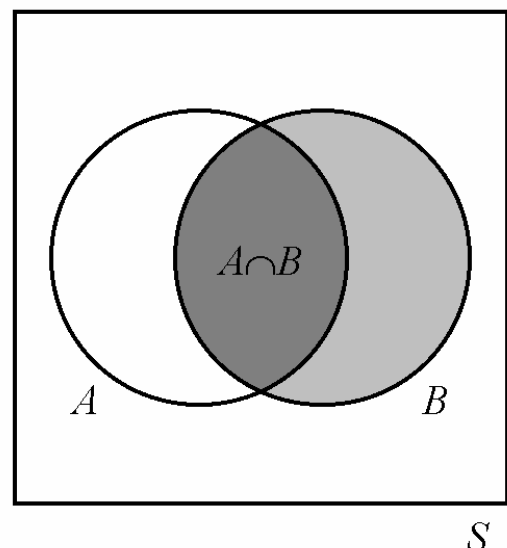
Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman B todennäköisyys

$$\Pr(B) \neq 0$$

ja leikkaustapahtuman

$$\begin{aligned} A \cap B &= \text{"}A \text{ sattuu ja } B \text{ sattuu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \in B\} \end{aligned}$$

todennäköisyys on $\Pr(A \cap B)$.



Tällöin **tapahtuman A ehdollinen todennäköisyys ehdolla, että tapahtuma B on sattunut** saadaan kaavalla

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

Tapahtuman A ehdollinen todennäköisyys ehdolla, että tapahtuma B on sattunut, saadaan määräämällä leikkaustapahtuman

$$A \cap B = \text{"}A \text{ sattuu ja } B \text{ sattuu"}$$

todennäköisyyden $\Pr(A \cap B)$ suhde tapahtuman B todennäköisyyteen $\Pr(B)$. *Tapahtuman A ehdollinen todennäköisyys ehdolla, että tapahtuma B on sattunut voidaan siis ymmärtää niin, että tapahtumaan A liittyviä alkeistapahtumia tarkastellaan rajoitettuna tapahtuman B indusoimaan otosavaruuteen.*

Esimerkki 1. Ehdollisen todennäköisyyden kaavan motivointi.

Tarkastellaan *toistuvaa* satunnaisilmiötä ja olkoot A ja B kaksi ko. satunnaisilmiöön liittyvää tapahtumaa.

Satunnaisilmiön toistuessa *jokaisella toistokerralla* pätee:

- (i) *Joko* tapahtuma A sattuu *tai* tapahtuma $A^c = ei-A$ sattuu.
- (ii) *Joko* tapahtuma B sattuu *tai* tapahtuma $B^c = ei-B$ sattuu.

Oletetaan, että ko. satunnaisilmiön toistuessa tuloksena on seuraava tapahtumaparien jono:

$$A^c B \quad AB^c \quad A^c B \quad AB \quad A^c B^c \quad A^c B \quad A^c B^c$$

Tässä tapahtumajonossa tapahtuman A todennäköisyys on

$$\Pr(A) = \frac{2}{7}$$

ja tapahtuman B todennäköisyys on

$$\Pr(B) = \frac{4}{7}$$

Lisäksi leikkaustapahtuman

$$A \cap B = \text{"}A \text{ sattuu ja } B \text{ sattuu"}$$

todennäköisyys on

$$\Pr(A \cap B) = \frac{1}{7}$$

Muodostetaan tässä tarkastellusta tapahtumaparien jonosta

$$A^c B \quad AB^c \quad A^c B \quad AB \quad A^c B^c \quad A^c B \quad A^c B^c$$

karsittu jono, johon otetaan vain ne tapahtumaparit, joissa tapahtuma B on sattunut. Nämä tapahtumaparit ovat

$$A^c B \quad A^c B \quad AB \quad A^c B$$

Tapahtuman A todennäköisyys tässä *karsitussa* jonossa on $1/4$.

Toisaalta *ehdollisen todennäköisyyden* kaavan mukaan

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{1/7}{4/7} = \frac{1}{4}$$

Siten tapahtuman A ehdollinen todennäköisyys ehdolla, että B on sattunut, on tapahtuman A todennäköisyys *karsitussa* jonossa, johon on mukaan otettu vain ne tapahtumaparit, joissa B on sattunut.

Esimerkki 2. Eduskunta.

Vuoden 1999 eduskunnan 200 kansanedustajasta 73 oli naisia. Valitaan satunnaisesti 1 edustaja. Mikä on todennäköisyys, että valittu edustaja on nainen?

Merkitään

$A = \text{"Edustaja on nainen"}$

Tällöin

$$\Pr(A) = \frac{73}{200} = 0.365$$

SDP:llä oli vuoden 1999 eduskunnassa 51 kansanedustajaa, 29 miestä ja 22 naista. Mikä on todennäköisyys, että satunnaisesti valittu edustaja, joka kuuluu SDP:n ryhmään, on nainen?

Ehdollisen todennäköisyyden kaavan mukaan

$$\Pr(\text{Nainen}|\text{SDP}) = \frac{\Pr(\text{Nainen ja SDP})}{\Pr(\text{SDP})} = \frac{22/200}{51/200} = \frac{22}{51} = 0.431$$

joten

$$\Pr(\text{Nainen}|\text{SDP}) = 0.431 > 0.365 = \Pr(\text{Nainen})$$

Tieto siitä, että satunnaisesti valittu kansanedustaja on SDP:stä on muuttanut käsitystämme todennäköisyydestä, että hän on nainen.

RKP:lla oli vuoden 1999 eduskunnassa 12 kansanedustajaa, 9 miestä ja 3 naista. Mikä on todennäköisyys, että satunnaisesti valittu edustaja, joka kuuluu RKP:n ryhmään, on nainen?

Ehdollisen todennäköisyyden kaavan mukaan

$$\Pr(\text{Nainen}|\text{RKP}) = \frac{\Pr(\text{Nainen ja RKP})}{\Pr(\text{RKP})} = \frac{3/200}{12/200} = \frac{3}{12} = 0.25$$

joten

$$\Pr(\text{Nainen}|\text{RKP}) = 0.25 < 0.365 = \Pr(\text{Nainen})$$

Tieto siitä, että satunnaisesti valittu kansanedustaja on RKP:stä on muuttanut käsitystämme todennäköisyydestä, että hän on nainen.

Edellisen esimerkin mukaan tapahtuman A ehdollinen todennäköisyys ehdolla, että tapahtuma B on sattunut voi olla *pienempi, yhtä suuri tai suurempi* kuin tapahtuman A todennäköisyys.

Siten mikä tahansa seuraavista kolmesta vaihtoehdosta on mahdollinen:

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

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

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

Jos

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

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

Sen sijaan, jos

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

niin tieto siitä, että *tapahtuma B on sattunut, ei sisällä informaatiota*, jota voidaan käyttää hyväksi tapahtuman A todennäköisyyttä määrittäessä. Tällöin on järkevää pitää tapahtumaa A *riippumattomana* tapahtumasta B.

Riippumattomuus

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman A todennäköisyys on $\Pr(A)$ ja tapahtuman B todennäköisyys on $\Pr(B)$.

Tapahtumat A ja B ovat **riippumattomia**, jos ja vain jos tapahtumien A ja B leikkauksen

$$\begin{aligned} A \cap B &= \text{"A sattuu ja B sattuu"} \\ &= \{s \in S \mid s \in A \text{ ja } s \in B\} \end{aligned}$$

todennäköisyys on

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

Lause 1.

Olkoon tapahtuman A todennäköisyys ehdolla, että B on sattunut $\Pr(A|B)$. Tällöin tapahtumat A ja B ovat **riippumattomia**, jos ja vain jos

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

Todistus:

- (i) Oletetaan, että tapahtumat A ja B ovat *riippumattomia* ja todistetaan, että

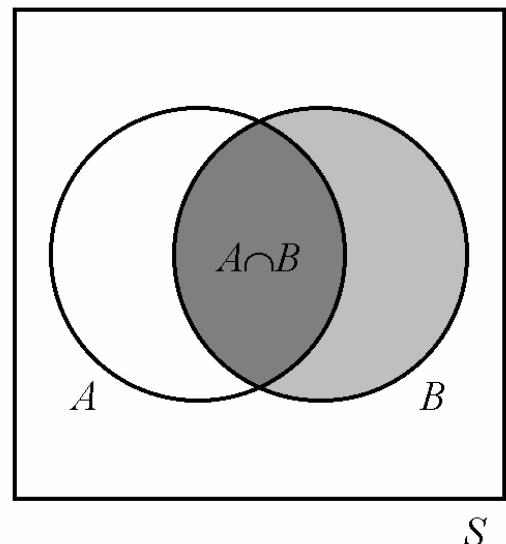
$$\Pr(A|B) = \Pr(A)$$

Riippumattomuuden määritelmän mukaan

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

Tällöin

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



(ii) Oletetaan, että

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

ja todistetaan, että tapahtumat A ja B ovat *riippumattomia*.

Oletuksesta ja ehdollisen todennäköisyyden määritelmästä seuraa, että

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

joten tapahtumat A ja B ovat riippumattomia.

■

Koska riippumattomuus on symmetrinen ominaisuus on helppo nähdä, että tapahtumat A ja B ovat **riippumattomia**, jos ja vain jos mikä tahansa seuraavista *ehdoista* pätee:

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

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

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

Esimerkki 3. Rutenian parlamentti.

Rutenian tasavallassa on 2 puoluetta: repijät ja säilyttäjät.

Puolueiden paikkajakauma 200-paikkaisessa parlamentissa:

Puolue	Miehet	Naiset	Paikat
Repijät	20	30	50
Säilyttäjät	60	90	150
Yhteensä	80	120	200

Valitaan parlamentista satunnaisesti 1 edustaja ja olkoon

A = ”Valittu edustaja on mies”

B = ”Valittu edustaja on repijä”

Tällöin

$$\Pr(A) = \frac{80}{200} = 0.4$$

$$\Pr(A | B) = \frac{20}{50} = 0.4$$

Koska

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

niin tapahtumat A ja B ovat riippumattomia.

On helppo nähdä, että myös tapahtumat

A ja B^c , A^c ja B , A^c ja B^c

ovat riippumattomia.

Koska

$$\Pr(A \cap B) = \frac{20}{200}$$

$$\Pr(B) = \frac{50}{200}$$

niin ehdollisen todennäköisyyden kaavasta seuraa, että

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{20/200}{50/200} = \frac{20}{50} = 0.4$$

On syytä huomata, että tapahtumien A ja B riippumattomuus ”leviää” aina myös niiden *komplementtitapahtumiin*; ks. edellistä esimerkkiä. Jos siis oletamme, että tapahtumat A ja B ovat riippumattomia, niin tällöin pätee seuraava:

- (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.

Yleinen tulosääntö

Olkoot A ja B otosavaruuden S tapahtumia ja oletetaan, että tapahtuman B todennäköisyys

$$\Pr(B) \neq 0$$

ja tapahtuman A ehdollinen todennäköisyys ehdolla, että tapahtuma B on sattunut on

$$\Pr(A|B)$$

Tällöin tapahtumien A ja B **leikkauksen**

$$\begin{aligned} A \cap B &= \text{”}A \text{ sattuu ja } B \text{ sattuu”} \\ &= \{s \in S \mid s \in A \text{ ja } s \in B\} \end{aligned}$$

todennäköisyys on

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

Kutsumme tätä sääntöä **yleiseksi tulosäännöksi**.

Yleistetty yleinen tulosääntö

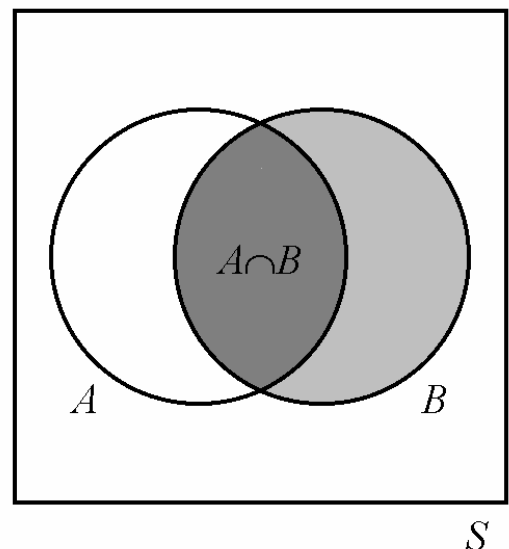
Olkoot $A_1, A_2, \dots, A_k$ otosavaruuden S tapahtumia.

Tällöin tapahtumien $A_1, A_2, \dots, A_k$ **leikkauksen**

$$\begin{aligned} A_1 \cap A_2 \cap \dots \cap A_k &= \text{”}A_1 \text{ sattuu ja } A_2 \text{ sattuu ja } \dots \text{ ja } A_k \text{ sattuu”} \\ &= \{s \in S \mid s \in A_1 \text{ ja } s \in A_2 \text{ ja } \dots \text{ ja } s \in A_k\} \end{aligned}$$

todennäköisyys on

$$\begin{aligned} \Pr(A_1 \cap A_2 \cap \dots \cap A_k) \\ = \Pr(A_1) \times \Pr(A_2 | A_1) \times \Pr(A_3 | A_1 \cap A_2) \times \dots \times \Pr(A_k | A_1 \cap A_2 \cap \dots \cap A_{k-1}) \end{aligned}$$



Esimerkki 4. Korttipakka.

Nostetaan korttipakasta *peräkkäin* 3 korttia. Mikä on todennäköisyys, että ne ovat kaikki patoja?

Olkoon tapahtuma

$$A_i = \text{"}i\text{. kortti on pata"}, i = 1, 2, 3$$

Koska korttipakassa on 52 korttia, joista 13 on patoja, leikkauksen

$$\text{"}A_1 \text{ sattuu ja } A_2 \text{ sattuu ja } A_3 \text{ sattuu"}$$

todennäköisyys on

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

Huomaa, että korttien nosto toteutettiin tässä *ilman takaisinpanoa*.

Yksinkertainen satunnaisotanta ja tulosäännöt

Olkoon perusjoukko S äärellinen. **Yksinkertaisessa satunnaisotannassa** perusjoukosta S poimitaan osajoukko B *arpomalla* perusjoukosta alkioita osajoukkoon B yksi alkio kerrallaan. Osajoukkoa B kutsutaan **otokseksi** ja arvonnassa käytettyä menetelmää **otantamenetelmäksi**; lisätietoja otannasta: ks. monistetta **Tilastolliset menetelmät**.

Tarkastellaan todennäköisyyttä saada otokseen B alkioita perusjoukon S osajoukosta A . Jos *otanta tehdään ilman takaisinpanoa* eli *palauttamatta poimittua alkioita takaisin perusjoukkoon*, poimintatodennäköisyyksiä määrättäessä on sovellettava *yleistä tulosääntöä*. Jos *otanta tehdään takaisinpanolla* eli *palauttamalla poimittu alkio aina takaisin perusjoukkoon*, poimintatodennäköisyyksiä määrättäessä on sovellettava *riippumattomien tapahtumien tulosääntöä*.

Luvussa **Diskreettejä jakaumia** todetaan, että äärellisen perusjoukon S tapahtumien todennäköisyyksiä hallitaan ns. **binomijakaumalla**, jos otanta on tapahtunut palauttaen ja ns. **hypergeometrisella jakaumalla**, jos otanta on tapahtunut palauttamatta.

5. Klassinen todennäköisyys ja kombinatoriikka

- 5.1. Klassinen todennäköisyys
- 5.2. Kombinatoriikan perusperiaatteet
- 5.3. Kombinatoriikan perusongelmat
- 5.4. Kombinatoriikan perusongelmien ratkaiseminen
- 5.5. Multinomikerroin

Tapahtuman **klassisen todennäköisyyden** määrittäminen vaatii sekä ko. *tapahtumalle suotuisien alkeistapahtumien lukumäärän* että *kaikkien mahdollisten alkeistapahtumien lukumäärän laskemista*. Tässä luvussa näytämme miten tapahtumavaihtoehtojen lukumääriä voidaan laskea soveltamalla **kombinatoriikkaa**.

Saaduilla tuloksilla on suuri merkitys – paitsi klassisen todennäköisyyden määrittämisestä vaativissa tehtävissä – johdettaessa **diskreettien todennäköisyysjakaumien pistetodennäköisyysfunktioita**; ks. lukua **Diskreettejä jakaumia**.

Avainsanat:

Binomikaava, Binomikerroin, Jono, Joukko, Järjestys, Kertolaskuperiaate, Kertoma, Klassinen todennäköisyys, Kombinaatio, Kombinatoriikka, Lokeromalli, Lukumääräfunktio, Multinomikerroin, Osajono, Osajoukko, Pascalin kolmio, Permutaatio, Riippumattomuus, Suotuisa alkeistapahtuma, Symmetrisyys, Variaatio, Yhteenlaskuperiaate

5.1. Klassinen todennäköisyys

Oletetaan, että äärellisen otosavaruuden

$$S = \{s_1, s_2, \dots, s_n\}$$

alkeistapahtumat

$$s_i, i = 1, 2, \dots, n$$

ovat yhtä todennäköisiä eli

$$\Pr(s_i) = \frac{1}{n}, i = 1, 2, \dots, n$$

jolloin sanomme, että alkeistapahtumat $s_i, i = 1, 2, \dots, n$ ovat **symmetrisiä**.

Olkoon

$$n_A = n(A)$$

funktio, joka kertoo joukon A alkioden lukumäärän. Jos siis

$$A = \{a_1, a_2, \dots, a_k\}$$

on äärellinen joukko, jonka (erilaisten) alkioden lukumäärä on k , niin

$$n_A = n(A) = k$$

Kutsumme funktiota $n(\cdot)$ **lukumääräfunktioksi**.

Olkoon tapahtuma A otosavaruuden S osajoukko. Tällöin tapahtuman A **klassinen todennäköisyys** $\Pr_c(A)$ saadaan määrittämällä tapahtumalle A *suotuisien alkeistapahtumien suhteellinen osuus* kaikista mahdollisista alkeistapahtumista eli

$$\Pr_c(A) = \frac{n_A}{n_S} = \frac{n(A)}{n(S)}$$

jossa

$$\begin{aligned} n_A = n(A) &= \text{ tapahtumalle } A \text{ suotuisien alkeistapahtumien lukumäärä} \\ &= \text{ joukkoon } A \text{ kuuluvien alkeistapahtumien lukumäärä} \end{aligned}$$

ja

$$\begin{aligned} n_S = n(S) &= \text{ kaikkien mahdollisten alkeistapahtumien lukumäärä} \\ &= \text{ otosavaruuteen } S \text{ kuuluvien alkeistapahtumien lukumäärä} \end{aligned}$$

Jos perusjoukko (otosavaruus) on kooltaan vähänkin isompi, perusjoukon ja sen osajoukkojen (tapahtumien) alkioden (alkeistapahtumien) *lukumäärien* laskemisessa tarvitaan apuna jotakin järjestelmällistä menetelmää. Tällaisen järjestelmällisen menetelmän joukon alkioden lukumäärän laskemiseen tarjoaa **kombinatoriikaksi** kutsuttu matematiikan osa-alue.

5.2. Kombinatoriikan perusperiaatteet

Tarkastellaan kahta **operaatiota**, M ja N . Tehdään operaatioista M ja N seuraavat oletukset:

Operaatio M voidaan suorittaa m :llä erilaisella tavalla.

Operaatio N voidaan suorittaa n :llä erilaisella tavalla.

Operaatiot M ja N voidaan yhdistää uudeksi, **yhdistetyksi operaatioksi** seuraavilla tavoilla:

- (i) ”Suoritetaan operaatio M ja operaatio N ”
- (ii) ”Suoritetaan operaatio M tai operaatio N ”

Kombinatoriikan perusperiaatteet liittyvät näiden kahden yhdistetyn operaation *suoritustapojen lukumäärien* laskemiseen.

- (i) **Kertolaskuperiaate:** Oletetaan, että operaatio M voidaan suorittaa m :llä erilaisella tavalla ja operaatio N voidaan suorittaa n :llä erilaisella tavalla ja oletetaan lisäksi, että operaatiot M ja N voidaan suorittaa toisistaan **riippumatta**. Tällöin yhdistetty operaatio

”Suoritetaan operaatio M ja operaatio N ”

voidaan suorittaa $m \times n$:llä erilaisella tavalla.

- (ii) **Yhteenlaskuperiaate:** Oletetaan, että operaatio M voidaan suorittaa m :llä erilaisella tavalla ja operaatio N voidaan suorittaa n :llä erilaisella tavalla ja oletetaan lisäksi, että operaatiot M ja N ovat **toisensa poissulkevia**. Tällöin yhdistetty operaatio

”Suoritetaan operaatio M tai operaatio N ”

voidaan suorittaa $(m + n)$:llä erilaisella tavalla.

Esimerkki 1. Vaihtoehtojen lukumäärien laskeminen.

Oletetaan, että kaupunkien X ja Y välillä on 2 suoraa lentoa, mutta että X :stä Y :hyn pääsee myös kaupungin Z kautta:

- (i) Kaupunkien X ja Z välillä on 3 lentoa.
- (ii) Kaupunkien Z ja Y välillä on 2 lentoa.

Ks. kuvaa oikealla.

Oletetaan vielä, että lentojen valinnat voidaan tehdä toisistaan **riippumatta**.

Kuinka monella erilaisella tavalla voidaan lentää X :stä Y :hyn?

Koska lentojen valinnat voidaan tehdä toisistaan **riippumatta**, Z :n kautta tapahtuviin lentoihin voidaan soveltaa kombinatoriikan **kertolaskuperiaatetta**. Sen mukaan X :stä Y :hyn pääsee lentämään Z :n kautta

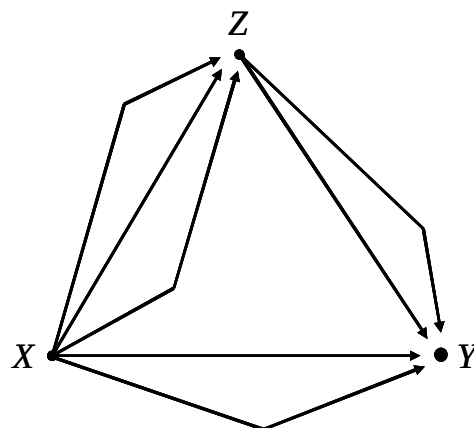
$$3 \times 2 = 6$$

erilaisella tavalla.

Koska suorat lennot X :stä Y :hyn ja lennot Z :n kautta ovat **toisensa poissulkevia**, lentojen kokonaismäärä saadaan soveltamalla kombinatoriikan **yhteenlaskuperiaatetta**. Sen mukaan X :stä Y :hyn pääsee lentämään kaikkiaan

$$2 + 6 = 8$$

erilaisella tavalla.



5.3. Kombinatoriikan perusongelmat

Olkoon S äärellinen joukko, jonka (erilaisten) alkioiden lukumäärä on

$$n = n_S = n(S)$$

Kombinatoriikan perusongelmat:

- (1a) Kuinka monella erilaisella tavalla joukon S alkiot voidaan järjestää **jonoon**?
- (1b) Kuinka monella erilaisella tavalla joukon S alkioista voidaan muodostaa k :n alkion **osajono**?
- (2) Kuinka monella erilaisella tavalla joukon S alkioista voidaan muodostaa k :n alkion **osajoukko**?

Huomaa, että ongelma (1a) on ongelman (1b) erikoistapaus.

Joukko

Joukko on täysin määrätty, jos sen *alkiot* tunnetaan. Olkoot äärellisen joukon S (erilaiset) alkiot

$$s_1, s_2, \dots, s_n$$

Tällöin merkitsemme joukkoa S seuraavalla tavalla:

$$S = \{s_1, s_2, \dots, s_n\}$$

Joukot A ja B ovat **samat**, jos niissä on samat alkiot eli

$$A = B$$

jos ja vain jos

$$x \in A \Leftrightarrow x \in B$$

Jono

Jono on täysin määrätty, jos sen *alkiot* ja niiden *järjestys* tunnetaan. Olkoon äärellisen jonon s i . alkio

$$s_i, i = 1, 2, \dots, n$$

Tällöin merkitsemme jonoa s seuraavalla tavalla:

$$s = (s_1, s_2, \dots, s_n)$$

tai hieman yksinkertaisemmin

$$s = s_1 s_2 \cdots s_n$$

Jonot $a = (a_1, a_2, \dots, a_n)$ ja $b = (b_1, b_2, \dots, b_n)$ ovat **samat**, jos niissä on samat alkiot samassa järjestyksessä eli

$$a = b$$

jos ja vain jos

$$a_i = b_i, i = 1, 2, \dots, n$$

Esimerkki 1. Joukkojen ja jonojen samuus.

Joukot

$$\{1, 2, 3\}, \{1, 3, 2\}, \{3, 1, 2\}$$

ovat samat, koska niissä on samat alkiot:

$$\{1, 2, 3\} = \{1, 3, 2\} = \{3, 1, 2\}$$

Jonot

123

132

ovat eri jonoja, koska niiden alkiot ovat eri järjestyksessä:

$$(1, 2, 3) \neq (1, 3, 2)$$

Esimerkki 2. Osajoukkojen ja osajonojen lukumäärien laskeminen.

Olkoon

$$S = \{1, 2, 3\}$$

Kaikki joukon S alkioden muodostamat *osajoukot*:

Kolmen alkion osajoukot:

$$\{1, 2, 3\} \quad 1 \text{ kpl}$$

Kahden alkion osajoukot:

$$\{1, 2\}, \{1, 3\}, \{2, 3\} \quad 3 \text{ kpl}$$

Yhden alkion osajoukot:

$$\{1\}, \{2\}, \{3\} \quad 3 \text{ kpl}$$

Kaikki joukon S :n osajoukot:

$$\{1, 2, 3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1\}, \{2\}, \{3\}, \emptyset \quad 8 \text{ kpl}$$

Kaikki joukon S alkioden muodostamat *osajonot*:

Kolmen alkion osajonot:

$$123, 132, 213, 231, 312, 321 \quad 6 \text{ kpl}$$

Kahden alkion osajonot:

$$12, 21, 13, 31, 23, 32 \quad 6 \text{ kpl}$$

Yhden alkion osajonot:

$$1, 2, 3 \quad 3 \text{ kpl}$$

5.4. Kombinatoriikan perusongelmien ratkaiseminen

Permutaatio

Olkoon S äärellinen joukko, jonka (erilaisten) alkioden lukumäärä on

$$n = n(S)$$

Kutsumme joukon S kaikkien alkioden jonoja joukon S alkioden **permutaatioiksi**. Joukon S alkioden kaikkien mahdollisten permutaatioiden lukumäärä on

$$n!$$

jossa

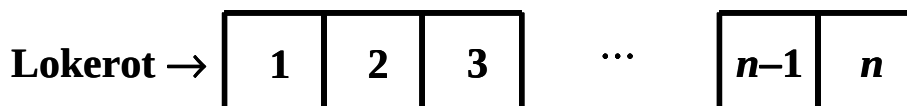
$$n! = n \times (n-1) \times \dots \times 2 \times 1$$

on ns. *n-kertoma*.

Perustelu:

Käytämme permutaatioiden lukumäärän kaavan johdossa apuna ns. **lokeromallia**.

Olkoon joukon S (erilaisten) alkioiden lukumäärä n ja oletetaan, että käytettävissämme on *lokerikko*, jossa on n lokeroa:



Asetetaan joukon S alkiot lokerikkoon yksi kerrallaan niin, että jokaiseen lokeroon tulee täsmälleen yksi alkio. Lokeroiden täyttäminen voidaan tehdä *vaiheittain*.

Vaiheessa $k = 1, 2, \dots, n$:

- (i) Lokeroista on täytetty $(k - 1)$ kpl.
- (ii) Joukossa S on jäljellä $(n - k + 1)$ alkioita.
- (iii) Suoritetaan operaatio

”Valitaan joukon S jäljellä olevista alkioista yksi lokeroon k ”

Kohdan (iii) operaatio voidaan suorittaa $(n - k + 1)$:llä erilaisella tavalla:

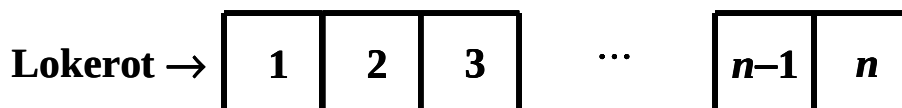
$k = 1$: Joukosta S voidaan valita alkio n :llä tavalla.

$k = 2$: Joukosta S voidaan valita alkio $(n - 1)$:llä tavalla.

$\dots$

$k = n - 1$: Joukosta S voidaan valita alkio 2:lla tavalla.

$k = n$: Joukosta S voidaan valita alkio 1:llä tavalla.



Operaatioiden lkm $\rightarrow$ $n \quad n-1 \quad n-2 \quad \dots \quad 2 \quad 1$

Tarkastellaan *yhdistettyä* operaatiota, jossa vaiheet $k = 1, 2, \dots, n$ käydään läpi peräkkäin. Kuinka monella erilaisella tavalla tämä *yhdistetty operaatio* voidaan suorittaa?

Koska jokainen kohdan (iii) operaatio voidaan suorittaa *edellisistä operaatioista riippumatta*, kombinatoriikan *kertolaskuperiaatteesta* seuraa, että lokeroiden täyttäminen voidaan tehdä

$$n \times (n - 1) \times \dots \times 2 \times 1 = n!$$

erilaisella tavalla.

■

Tulos ratkaisee *kombinatoriikan ensimmäisen perusongelman* (1a): Kuinka monella erilaisella tavalla $n:n$ (erilaisen) alkion joukon S alkiot voidaan järjestää *jonoon*?

n -kertoma

Permutaatioiden lukumäärän antava **n -kertoma** voidaan laskea seuraavalla *palautuskaavalla*:

$$n! = n \times (n-1)! , n = 1, 2, \dots$$

Määritellään:

$$0! = 1$$

Palautuskaavasta:

$$1! = 1 \times 0! = 1 \times 1 = 1$$

$$2! = 2 \times 1! = 2 \times 1 = 2$$

$$3! = 3 \times 2! = 3 \times 2 \times 1 = 6$$

$$4! = 4 \times 3! = 4 \times 3 \times 2 \times 1 = 24$$

$$5! = 5 \times 4! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$

...

Variaatio

Olkoon S äärellinen joukko, jonka (erilaisten) alkioiden lukumäärä on

$$n = n(S)$$

Kutsumme joukon S $k:n$ *alkion osajonoja* joukon S alkioiden **k -permutaatioiksi** eli **variaatioiksi**. Joukon S alkioiden kaikkien mahdollisten k -permutaatioiden lukumäärä on

$$P(n, k) = \frac{n!}{(n-k)!} = n \times (n-1) \times \dots \times (n-k+1)$$

Perustelu:

Olkoon joukon S (erilaisten) alkioiden lukumäärä n .

Joukon S kaikkien alkioiden permutaatioiden lukumäärää koskevasta todistuksesta nähdään, että n :stä alkioista voidaan valita k alkioita k :hon *ensimmäiseen* lokeroon

$$n \times (n-1) \times \dots \times (n-k+1)$$

erilaisella tavalla. Laventamalla saadaan

$$n \times (n-1) \times \dots \times (n-k+1) = \frac{n \times (n-1) \times \dots \times (n-k+1) \times (n-k)!}{(n-k)!} = \frac{n!}{(n-k)!}$$

■

Tulos ratkaisee *kombinatoriikan toisen perusongelman* (1b): Kuinka monella erilaisella tavalla $n:n$ (erilaisten) alkion joukon S alkioista voidaan muodostaa $k:n$ *alkion osajono*?

Jos

$$k = n$$

niin kombinatoriikan perusongelma (1b) kutistuu perusongelmaksi (1a). Koska olemme sopineet, että $0! = 1$, niin

$$P(n, n) = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$$

Kombinaatio

Olkoon S äärellinen joukko, jonka (erilaisten) alkoiden lukumäärä on

$$n = n(S)$$

Kutsumme joukon S k :n alkion osajoukkoja joukon S alkoiden k alkioita sisältäviksi **kombinaatioiksi**.

Joukon S alkoiden kaikkien mahdollisten k alkioita sisältävien kombinaatioiden lukumäärä on

$$C(n, k) = \binom{n}{k}$$

jossa

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

on ns. *binomikerroin*.

Perustelu:

Olkoon joukon S (erilaisten) alkoiden lukumäärä n .

Perustelemme kombinaatioiden lukumäärää koskevan kaavan *määrämällä* joukon S alkoiden k alkioita sisältävien *permutaatioiden lukumäärän kahdella erilaisella tavalla* ja merkitsemällä tulokset yhtä suuriksi. Saamme näin yhtälön, josta (toistaiseksi tuntematon) kombinaatioiden lukumäärä $C(n, k)$ voidaan ratkaista.

Joukon S , jossa on n alkioita, k -permutaatioiden lukumäärä on edellä esitetyn mukaan

$$P(n, k) = \frac{n!}{(n-k)!} = n \times (n-1) \times \dots \times (n-k+1)$$

Toisaalta joukon S alkoiden k -permutaatio voidaan muodostaa *kahdessa vaiheessa*:

(1) Valitaan joukon S alkioista k alkioita sisältävä osajoukko. Tämä operaatio voidaan tehdä

$$C(n, k)$$

erilaisella tavalla, jossa $C(n, k)$ on toistaiseksi tuntematon joukon S alkoiden k alkioita sisältävien permutaatioiden lukumäärä.

(2) Järjestetään valitun osajoukon k alkioita jonoon. Tämä operaatio voidaan tehdä $k!$

erilaisella tavalla.

Koska operaatiot (1) ja (2) voidaan suorittaa *toisistaan riippumatta*, niin kombinatoriikan *kertolaskuperiaatteesta* seuraa, että joukon S alkioden k alkioita sisältävien *permutaatioiden* lukumäärä on

$$P(n, k) = C(n, k) \times k!$$

Olemme siten saaneet yhtälön

$$P(n, k) = \frac{n!}{(n-k)!} = C(n, k) \times k!$$

josta $C(n, k)$ ratkaisemalla saadaan haluttu tulos:

$$C(n, k) = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

■

Tulos ratkaisee *kombinatoriikan kolmannen perusongelman* (2): Kuinka monella erilaisella tavalla $n:n$ (erilaisen) alkion joukon S alkioista voidaan muodostaa $k:n$ alkion *osajoukko*?

Kerrointa

$$C(n, k) = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

kutsutaan **binomikertoimeksi** ja se luetaan ” n yli $k:n$ ”. Koska olemme sopineet, että $0! = 1$, niin

$$\binom{n}{0} = \frac{n!}{0!n!} = 1 = \frac{n!}{n!0!} = \binom{n}{n}$$

Permutaatiot, variaatiot, kombinaatiot

Joukon alkioden *permutaatioissa* ja *variaatioissa* alkioden *järjestys on merkityksellinen*, mutta sen sijaan joukon alkioden *kombinaatioissa* alkioden *järjestys ei ole merkityksellinen*.

Esimerkki 1. Permutaatiot, variaatiot, kombinaatiot.

Ongelma 1:

Kuinka monta erilaista 3-numeroista kokonaislukua voidaan muodostaa numeroista

0, 1, 2, 3, 4, 5, 6, 7, 8, 9?

kun lukuja muodostettaessa ”etunollat” merkitään näkyviin.

Esimerkkejä:

$$5 = 005$$

$$19 = 019$$

Kaikki näin saatavat 3-numeroiset kokonaisluvut ovat muotoa

xyz

olevia *jonoja*, joissa numerot x, y ja z valitaan joukosta

$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Numeroiden x, y ja z valinta jonoon xyz voidaan tehdä *kahdella erilaisella tavalla*:

(i) Aikaisemmin valitun numeron *saa* valita uudelleen.

(ii) Aikaisemmin valittua numeroa *ei saa* valita uudelleen.

Tarkastellaan ensin tapausta:

(i) Aikaisemmin valitun numeron *saa* valita uudelleen.

Käytetään apuna *lokeromallia*: Kokonaisluku xyz muodostuu kolmesta lokerosta, joista jokainen voidaan täyttää *toisistaan riippumatta* 10:llä erilaisella objektilla.

Kertolaskuperiaatteen mukaan lokerot xyz voidaan täyttää

$$10 \times 10 \times 10 = 1000$$

erilaisella tavalla. Siten erilaisia 3-numeroisia lukuja, *joissa saa olla samoja numeroita*, on 1000 kpl. Tulos on tietysti sopusoinnussa sen kanssa, että kokonaislukujen

$$000, 001, 002, \dots, 010, 011, 012, \dots, 100, 101, 102, \dots, 999$$

lukumäärä on 1000.

Tarkastellaan seuraavaksi tapausta:

(ii) Aikaisemmin valittua numeroa *ei saa* valita uudelleen.

Käytetään apuna *lokeromallia*: Kokonaisluku xyz muodostuu kolmesta lokerosta, jotka voidaan täyttää vaiheittain seuraavalla tavalla:

1. lokero x voidaan täyttää 10:llä erilaisella objektilla.

2. lokero y voidaan täyttää *vaiheesta (1) riippumatta* 9:llä erilaisella objektilla, koska 1 objektista on käytetty.

3. lokero z voidaan täyttää *vaiheesta (2) riippumatta* 8:lla erilaisella objektilla, koska 2 objektista on käytetty.

Kertolaskuperiaatteen mukaan lokerot xyz voidaan täyttää

$$10 \times 9 \times 8 = 720$$

erilaisella tavalla. Siten erilaisia 3-numeroisia lukuja, *joissa sama numero ei saa esiintyä kuin kerran*, on 720 kpl.

Huomaa, että sama tulos saadaan huomaamalla, että tapauksessa (ii) on itse asiassa määrättävä joukon $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ 3-permutaatioiden lukumäärä:

3-permutaatioiden lukumääräksi saadaan

$$P(10, 3) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 10 \cdot 9 \cdot 8 = 720$$

mikä tietysti yhtyy aikaisemmin saatuun tulokseen.

Ongelma 2:

Kuinka monta erilaista 3:n alkion *osajoukkoa* voidaan valita numeroiden 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 muodostamasta joukosta

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

Ratkaisun antaa binomikerroin $C(10,3)$:

$$C(10,3) = \frac{10!}{3!(10-3)!} = \frac{10!}{3!7!} = \frac{10 \cdot 9 \cdot 8}{3 \cdot 2} = 120$$

Siten joukosta

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

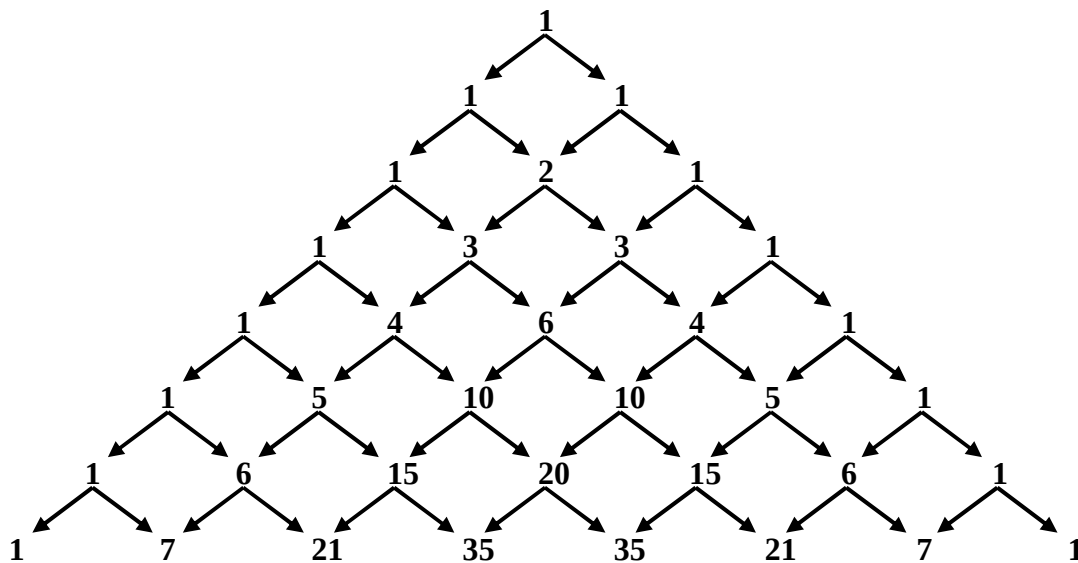
voidaan valita 3:n alkion *osajoukko* 120:llä erilaisella tavalla.

Huomaa asetettujen ehtojen vaikutus:

- (i) Numeroista 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 voidaan muodostaa 1000 kpl 3-numeroisia lukuja, joissa sama numero saa esiintyä useamman kerran.
- (ii) Joukosta $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ voidaan muodostaa 720 kpl 3:n numeron osajonoja.
- (iii) Joukosta $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ voidaan muodostaa 120 kpl 3:n numeron osajoukkoja.

Pascalin kolmio

Binomikertoimet saadaan ns. *Pascalin kolmiosta*. Alla on annettu Pascalin kolmion 8 ensimmäistä riviä.



Lukuun ottamatta kolmion reunoilla olevia ykkösiä jokainen kolmion luvuista on saatu laskemalla yhteen kaksi edeltävän rivin lukua nuolten suuntaan.

Pascalin kolmio ja binomikertoimet

Pascalin kolmion $(n+1)$. rivillä ovat binomikertoimet

$$\binom{n}{k}, k = 0, 1, 2, \dots, n$$

Pascalin kolmion *muodostamissääntö* voidaan ilmaista binomikertoimien avulla seuraavassa muodossa:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

Kaavan mukaan Pascalin kolmion n . rivin k . luku saadaan laskemalla yhteen $(n-1)$. rivin $(k-1)$. ja k . luku.

Perustelu:

$$\begin{aligned} \binom{n-1}{k-1} + \binom{n-1}{k} &= \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} + \frac{(n-1)!}{k!((n-1)-k)!} \\ &= \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-k-1)!} \\ &= \frac{k(n-1)!}{k!(n-k)!} + \frac{(n-k)(n-1)!}{k!(n-k)!} \\ &= \frac{(k+(n-k))(n-1)!}{k!(n-k)!} \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k} \end{aligned}$$

■

Pascalin kolmio on *symmetrinen* kolmion rivien keskikohdan suhteen, mikä voidaan ilmaista binomikertoimien avulla seuraavassa muodossa:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$$

Perustelu:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!(n-(n-k))!} = \binom{n}{n-k}$$

■

Binomikaava

Binomikaavan mukaan *binomin*

$$x + y$$

n . *potenssi* voidaan esittää muodossa

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

Perustelu:

Kun binomi

$$x + y$$

korotetaan potenssiin n , saadaan summalauseke, jonka termit ovat muotoa

$$x^{n-k}y^k, k = 0, 1, 2, \dots, n$$

Yhdistetään sellaiset termit, joissa esiintyy sama x :n potenssi ja järjestetään näin saadut termit x :n alenevien potenssien mukaiseen järjestykseen. Yhdistämisen tuloksena saadaan $(n+1)$ termiä sisältävä summalauseke, jonka $(k+1)$. termi on muotoa

$$D(n, k)x^{n-k}y^k, k = 0, 1, 2, \dots, n$$

ja jossa $D(n, k)$ on muotoa $x^{n-k}y^k$ olevien termien lukumäärä.

Tehtävänä on määrätä $D(n, k)$ eli se kuinka monella erilaisella tavalla muotoa $x^{n-k}y^k$ oleva termi syntyy, kun binomi

$$x + y$$

korotetaan potenssiin n .

Käytetään tehtävän ratkaisemisessa *lokeromallia*: Täytetään lokerikko, jossa on n lokeroa, tyyppiä x ja tyyppiä y olevilla objekteilla, kun tyyppiä x olevia objekteja on $(n-k)$ kpl ja tyyppiä y olevia objekteja on k kpl. Haluamme tietää, kuinka monella erilaisella tavalla tämä operaatio voidaan suorittaa.

Koska tyyppiä y olevien objektien paikat on määrätty sen jälkeen, kun tyyppiä x olevat objektit on sijoitettu lokerikkoon, riittää tarkastella sitä, kuinka monella erilaisella tavalla $(n-k)$ kpl tyyppiä x olevia objekteja voidaan sijoittaa lokerikkoon, jossa on n lokeroa.

Tämä tehtävä on selvästi ekvivalentti seuraavan tehtävän kanssa: Kuinka monella erilaisella tavalla joukosta, jossa on n alkioita, voidaan valita *osajoukko*, jossa on $(n-k)$ alkioita? Tämä on *kombinatoriikan perusongelma* (2) (ks. kappaletta **Kombinatoriikan perusongelmat**), joten ratkaisuksi saadaan *binomikerroin*

$$D(n, k) = \binom{n}{n-k} = \binom{n}{k}$$

■

Esimerkki 2. Binomikaava.

Binomikaavan mukaan binomin

$$x + y$$

4. potenssi voidaan esittää muodossa

$$\begin{aligned} (x+y)^4 &= \sum_{k=0}^4 \binom{4}{k} x^{4-k} y^k = \binom{4}{0} x^4 + \binom{4}{1} x^3 y + \binom{4}{2} x^2 y^2 + \binom{4}{3} x y^3 + \binom{4}{4} y^4 \\ &= x^4 + 4x^3 y + 6x^2 y^2 + 4x y^3 + y^4 \end{aligned}$$

Tulos on sopusoinnussa sen kanssa, että *Pascalin kolmion* 5. rivin luvut ovat

$$1, 4, 6, 4, 1$$

Tarkastellaan vielä sitä, miten tyyppiä $x^2 y^2$ olevat termit syntyvät: Kaikki mahdolliset muotoa $x^2 y^2$ olevat tulot ovat

$$\begin{array}{ccc} xxyy & xyxy & xyyx \\ yxxy & yxyx & yyxx \end{array}$$

Tuloja on siis 6 kappaletta.

Koska tässä $n = 4$ ja $k = 2$, binomikertoimen kaavasta saadaan tämän tuloksen kanssa yhtäpitävästi

$$C(4, 2) = \binom{4}{2} = \frac{4!}{2!2!} = \frac{4 \times 3 \times 2 \times 1}{2 \times 1 \times 2 \times 1} = 6$$

Äärellisen joukon osajoukkojen lukumäärä

Olkoon joukon S (erilaisten) alkioiden lukumäärä

$$n = n(S)$$

Tällöin joukon S osajoukkojen lukumäärä on

$$N = 2^n$$

Tässä lukumäärässä ovat mukana:

- (1) Tyhjä joukko $\emptyset$
- (2) Kaikki *yhden* alkion osajoukot
- (3) Kaikki *kahden* alkion osajoukot
- (4) Kaikki *kolmen* alkion osajoukot
- ...
- (n) Kaikki $(n - 1)$:n alkion osajoukot
- ($n + 1$) Joukko S

Perustelu:

Olkoon joukon S (erilaisten) alkioiden lukumäärä

$$n = n(S)$$

Joukolla S on k alkia sisältävien kombinaatioiden lukumäärää koskevan tuloksen mukaan

$$C(n, k) = \binom{n}{k}$$

osajoukkoa, jossa on k alkia, $k = 0, 1, 2, \dots, n$.

Joukon S osajoukkojen kokonaislukumäärä N saadaan laskemalla kaikki binomikertoimet

$$C(n, k), k = 0, 1, 2, \dots, n$$

yhteen:

$$N = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n-1} + \binom{n}{n}$$

Toisaalta *binomikaavan* mukaan, kun kaavaan sijoitetaan $x = y = 1$:

$$(1+1)^n = 2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n-1} + \binom{n}{n}$$

Yhdistämällä nämä tulokset saadaan joukon S kaikkien osajoukkojen lukumääräksi

$$N = 2^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n-1} + \binom{n}{n}$$

jossa binomikerroin

$$\binom{n}{k}$$

kertoo joukon S sellaisten osajoukkojen lukumäärän, joissa on k alkioita, $k = 0, 1, 2, \dots, n$.

■

Esimerkkejä lukumäärien laskemisesta

Esimerkki 3. Lotto.

Lotossa 1 ruudukko lototaan valitsemalla 7 numeroa 39:stä. Kuinka monta erilaista lotto-ruudukkoa on olemassa?

Tämä kysymys voidaan muotoilla myös seuraavalla tavalla: Kuinka monta erilaista 7:n alkion osajoukkoa voidaan valita 39:n erilaisen alkion joukosta?

Vastauksen antaa *kombinaatioiden lukumäärä* koskeva tulos:

$$\binom{39}{7} = \frac{39!}{7!32!} = 15\,380\,937$$

Esimerkki 4. Lotto.

Lotossa 1 ruudukko lototaan valitsemalla 7 numeroa 39:stä. Kuinka montako sellaista lotto-ruudukkoa on olemassa, joissa on *täsmälleen* 5 oikein?

5 oikein saadaan, jos on valittu 5 numeroa 7:n oikean numeron joukosta ja 2 numeroa 32:n väärän numeron joukosta.

5 numeroa voidaan valita

$$\binom{7}{5}$$

erilaisella tavalla 7:n oikean numeron joukosta ja 2 numeroa voidaan valita

$$\binom{32}{2}$$

erilaisella tavalla 32:n väärän numeron joukosta. Lisäksi valinnat voidaan tehdä *toisistaan riippumatta*. Siten *kertolaskuperiaatteesta* seuraa, että 5 oikein sisältävien rivien lukumäärä on

$$\binom{7}{5} \binom{32}{2} = \frac{7!}{5!2!} \times \frac{32!}{2!30!} = 21 \times 496 = 10\,416$$

Esimerkki 5. Pokeri.

Oletetaan, että pelaamme pokerin muotoa, jossa jokaiselle pelaajalle jaetaan 52:n kortin pakasta (ts. pakassa ei ole mukana jokereita) aluksi 5 korttia. Kuinka monta sellaista kättä on olemassa, joissa on 5 korttia?

Kombinaatioiden lukumäärää koskevan tuloksen mukaan vastauksena on

$$\binom{52}{5} = 2\,598\,960$$

kappaletta.

Esimerkki 6. Bridge.

Bridgessä on neljä pelaajaa ja jokaiselle pelaajalle jaetaan 52:n kortin pakasta (ts. pakassa ei ole mukana jokereita) 13 korttia. Kuinka monta sellaista kättä on olemassa, joissa on 13 korttia?

Kombinaatioiden lukumäärää koskevan tuloksen mukaan vastauksena on

$$\binom{52}{13} = 635\,013\,559\,600$$

kappaletta.

5.5. Multinomikerroin

Olkoon joukon S (erilaisten) alkioden lukumäärä

$$n = n(S)$$

Oletetaan, että positiiviset kokonaisluvut

$$n_i, i = 1, 2, \dots, k$$

toteuttavat ehdon

$$n_1 + n_2 + \dots + n_k = n$$

Ositetaan joukko S pistevieraisiin osajoukkoihin

$$A_i, i = 1, 2, \dots, k$$

niin, että joukossa A_i on

$$n_i = n(A_i)$$

alkiota. Kuinka monella erilaisella tavalla tällainen ositus voidaan tehdä?

Vastauksen antaa **multinomikerroin**

$$\binom{n}{n_1 \ n_2 \ \dots \ n_k} = \frac{n!}{n_1! n_2! \dots n_k!}$$

jossa siis

$$n_1 + n_2 + \dots + n_k = n$$

Multinomikertoimen kaava voidaan johtaa samanlaisella tekniikalla kuin binomikertoimen kaava.

Jos $k = 2$, multinomikertoimesta saadaan *binomikerroin*

$$\binom{n}{n_1 \ n_2} = \frac{n!}{n_1! n_2!} = \binom{n}{n_1} = \binom{n}{n_2}$$

jossa

$$n_1 + n_2 = n$$

Multinomikertoimet ovat kertoimina *multinomin*

$$x_1 + x_2 + \dots + x_k$$

n . *potenssin*

$$(x_1 + x_2 + \dots + x_k)^n$$

kehityskaavan muotoa

$$x_1^{n_1} x_2^{n_2} \dots x_k^{n_k}$$

olevissa termeissä, joissa siis potensseja $n_1, n_2, \dots, n_k$ sitoo ehto

$$n_1 + n_2 + \dots + n_k = n$$

Esimerkki. Multinomikerroin.

Sanassa *kasa* on 4 kirjainta, joiden joukossa on 3 *erilaista* kirjainta:

k	1 kpl
a	2 kpl
s	1 kpl

Kuinka monta *erilaista* neljän kirjaimen mittaista ”sanaa” voidaan muodostaa permutoimalla sanan *kasa* kirjaimia k, a, s ja a ?

Erilaisten sanojen lukumäärän antaa *multinomikerroin*

$$\binom{4}{1 \ 2 \ 1} = \frac{4!}{1!2!1!} = 12$$

Tässä tapauksessa erilaiset sanat on helppo *luetella*:

a -alkuiset sanat:

aaks aask akas aksa asak aska

k -alkuiset sanat:

kaas kasa ksaa

s -alkuiset sanat

saak saka skaa

Sanoja on todellakin 12 kpl kuten edellä todettiin.

6. Todennäköisyyden aksioomat

- 6.1. Todennäköisyys äärellisissä otosavaruuksissa
- 6.2. Klassinen todennäköisyys, suhteellinen frekvenssi ja ehdollinen todennäköisyys todennäköisyyksinä
- 6.3. Todennäköisyys mielivaltaisissa otosavaruuksissa

Tarkastelemme tässä luvussa todennäköisyyslaskentaa **aksiomaattisena matemaattisena teoriana**.

Käsittely on jaettu kahteen osaan:

- **Äärelliset otosavaruudet.**
- **Mielivaltaiset (äärettömät) otosavaruudet.**

Todennäköisyyden aksioomien olennaisena sisältönä on se, että **todennäköisyys on mitta matemaattisen mittateorian tarkoittamassa mielessä**.

Näytämme myös sen, että **klassista todennäköisyyttä ja suhteellista frekvenssiä (empiiristä todennäköisyyttä)** voidaan pitää sopivasti määriteltynä todennäköisyyden käsitteen *erikoistapauksina* ja sen, että **ehdollinen todennäköisyys** on todennäköisyyden käsitteen *laajennus*.

Avainsanat:

Additiivinen mitta, Boolean algebra, Ehdollinen todennäköisyys, Empiirinen todennäköisyys, Epämitallinen joukko, Jatkuvuusaksiooma, Joukko, Klassinen todennäköisyys, Kolmogorovin aksioomat, Komplementti, Mahdoton tapahtuma, Mitallinen joukko, Mitta, Mitallisuus, Normeerattu mitta, Osajoukko, Otosavaruus, Perusjoukko, Positiivinen mitta, Riippumattomuus, σ -algebra, Suhteellinen frekvenssi, Tapahtuma, Todennäköisyyden aksioomat, Todennäköisyys, Todennäköisyyskenttä, Todennäköisyysmitta, Toisensa poissulkevuus, Yhdistetty tapahtuma, Yhteenlaskusääntö, Yleinen yhteenlaskusääntö, Äärellinen todennäköisyyskenttä, Äärellisen otosavaruuden aksioomat

6.1. Todennäköisyys äärellisissä otosavaruuksissa

Tarkastelemme ensin todennäköisyyden määrittelemistä **äärellisissä otosavaruuksissa**. Suuri osa tässä todennäköisyyslaskennan esityksessä sovellettavista todennäköisyyden laskusäännöistä voidaan todistaa *äärellisten otosavaruuksien aksioomista*.

Boolean algebrat

Olkoon S joukko ja jokin F joukon S osajoukkojen muodostama *perhe* eli

$$A \in F \Rightarrow A \subset S$$

Joukkoperhe F on **Boolean algebra**, jos seuraavat ehdot pätevät:

(i) Tyhjä joukko $\emptyset$ on joukkoperheen F alkio:

$$\emptyset \in F$$

(ii) Jos joukko A on joukkoperheen F alkio, niin myös sen *komplementti* A^c on joukkoperheen F alkio:

$$A \in F \Rightarrow A^c \in F$$

(iii) Jos joukot A ja B ovat joukkoperheen F alkioita, niin myös niiden *yhdiste* $A \cup B$ on joukkoperheen F alkio:

$$A \in F, B \in F \Rightarrow A \cup B \in F$$

Lause 1.

Olkoon F jokin perusjoukon S osajoukoille määritelty *Boolean algebra* ja oletetaan, että

$$A \in F, B \in F$$

Tällöin

$$(a) \quad S \in F$$

$$(b) \quad A \cap B \in F$$

$$(c) \quad A \setminus B \in F$$

$$(d) \quad B \setminus A \in F$$

Todistus:

Olkoon F jokin perusjoukon S osajoukoille määritelty Boolean algebra ja oletetaan, että

$$A \in F, B \in F$$

Huomaa, että *Boolean algebran aksioomista* seuraa, että

$$\emptyset \in F, A^c \in F, B^c \in F, A \cup B \in F$$

(a) Osoitetaan, että

$$S \in F$$

Todetaan ensin, että

$$S = \emptyset^c$$

Boolean algebran aksiooman (i) mukaan

$$\emptyset \in F$$

Boolean algebran aksiooman (ii) mukaan

$$\emptyset \in F \Rightarrow \emptyset^c = S \in F$$

(b) Osoitetaan, että $A \cap B \in F$.

Todetaan ensin, että *De Morganin lain* mukaan

$$A \cap B = (A^c \cup B^c)^c$$

Boolean algebran aksiooman (ii) mukaan

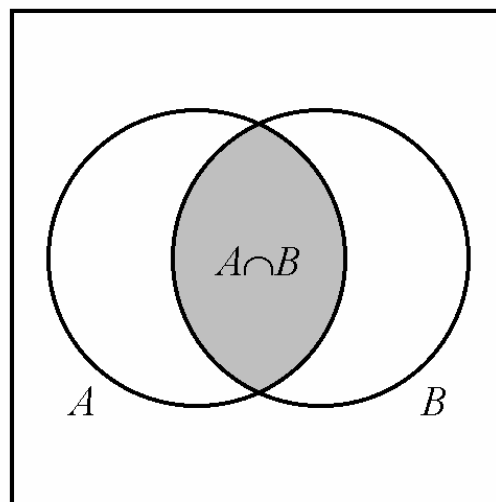
$$A \in F, B \in F \Rightarrow A^c \in F, B^c \in F$$

Boolean algebran aksiooman (iii) mukaan

$$A^c \in F, B^c \in F \Rightarrow A^c \cup B^c \in F$$

Boolean algebran aksiooman (ii) mukaan

$$A^c \cup B^c \in F \Rightarrow (A^c \cup B^c)^c = A \cap B \in F$$



S

(c) Osoitetaan, että $A \setminus B \in F$.

Todetaan ensin, että

$$A \setminus B = A \cap B^c$$

Boolean algebran aksiooman (ii) mukaan

$$B \in F \Rightarrow B^c \in F$$

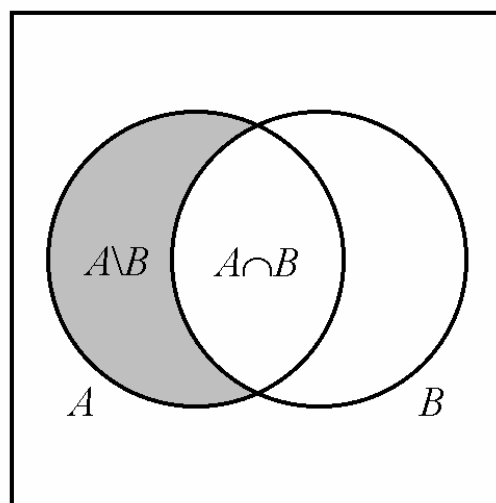
Kohdan (b) mukaan

$$A \in F, B^c \in F \Rightarrow A \cap B^c = A \setminus B \in F$$

(d) Sen osoittaminen, että

$$B \setminus A \in F$$

tapahtuu samalla tavalla kuin (c)-kohdan todistus.



S

■

Boolean algebran aksioomista ja lauseesta 1 seuraa, että Boolean algebrat ovat **suljettuja** tavanomaisten joukko-opin operaatioiden suhteen, kun operaatioita tehdään *äärellinen määrä*. Tällä tarkoitetaan siitä, että *äärellinen määrä* tavanomaisia joukko-opin operaatioita *ei vie Boolean algebran ulkopuolelle*: Jos Boolean algebran F joukkoihin sovelletaan *äärellinen määrä* tavanomaisia joukko-opin operaatioita *komplementti, yhdiste, leikkaus* ja *erotus*, niin tuloksena saatavat joukot kuuluvat edelleen Boolean algebraan F .

Esimerkki. Boolean algebra.

Olkoon S mielivaltainen joukko ja olkoon

$$A \subset S$$

mielivaltainen joukon S osajoukko. Tällöin joukkoperhe

$$\mathcal{F} = \{\emptyset, A, A^c, S\}$$

muodostaa Boolean algebran joukossa S , koska

$$(i) \quad \emptyset \in \mathcal{F}$$

$$(ii) \quad B \in \mathcal{F} \Rightarrow B^c \in \mathcal{F}$$

$$(iii) \quad B \in \mathcal{F}, C \in \mathcal{F} \Rightarrow B \cup C \in \mathcal{F}$$

jossa B ja C voivat olla mitkä tahansa kaksi joukoista

$$\emptyset, A, A^c, S$$

Todennäköisyyslaskennassa perusjoukkoa S kutsutaan **otosavaruudeksi** ja otosavaruuden S alkioita kutsutaan **alkeistapahtumiksi**. Jos siis s on *alkeistapahtuma*, niin

$$s \in S$$

Jos $\mathcal{F}$ on jokin otosavaruudessa S määritelty *Boolean algebra*, niin kutsumme Boolean algebraan $\mathcal{F}$ kuuluvia otosavaruuden S osajoukkoja **tapahtumiksi**. Jos siis joukko A on *tapahtuma*, niin

$$A \in \mathcal{F}$$

Olkoon $\mathcal{F}$ otosavaruudessa S määritelty *Boolean algebra*. Olkoot otosavaruuden S osajoukot A ja B *tapahtumia* eli oletetaan, että

$$A \in \mathcal{F}, B \in \mathcal{F}$$

Koska edellä esitetyn mukaan Boolean algebra $\mathcal{F}$ on *suljettu* tavanomaisten joukko-opin operaatioiden suhteen, niin myös joukot

$$A^c, B^c, A \cup B, A \cap B, A \setminus B, B \setminus A$$

ovat *tapahtumia*. Tällä tarkoitetaan siis sitä, että

$$A^c \in \mathcal{F}, B^c \in \mathcal{F}, A \cup B \in \mathcal{F}, A \cap B \in \mathcal{F}, A \setminus B \in \mathcal{F}, B \setminus A \in \mathcal{F}$$

Tämä merkitsee siis sitä, että otosavaruuden tapahtumista voidaan johtaa *uusia tapahtumia* soveltamalla niihin *äärellinen määrä* tavanomaisia joukko-opin operaatioita.

Olkoon S *äärellinen otosavaruus*, jossa on n alkioita ja olkoon

$$2^S = \{A \mid A \subset S\}$$

otosavaruuden S kaikkien osajoukkojen *perhe*. Joukkoperheessä 2^S on

$$2^n$$

alkioita; ks. luvun **Klassinen todennäköisyys ja kombinatoriikka** kappaletta **Kombinatoriikan perusongelmien ratkaiseminen**. Otosavaruuden S kaikkien osajoukkojen perhe 2^S muodostaa *triviaalin Boolean algebran* joukossa S .

Todennäköisyyden aksioomat äärellisissä otosavaruuksissa

Olkoon S äärellinen joukko ja $\mathcal{F}$ jokin joukon S osajoukkojen muodostama *Boolean algebra*. Olkoon lisäksi $\Pr$ *joukkofunktio*, joka liittää jokaiseen Boolean algebraan $\mathcal{F}$ kuuluvaan joukon S osajoukkoon A reaalikuvun eli

$$A \in \mathcal{F} \Rightarrow A \subset S \Rightarrow \Pr(A) \in \mathbb{R}$$

Joukkofunktio $\Pr$ on *äärellisen otosavaruuden todennäköisyysmitta*, jos

- (i) $\Pr(S) = 1$
- (ii) $0 \leq \Pr(A) \leq 1$ kaikille $A \in \mathcal{F}$
- (iii) $A \in \mathcal{F}, B \in \mathcal{F}, A \cap B = \emptyset \Rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$

Äärellisen otosavaruuden todennäköisyyden aksioomien (i)-(iii) mukaan *todennäköisyys* $\Pr$ on *positiivinen, äärellisesti additiivinen ja normeerattu mitta*.

Aksioomat (i) ja (ii), *normeeraus ja positiivisuus*:

$$A \subset S \Rightarrow 0 \leq \Pr(A) \leq \Pr(S) = 1$$

Aksiooma (iii), *äärellinen additiivisuus*:

$$A \cap B = \emptyset \Rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$$

Aksiooma (iii) on *yhteenlaskusääntö toisensa poissulkeville tapahtumille*.

Aksioomien (i)-(iii) olennaisena sisältönä on siis se, että *todennäköisyys on mitta* matematiikan tarkoittamassa mielessä. Todennäköisyyslaskentaa voidaankin pitää matemaattisen *mittateorian* erikoistuneena osana. Aksioomien (i)-(iii) mukaan todennäköisyysmitalla on samat ominaisuudet kuin *pinta-alamitalla* paitsi, että todennäköisyysmitta on *normeerattu* niin, että sen ylärajana on 1.

Äärellisen otosavaruuden tapahtumista voidaan muodostaa *uusia tapahtumia* soveltamalla niihin *Boolean algebran aksioomia* ja niistä johdettuja joukko-opin laskusääntöjä. Uusien tapahtumien *todennäköisyydet* saadaan soveltamalla niihin *äärellisen otosavaruuden todennäköisyyden aksioomia* ja niistä johdettuja *todennäköisyyslaskennan laskusääntöjä*.

Alkeistodennäköisyyslaskennan laskusääntöjen todistaminen

Todistamme alla seuraavat todennäköisyyslaskennan laskusäännöt lähtien äärellisten otosavaruuksien todennäköisyyden aksioomista:

- (i) **Mahdottoman tapahtuman** todennäköisyys.
- (ii) **Komplementtitapahtuman** todennäköisyys.
- (iii) **Osajoukon** todennäköisyys.
- (iv) **Yleinen yhteenlaskusääntö**.

Lause 2.

Olkoon S äärellinen otosavaruus ja olkoon $\mathcal{F}$ otosavaruuden S osajoukoille määritelty Boolean algebra. Tällöin *mahdottoman tapahtuman* $\emptyset$ todennäköisyys on nolla:

$$\Pr(\emptyset) = 0$$

Todistus:

Olkoon $\mathcal{F}$ äärellisen otosavaruuden S osajoukoille määritelty Boolean algebra ja olkoon $\emptyset$ mahdoton tapahtuma.

Boolean algebran aksiooman (i) mukaan

$$\emptyset \in \mathcal{F}$$

Lauseessa 1 on todistettu, että

$$S \in \mathcal{F}$$

Joukot $\emptyset$ ja S muodostavat otosavaruuden S osituksen eli

$$\emptyset \cup S = S$$

$$\emptyset \cap S = \emptyset$$

Todennäköisyyden aksioomien (i) ja (iii) mukaan

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

mikä ei voi olla totta, ellei

$$\Pr(\emptyset) = 0$$

■

Lause 3.

Olkoon S äärellinen otosavaruus ja olkoon $\mathcal{F}$ otosavaruuden S osajoukoille määritelty Boolean algebra. Oletetaan, että

$$A \in \mathcal{F}$$

Tällöin tapahtuman A komplementille A^c pätee:

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

Todistus:

Olkoon $\mathcal{F}$ äärellisen otosavaruuden S osajoukoille määritelty Boolean algebra ja olkoon

$$A \in \mathcal{F}$$

Boolean algebran aksiooman (ii) mukaan

$$A^c \in \mathcal{F}$$

Joukot A ja A^c muodostavat otosavaruuden S osituksen eli

$$A \cup A^c = S$$

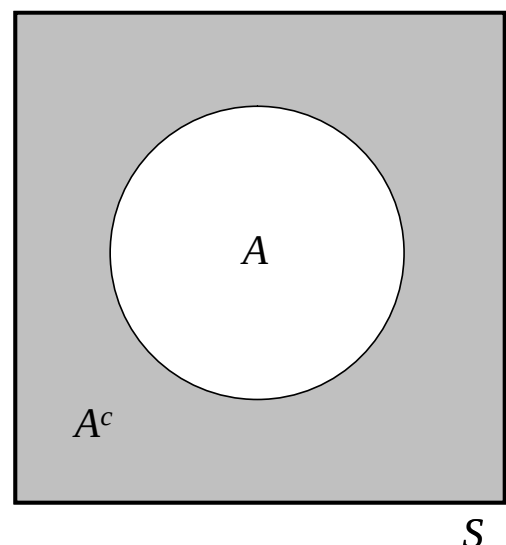
$$A \cap A^c = \emptyset$$

Todennäköisyyden aksioomien (i) ja (iii) mukaan

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

joten

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



■

Lause 4.

Olkoon S äärellinen otosavaruus ja olkoon $\mathcal{F}$ otosavaruuden S osajoukoille määritelty Boolean algebra. Oletetaan, että

$$A \in \mathcal{F}, B \in \mathcal{F}$$

ja

$$B \subset A$$

Tällöin

$$\Pr(B) \leq \Pr(A)$$

Todistus:

Olkoon $\mathcal{F}$ äärellisen otosavaruuden S osajoukoille määritelty Boolean algebra ja olkoon

$$A \in \mathcal{F}, B \in \mathcal{F}$$

Oletetaan lisäksi, että

$$B \subset A$$

Lauseessa 1 on todistettu, että

$$A \in \mathcal{F}, B \in \mathcal{F} \Rightarrow A \setminus B = A \cap B^c \in \mathcal{F}$$

Koska oletimme, että

$$B \subset A$$

niin joukot B ja $A \setminus B$ muodostavat joukon A osituksen:

$$B \cup (A \setminus B) = A$$

$$B \cap (A \setminus B) = \emptyset$$

Todennäköisyyden aksioomien (iii) ja (ii) mukaan

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

■

Lause 5.

Olkoon S äärellinen otosavaruus ja olkoon $\mathcal{F}$ otosavaruuden S osajoukoille määritelty Boolean algebra. Oletetaan, että

$$A \in \mathcal{F}, B \in \mathcal{F}$$

Tällöin

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

Todistus:

Olkoon $\mathcal{F}$ äärellisen otosavaruuden S osajoukoille määritelty Boolean algebra ja olkoon

$$A \in \mathcal{F}, B \in \mathcal{F}$$

Boolean algebran aksioomien ja lauseen 1 mukaan

$$A \in \mathcal{F}, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}, A \cap B \in \mathcal{F}, A \setminus B \in \mathcal{F}, B \setminus A \in \mathcal{F}$$

Joukot A ja $B \setminus A$ muodostavat joukon $A \cup B$ osituksen:

$$A \cup B = A \cup (B \setminus A)$$

$$A \cap (B \setminus A) = \emptyset$$

Ks. viereistä kuvaa.

Todennäköisyyden aksiooman (iii) mukaan

$$(1) \quad \Pr(A \cup B) = \Pr(A) + \Pr(B \setminus A)$$

Joukot $A \cap B$ ja $B \setminus A$ muodostavat joukon B osituksen:

$$B = (A \cap B) \cup (B \setminus A)$$

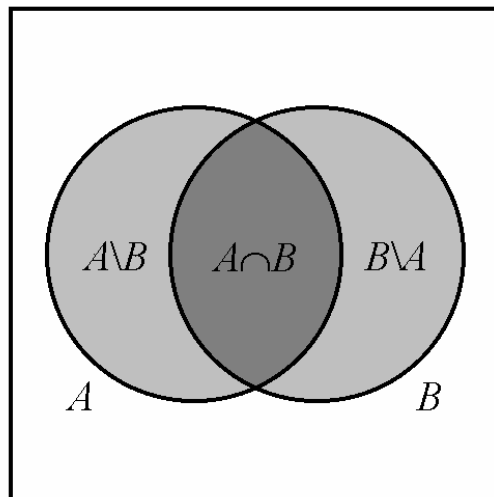
$$(A \cap B) \cap (B \setminus A) = \emptyset$$

Todennäköisyyden aksiooman (iii) mukaan

$$(2) \quad \Pr(B) = \Pr(A \cap B) + \Pr(B \setminus A)$$

Ratkaisemalla $\Pr(B \setminus A)$ yhtälöstä (2) ja sijoittamalla ratkaisu yhtälöön (1) saadaan väite:

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



S

■

Äärellinen todennäköisyyskenttä

Kolmikko

$$(S, \mathcal{F}, \Pr)$$

muodostaa **äärellisen todennäköisyyskentän**, jos S on äärellinen otosavaruus, $\mathcal{F}$ on otosavaruudessa S määritelty *Boolean algebra* ja $\Pr$ on Boolean algebrassa $\mathcal{F}$ määritelty *todennäköisyysmitta*.

Riippumattomuus ja riippumattomien tapahtumien tulosääntö

Sanomme, että tapahtumat A ja B ovat **riippumattomia**, jos ja vain jos *riippumattomien tapahtumien tulosääntö*

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

pätee.

6.2. Klassinen todennäköisyys, suhteellinen frekvenssi ja ehdollinen todennäköisyys todennäköisyyksinä

Klassinen todennäköisyys todennäköisyytenä

Olkoon

$$S = \{s_1, s_2, \dots, s_n\}$$

johonkin satunnaiskokeeseen liittyvä äärellinen *otosavaruus*, jossa on

$$n = n(S)$$

alkeistapahtumaa. Oletetaan lisäksi, että otosavaruuden S alkeistapahtumat ovat *symmetrisiä* eli yhtä todennäköisiä, jolloin

$$\Pr(s_i) = \frac{1}{n}, i = 1, 2, K, n$$

Olkoon

$$A = \{s_{i_1}, s_{i_2}, K, s_{i_k}\} \subset S$$

ko. satunnaiskokeeseen liittyvä *tapahtuma*, jossa on

$$k = n(A)$$

alkeistapahtumaa, joita sanotaan tapahtumalle A *suotuisiksi*. Määritellään tapahtuman A **klassinen todennäköisyys** $\Pr_c(A)$ kaavalla

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

jossa siis

$$k = n(A)$$

ja

$$n = n(S)$$

Perustelu:

Klassisen todennäköisyyden määritelmä voidaan perustella seuraavalla tavalla:

Olkoon

$$A = \{s_{i_1}, s_{i_2}, K, s_{i_k}\} \subset S = \{s_1, s_2, K, s_n\}$$

Tällöin

$$A = \{s_{i_1}\} \cup \{s_{i_2}\} \cup L \cup \{s_{i_k}\}$$

ja lisäksi alkeistapahtumat

$$s_{i_j}, j = 1, 2, K, k$$

ovat *pareittain toisensa poissulkevia*. Koska alkeistapahtumat $s_1, s_2, \dots, s_n$ on oletettu *symmetrisiksi*, todennäköisyyden aksioomasta (iii) seuraa:

$$\Pr(A) = \Pr(s_{i_1}) + \Pr(s_{i_2}) + L + \Pr(s_{i_k}) = \frac{1}{n} + L + \frac{1}{n} = \frac{k}{n} = \Pr_c(A)$$

1 4 2 43
k kpl

■

Klassinen todennäköisyys $\Pr_c(A)$ toteuttaa todennäköisyyden aksioomat (i)-(iii) äärellisille otosavaruuksille, joten **klassinen todennäköisyys on todennäköisyys**.

Perustelu:

Aksiooma (i):

Koska

$$n(S) = n$$

niin

$$\Pr_c(S) = \frac{n}{n} = 1$$

Aksiooma (ii):

Kaikille tapahtumille $A \subset S$ pätee

$$0 \leq n(A) = k \leq n = n(S)$$

joten

$$0 \leq \Pr_c(A) = \frac{k}{n} \leq 1$$

Aksiooma (iii):

Jos tapahtumat A ja B ovat *toisensa poissulkevia* eli, jos

$$A \cap B = \emptyset$$

niin

$$k_{A \cup B} = k_A + k_B$$

Tällöin

$$\Pr_c(A \cup B) = \frac{k_{A \cup B}}{n} = \frac{k_A + k_B}{n} = \frac{k_A}{n} + \frac{k_B}{n} = \Pr_c(A) + \Pr_c(B)$$

■

Siten klassinen todennäköisyys on äärellisissä otosavaruuksissa määritellyn todennäköisyyden *erikoistapaus*.

Suhteellinen frekvenssi todennäköisyytenä

Olkoon S johonkin satunnaiskokeeseen liittyvä äärellinen *otosavaruus* ja olkoon

$$A \subset S$$

ko. satunnaiskokeeseen liittyvä *tapahtuma*. Toistetaan satunnaiskoetta n kertaa ja olkoon f_A tapahtuman A *frekvenssi* eli *lukumäärä* koetoistojen joukossa. Tällöin

$$\frac{f_A}{n}$$

on tapahtuman A **suhteellinen frekvenssi**.

Käytetään tapahtuman A suhteelliselle frekvenssille f_A/n merkintää:

$$\Pr_f(A) = \frac{f_A}{n}$$

Suhteellinen frekvenssi $\Pr_f(A)$ toteuttaa todennäköisyyden aksioomat (i)-(iii) äärellisille otosavaruuksille, joten **suhteellinen frekvenssi on todennäköisyys**.

Perustelu:

Toistetaan satunnaiskoetta n kertaa.

Aksiooma (i):

Koska otosavaruus S on *varma tapahtuma*, se sattuu jokaisessa koetoistossa ja

$$f_S = n$$

Siten

$$\Pr_f(S) = \frac{f_S}{n} = \frac{n}{n} = 1$$

Aksiooma (ii):

Kaikille tapahtumille $A \subset S$ pätee

$$0 \leq f_A \leq n$$

Siten

$$0 \leq \Pr_f(A) = \frac{f_A}{n} \leq 1$$

Aksiooma (iii):

Jos tapahtumat A ja B ovat *toisensa poissulkevia* eli, jos

$$A \cap B = \emptyset$$

niin

$$f_{A \cup B} = f_A + f_B$$

Tällöin

$$\Pr_f(A \cup B) = \frac{f_{A \cup B}}{n} = \frac{f_A + f_B}{n} = \frac{f_A}{n} + \frac{f_B}{n} = \Pr_f(A) + \Pr_f(B)$$

■

Siten suhteellinen frekvenssi on äärellisissä otosavaruuksissa määritellyn todennäköisyyden *erikoistapaus*.

Empiirinen todennäköisyys

Jos tapahtuman A suhteellinen frekvenssi

$$\Pr_f(A) = \frac{f_A}{n}$$

lähestyy (jossakin mielessä) toistokokeiden lukumäärän n rajatta kasvaessa kiinteätä lukua p , sanotaan lukua p tapahtuman A **empiiriseksi todennäköisyydeksi**. Jos siis p on tapahtuman A empiirinen todennäköisyys, niin

$$\Pr_f(A) = \frac{f_A}{n} \rightarrow p, n \rightarrow \infty$$

jossa konvergenssi ei ole tavanomaista lukujonokonvergenssia; konvergenssikäsité täsmennetään luvussa **Stokastiikan konvergenssikäsitteet ja raja-arvolauseet**.

Todennäköisyyden frekvenssitulkinta

Oletetaan, että tapahtuman A todennäköisyys on p . Toistetaan sitä satunnaiskoetta, johon tapahtuma A liittyy, n kertaa. **Todennäköisyyden frekvenssitulkinnan** mukaan tapahtuman A suhteellinen frekvenssi

$$\Pr_f(A) = \frac{f_A}{n}$$

vaihtelee satunnaisesti koetoistosta toiseen, mutta saa keskimäärin tapahtuman todennäköisyyttä p lähellä olevia arvoja.

Ehdollinen todennäköisyys todennäköisyytenä

Olkoon S johonkin satunnaiskokeeseen liittyvä äärellinen otosavaruus. Olkoot

$$A \subset S, C \subset S$$

ko. satunnaiskokeeseen liittyviä *tapahtumia* ja olkoon

$$\Pr(C) \neq 0$$

Määritellään tapahtuman A **ehdollinen todennäköisyys** $\Pr(A|C)$ kaavalla

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

Ehdollinen todennäköisyys $\Pr(A|C)$ toteuttaa todennäköisyyden aksioomat (i)-(iii) äärellisille otosavaruuksille, joten **ehdollinen todennäköisyys on todennäköisyys**.

Perustelu:

Aksiooma (i):

Kaikille tapahtumille $C \subset S$ pätee

$$S \cap C = C$$

Siten

$$\Pr(S|C) = \frac{\Pr(S \cap C)}{\Pr(C)} = \frac{\Pr(C)}{\Pr(C)} = 1$$

Aksiooma (ii):

Koska kaikille tapahtumille $A \subset S$ ja $C \subset S$ pätee

$$A \cap C \subset C$$

niin

$$0 \leq \Pr(A \cap C) \leq \Pr(C)$$

Siten

$$0 \leq \Pr(A|C) = \frac{\Pr(A \cap C)}{\Pr(C)} \leq \frac{\Pr(C)}{\Pr(C)} = 1$$

Aksiooma (iii):

Kaikille tapahtumille $A \subset S$, $B \subset S$ ja $C \subset S$ pätee ns. *distribuutiolaki*

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$

Jos tapahtumat A ja B ovat *toisensa poissulkevia* eli, jos

$$A \cap B = \emptyset$$

niin kaikille $C \subset S$ pätee

$$(A \cap C) \cap (B \cap C) = \emptyset$$

Tällöin

$$\begin{aligned} \Pr(A \cup B | C) &= \frac{\Pr((A \cup B) \cap C)}{\Pr(C)} \\ &= \frac{\Pr((A \cap C) \cup (B \cap C))}{\Pr(C)} \\ &= \frac{\Pr(A \cap C)}{\Pr(C)} + \frac{\Pr(B \cap C)}{\Pr(C)} \\ &= \Pr(A | C) + \Pr(B | C) \end{aligned}$$

■

Siten ehdollinen todennäköisyys on todennäköisyyden käsitteen *laajennus*.

6.3. Todennäköisyys mielivaltaisissa otosavaruuksissa

Tarkastellaan todennäköisyyden määrittelemistä **mielivaltaisissa otosavaruuksissa**.

σ -algebrat

Olkoon S joukko ja jokin $\mathcal{F}$ joukon S osajoukkojen muodostama *perhe* eli

$$A \in \mathcal{F} \Rightarrow A \subset S$$

Joukkoperhe $\mathcal{F}$ on **σ -algebra**, jos seuraavat ehdot pätevät:

(i) Tyhjä joukko $\emptyset$ on joukkoperheen $\mathcal{F}$ alkio:

$$\emptyset \in \mathcal{F}$$

(ii) Jos joukko A on joukkoperheen $\mathcal{F}$ alkio, niin myös sen *komplementti* A^c on joukkoperheen $\mathcal{F}$ alkio:

$$A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$$

(iii) Jos joukot $A_1, A_2, A_3, \dots$ ovat joukkoperheen $\mathcal{F}$ alkioita, niin myös niiden *yhdiste* $\cup A_i$ on joukkoperheen $\mathcal{F}$ alkio:

$$A_1, A_2, A_3, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$$

Jos joukkoperhe $\mathcal{F}$ toteuttaa σ -algebran aksioomat, niin se toteuttaa myös *Boolean algebran aksioomat*. Siten **kaikki Boolean algebroiden laskusäännöt pätevät σ -algebroidille**. Erityisesti siis tämän luvun kappaleen **Todennäköisyys äärellisissä otosavaruuksissa** lausetta 1 vastaava lause pätee myös σ -algebroidille:

Lause 1.

Olkoon $\mathcal{F}$ jokin perusjoukon S osajoukoille määritelty σ -algebra ja oletetaan, että

$$A \in \mathcal{F}, B \in \mathcal{F}$$

Tällöin

$$(a) \quad S \in \mathcal{F}$$

$$(b) \quad A \cap B \in \mathcal{F}$$

$$(c) \quad A \setminus B \in \mathcal{F}$$

$$(d) \quad B \setminus A \in \mathcal{F}$$

Lauseen 1 (b)-kohta voidaan laajentaa koskemaan *numeroituvaa* määrää σ -algebran joukkoja:

Lause 2.

Olkoon $\mathcal{F}$ jokin perusjoukon S osajoukoille määritelty σ -algebra ja oletetaan, että

$$A_1, A_2, A_3, \dots \in \mathcal{F}$$

Tällöin joukkojen $A_1, A_2, A_3, \dots$ leikkaus $\cap A_i$ on σ -algebran $\mathcal{F}$ alkio:

$$A_1, A_2, A_3, \dots \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

Todistus:

Olkoon $\mathcal{F}$ jokin perusjoukon S osajoukoille määritelty σ -algebra ja oletetaan, että

$$A_1, A_2, A_3, \dots \in \mathcal{F}$$

Todetaan ensin, että *De Morganin lain yleistyksen* mukaan

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c \right)^c$$

σ -algebran aksiooman (ii) mukaan

$$A_1, A_2, A_3, \dots \in \mathcal{F} \Rightarrow A_1^c, A_2^c, A_3^c, \dots \in \mathcal{F}$$

σ -algebran aksiooman (iii) mukaan

$$A_1^c, A_2^c, A_3^c, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F}$$

σ -algebran aksiooman (ii) mukaan

$$\bigcup_{i=1}^{\infty} A_i^c \in \mathcal{F} \Rightarrow \left(\bigcup_{i=1}^{\infty} A_i^c \right)^c = \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$$

■

σ -algebran aksioomista ja lauseista 1 ja 2 seuraa, että σ -algebrat ovat **suljettuja** tavanomaisten joukko-opin operaatioiden suhteen, kun operaatioita tehdään *korkeintaan numeroituva määrä*. Tällä tarkoitetaan siitä, että *numeroituva määrä* tavanomaisia joukko-opin operaatioita *ei vie* σ -algebran *ulkopuolelle*: Jos σ -algebran F joukkoihin sovelletaan *korkeintaan numeroituva määrä* tavanomaisia joukko-opin operaatioita *komplementti, yhdiste, leikkaus ja erotus*, niin tuloksena saatavat joukot kuuluvat edelleen σ -algebraan F .

Todennäköisyyslaskennassa perusjoukkoa S kutsutaan **otosavaruudeksi** ja otosavaruuden S alkioita kutsutaan **alkeistapahtumiksi**. Jos siis s on *alkeistapahtuma*, niin

$$s \in S$$

Jos F on jokin otosavaruudessa S määritelty σ -algebra, niin kutsumme σ -algebraan F kuuluvia otosavaruuden S osajoukkoja **tapahtumiksi**. Jos siis joukko A on *tapahtuma*, niin

$$A \in F$$

Olkoon F otosavaruudessa S määritelty σ -algebra. Olkoot otosavaruuden S osajoukot

$$A_1, A_2, A_3, \dots$$

tapahtumia eli oletetaan, että

$$A_1, A_2, A_3, \dots \in F$$

Koska edellä esitetyn mukaan σ -algebra F on *suljettu* tavanomaisten joukko-opin operaatioiden suhteen, kun operaatioita tehdään *korkeintaan numeroituva määrä*, niin myös joukot

$$A_1^c, A_2^c, A_3^c, K, \bigcup_{i=1}^{\infty} A_i, \bigcap_{i=1}^{\infty} A_i$$

ovat *tapahtumia*. Tällä tarkoitetaan siis sitä, että

$$A_1^c, A_2^c, A_3^c, K \in F, \bigcup_{i=1}^{\infty} A_i \in F, \bigcap_{i=1}^{\infty} A_i \in F$$

Tämä merkitsee siis sitä, että otosavaruuden tapahtumista voidaan johtaa *uusia tapahtumia* soveltamalla niihin *korkeintaan numeroituva määrä* tavanomaisia joukko-opin operaatioita.

Kolmogorovin aksioomat

Olkoon S jokin joukko ja F jokin joukon S osajoukkojen muodostama σ -algebra. Olkoon lisäksi $\Pr$ *joukkofunktio*, joka liittyy jokaiseen σ -algebraan F kuuluvaan joukon S osajoukkoon A reaalikuvun eli

$$A \in F \Rightarrow A \subset S \Rightarrow \Pr(A) \in$$

Joukkofunktio $\Pr$ on **todennäköisyysmitta**, jos

- (i) $\Pr(S) = 1$
- (ii) $0 \leq \Pr(A) \leq 1$ kaikille $A \in F$
- (iii) $A_1, A_2, A_3, K \in F$ ja $A_i \cap A_j = \emptyset, i \neq j \Rightarrow \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$

Aksioomia (i)-(iii) kutsutaan **Kolmogorovin aksioomiksi todennäköisyydelle**.

Kolmogorovin aksioomien (i)-(iii) mukaan *todennäköisyys* $\Pr$ on *positiivinen*, *additiivinen* ja *normeerattu mitta*.

Aksioomat (i) ja (ii), *normeeraus* ja *positiivisuus*:

$$A \subset S \Rightarrow 0 \leq \Pr(A) \leq \Pr(S) = 1$$

Aksiooma (iii), *additiivisuus*:

$$A_1, A_2, A_3, K \in \mathcal{F} \text{ ja } A_i \cap A_j = \emptyset, i \neq j \Rightarrow \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$$

Aksiooma (iii) on *yhteenlaskusääntö numeroituvalle määrälle toisensa poissulkevia tapahtumia*.

Aksioomien (i)-(iii) olennaisena sisältönä on siis se, että *todennäköisyys on mitta* matematiikan tarkoittamassa mielessä. Todennäköisyyslaskentaa voidaankin pitää matemaattisen *mittateorian* erikoistuneena osana. Aksioomien (i)-(iii) mukaan todennäköisyysmitalla on samat ominaisuudet kuin *pinta-alamitalla* paitsi, että todennäköisyysmitta on *normeerattu* niin, että sen ylärajana on 1.

Otosavaruuden tapahtumista voidaan muodostaa *uusia tapahtumia* soveltamalla niihin σ -algebran aksioomia ja niistä johdettuja joukko-opin laskusääntöjä. Uusien tapahtumien *todennäköisyydet* saadaan soveltamalla niihin *Kolmogorovin aksioomia* ja niistä johdettuja *todennäköisyyslaskennan laskusääntöjä*.

Jos joukkofunktio $\Pr$ toteuttaa *Kolmogorovin aksioomat*, niin se toteuttaa *todennäköisyyden aksioomat äärellisille otosavaruuksille*. Siten **kaikki äärellisen otosavaruuden todennäköisyyden aksioomista johdetut laskusäännöt pätevät yleiselle todennäköisyysmitalle**. Erityisesti siis kappaleen **Todennäköisyys äärellisissä otosavaruuksissa** lauseet 2-5 äärellisen otosavaruuden todennäköisyysmitalle pätevät myös yleiselle todennäköisyysmitalle.

Todennäköisyyskenttä

Kolmikko

$$(S, \mathcal{F}, \Pr)$$

muodostaa **todennäköisyyskentän**, jos S on otosavaruus, $\mathcal{F}$ on otosavaruudessa S määritelty σ -algebra ja $\Pr$ on σ -algebrassa $\mathcal{F}$ määritelty *todennäköisyysmitta*.

Mitalliset ja epämitalliset joukot

On syytä huomata, että jos otosavaruus S on ääretön, sen kaikille osajoukoille *ei välttämättä voida määritellä todennäköisyyttä*. Niitä otosavaruuden S osajoukkoja, joille todennäköisyys voidaan määritellä kutsutaan *mitallisiksi* ja niitä, joille todennäköisyyttä ei voida määritellä kutsutaan *epämitallisiksi*. Voidaan osoittaa, että *otosavaruuden S mitalliset osajoukot muodostavat σ -algebran*.

Esimerkki.

Jos otosavaruutena S on *reaalilukujen joukko* , niin siinä voidaan määritellä todennäköisyysmitta, jolle esimerkiksi kaikki reaaliakselin *välit* sekä *avoimet* ja *suljetut joukot* ovat *mitallisia* ja kelpaavat siis tapahtumiksi. Voidaan kuitenkin osoittaa, että reaalilukujen joukossa on myös sellaisia osajoukkoja, jotka ovat *epämitallisia* ja eivät siis kelpaa tapahtumiksi.

Voidaan osoittaa, että *kaikki reaalilukujen joukon todennäköisyysmitan suhteen mitalliset osajoukot* saadaan tyyppiä

$$(-\infty, b] = \{x \in \mathbb{R} \mid -\infty < x \leq b\}$$

olevista puoliavoimista väleistä soveltamalla niihin *korkeintaan numeroituva määrä* komplementti-, yhdiste- ja leikkausoperaatiota.

Sanomme, että muotoa $(-\infty, b]$ olevat reaaliakselin välit **virittävät** reaalilukujen joukon mitallisten osajoukkojen muodostaman σ -algebran.

Tähän todennäköisyysmitan ominaisuuteen perustuu *kertymäfunktion* keskeinen asema *todennäköisyyslaskennassa* ja *matemaattisessa tilastotieteessä*; ks. lukua **Kertymäfunktio**.

Kolmogorovin aksioomien seurauksia

Todistamme seuraavassa kaksi myöhemmin tarvittavaa lausetta, joiden todistamiseen *eivät riitä* äärellisen otosavaruuden aksioomat todennäköisyydelle: Kummassakin todistuksessa joudutaan vetoamaan *Kolmogorovin aksioomaan* (iii).

Lause 3.

Olkoon $(S, F, \Pr)$ todennäköisyyskenttä ja olkoot

$$A_1, A_2, A_3, K \in F$$

(a) Jos

$$A_1 \subset A_2 \subset A_3 \subset L$$

niin

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} \Pr(A_i)$$

(b) Jos

$$A_1 \supset A_2 \supset A_3 \supset L$$

niin

$$\Pr\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{i \rightarrow \infty} \Pr(A_i)$$

Todistus:

(a) Olkoon F jokin perusjoukon S osajoukoille määritelty σ -algebra ja oletetaan, että

$$A_1, A_2, A_3, K \in F$$

ja

$$A_1 \subset A_2 \subset A_3 \subset L$$

Määritellään

$$B_0 = A_0 = \emptyset$$

$$B_1 = A_1$$

$$B_2 = A_2 \setminus A_1$$

$$B_3 = A_3 \setminus A_2$$

ja yleisesti

$$B_i = A_i \setminus A_{i-1} = A_i \cap A_{i-1}^c, i = 1, 2, 3, K$$

Joukot B_i , $i = 0, 1, 2, 3, \dots$ ovat *pareittain toisensa poissulkevia*, koska oletuksen mukaan

$$A_{i-1} \subset A_i, i = 1, 2, 3, \dots$$

Oletuksesta

$$A_{i-1} \subset A_i, i = 1, 2, 3, \dots$$

seuraa, että

$$\begin{aligned} \Pr(B_i) &= \Pr(A_i \setminus A_{i-1}) = \Pr(A_i) - \Pr(A_{i-1}) \\ i &= 1, 2, 3, K \end{aligned}$$

Lisäksi on selvää, että

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$

Siten

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \Pr\left(\bigcup_{i=1}^{\infty} B_i\right)$$

Koska joukot B_i , $i = 0, 1, 2, 3, \dots$ ovat pareittain toisensa poissulkevia, *Kolmogorovin aksioomasta* (iii) seuraa:

$$\Pr\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \Pr(B_i) = \lim_{n \rightarrow +\infty} \sum_{i=1}^n \Pr(B_i)$$

Sijoitetaan tähän

$$\Pr(B_i) = \Pr(A_i) - \Pr(A_{i-1}), i = 1, 2, 3, K$$

Tällöin saadaan

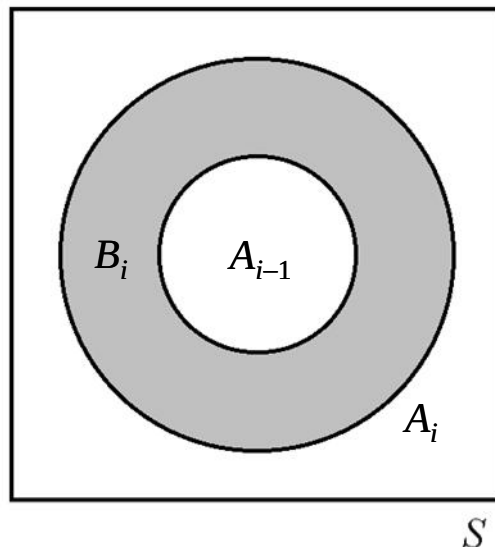
$$\begin{aligned} \lim_{n \rightarrow +\infty} \sum_{i=1}^n \Pr(B_i) &= \lim_{n \rightarrow +\infty} \{\Pr(B_1) + \Pr(B_2) + \Pr(B_3) + \\ &\quad \text{L} \\ &\quad + \Pr(B_{n-1}) + \Pr(B_n)\} \\ &= \lim_{n \rightarrow +\infty} \{\Pr(A_1) + [\Pr(A_2) - \Pr(A_1)] + [\Pr(A_3) - \Pr(A_2)] + \\ &\quad \text{L} \\ &\quad + [\Pr(A_{n-1}) - \Pr(A_{n-2})] + [\Pr(A_n) - \Pr(A_{n-1})]\} \\ &= \lim_{n \rightarrow +\infty} \Pr(A_n) \end{aligned}$$

Yhdistämällä edellä johdetut tulokset saadaan lopulta:

$$\Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \Pr\left(\bigcup_{i=1}^{\infty} B_i\right) = \lim_{n \rightarrow +\infty} \Pr(A_n)$$

- (b) Olkoon $\mathcal{F}$ jokin perusjoukon S osajoukoille määritelty σ -algebra ja oletetaan, että $A_1, A_2, A_3, K \in \mathcal{F}$

ja



$$A_1 \supset A_2 \supset A_3 \supset L$$

Määritellään

$$C_0 = S$$

$$C_1 = A_1^c$$

$$C_2 = A_2^c$$

$$C_3 = A_3^c$$

ja yleisesti

$$C_i = A_i^c, i = 1, 2, 3, K$$

Oletuksesta $A_{i-1} \supset A_i, i = 1, 2, 3, \dots$ seuraa:

$$C_{i-1} = A_{i-1}^c \subset A_i^c = C_i, i = 1, 2, 3, K$$

Sovelletaan joukkoihin $C_i, i = 1, 2, 3, \dots$ lauseen (a)-kohtaa:

$$\Pr\left(\bigcup_{i=1}^{\infty} C_i\right) = \lim_{n \rightarrow +\infty} \Pr(C_n)$$

Koska

$$C_i = A_i^c, i = 1, 2, 3, K$$

niin *komplementtitapahtuman todennäköisyyden kaavasta* seuraa, että

$$\Pr(C_i) = \Pr(A_i^c) = 1 - \Pr(A_i), i = 1, 2, 3, K$$

DeMorganin lain yleistyksen mukaan

$$\left(\bigcup_{i=1}^{\infty} C_i\right)^c = \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c = \bigcap_{i=1}^{\infty} A_i$$

Komplementtitapahtuman todennäköisyyden kaavan mukaan

$$\Pr\left[\left(\bigcup_{i=1}^{\infty} C_i\right)^c\right] = 1 - \Pr\left(\bigcup_{i=1}^{\infty} C_i\right)$$

Yhdistämällä edellä johdetut tulokset saadaan lopulta:

$$\begin{aligned} \Pr\left(\bigcap_{i=1}^{\infty} A_i\right) &= \Pr\left[\left(\bigcup_{i=1}^{\infty} C_i\right)^c\right] \\ &= 1 - \Pr\left(\bigcup_{i=1}^{\infty} C_i\right) \\ &= 1 - \lim_{n \rightarrow +\infty} \Pr(C_n) \\ &= 1 - \lim_{n \rightarrow +\infty} [1 - \Pr(A_n)] \\ &= \lim_{n \rightarrow +\infty} \Pr(A_n) \end{aligned}$$

■

Lause 4.

Olkoon

$$(S, \mathcal{F}, \Pr)$$

todennäköisyyskenttä ja olko

$$A_1, A_2, A_3, K \in \mathcal{F}$$

Jos

$$A_1 \supset A_2 \supset A_3 \supset L \rightarrow \emptyset$$

niin

$$\lim_{i \rightarrow \infty} \Pr(A_i) = 0$$

Todistus:

Olkoon $\mathcal{F}$ jokin perusjoukon S osajoukoille määritelty σ -algebra ja oletetaan, että

$$A_1, A_2, A_3, K \in \mathcal{F}$$

ja

$$A_1 \supset A_2 \supset A_3 \supset L \rightarrow \emptyset$$

Koska $A_1 \supset A_2 \supset A_3 \supset L \rightarrow \emptyset$, niin

$$\bigcap_{i=1}^{\infty} A_i = \emptyset$$

Siten

$$\Pr\left(\bigcap_{i=1}^{\infty} A_i\right) = \Pr(\emptyset) = 0$$

Lauseen 3 kohdan (b) mukaan

$$\Pr\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow +\infty} \Pr(A_n)$$

Yhdistämällä edellä johdetut tulokset saadaan lopulta:

$$\lim_{n \rightarrow +\infty} \Pr(A_n) = 0$$

■

Voidaan osoittaa, että Kolmogorovin aksiooma

$$(iii) \quad A_1, A_2, A_3, K \in \mathcal{F} \text{ ja } A_i \cap A_j = \emptyset, i \neq j \Rightarrow \Pr\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$$

on yhtäpitävä aksioomien

$$(iii)' \quad A \cap B = \emptyset \Rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B)$$

$$(iv) \quad A_1 \supset A_2 \supset A_3 \supset L \rightarrow \emptyset \Rightarrow \lim_{n \rightarrow +\infty} \Pr(A_n) = 0$$

kanssa.

Aksiooma (iii)' on tavanomainen *yhteenlaskusääntö kahdelle toisensa poissulkevalle tapahtumalle*. Aksioomaa (iv) kutsutaan tavallisesti todennäköisyyden **jatkuvuusaksioomaksi**.

7. Kokonaistodennäköisyyden ja Bayesin kaavat

7.1. Johdatteleva esimerkki

7.2. Ositukset

7.3. Kokonaistodennäköisyyden kaava

7.4. Bayesin kaava

7.5. Kokonaistodennäköisyyden kaavan ja Bayesin kaavan systeemiteoreettinen tulkinta

Tarkastelemme tässä luvussa kahta hyödyllistä todennäköisyyslaskennan kaavaa, joita kutsutaan **kokonaistodennäköisyyden kaavaksi** ja **Bayesin kaavaksi**. On syytä huomata, että kaavat liittyvät läheisesti toisiinsa.

Avainsanat:

Bayesin kaava, Ehdollinen todennäköisyys, Kokonaistodennäköisyyden kaava, Käänteistodennäköisyys, Ositus, Posteriori-todennäköisyys, Priori-todennäköisyys, Systeemiteoria, Verkkodiagrammi

7.1. Johdatteleva esimerkki

Ruuvitehtaalla on kaksi konetta A ja B, joilla tehdään samanlaisia ruuveja. A- ja B-koneen valmistamat ruuvit sekoitetaan keskenään ja pakataan laatikoihin. Koska A-kone toimii hitaammin, laatikoihin tulee A- ja B-koneiden valmistamia ruuveja *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* laatikollinen ruuveja tutkittavaksi ja poimitaan valitusta laatikosta *satunnaisesti* 1 ruuvi tutkittavaksi.

Kysymyksiä:

- (1) Mikä on todennäköisyys, että poimittu ruuvi on *viallinen*?
- (2a) Mikä on todennäköisyys, että ruuvin on valmistanut A-kone, jos ruuvi osoittautuu *vialliseksi*?
- (2b) Mikä on todennäköisyys, että ruuvin on valmistanut B-kone, jos ruuvi osoittautuu *vialliseksi*?

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*:

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

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

Seuraavia todennäköisyyksiä kysytään:

$$\Pr(V), \Pr(A|V), \Pr(B|V)$$

Tapahtumat A ja B muodostavat otosavaruuden S **osituksen**. Tällä tarkoitetaan seuraavaa:

- (i) A ja B ovat *epätyhjiä*:

$$A \neq \emptyset \text{ ja } B \neq \emptyset$$

- (ii) A ja B ovat *pistevieraita*:

$$A \cap B = \emptyset$$

- (iii) Joukkojen A ja B yhdisteenä saadaan *perusjoukko* S :

$$S = A \cup B$$

Otosavaruuden S ositus $S = A \cup B$ **indusoi osituksen** tapahtumaan V . Tällä tarkoitetaan seuraavaa:

- (i) Jos V on epätyhjä eli $V \neq \emptyset$, niin ainakin toinen joukoista $V \cap A$ ja $V \cap B$ on *epätyhjä*:

$$V \cap A \neq \emptyset \text{ tai } V \cap B \neq \emptyset$$

(ii) $V \cap A$ ja $V \cap B$ ovat pistevieraita:

$$(V \cap A) \cap (V \cap B) = \emptyset$$

koska $A \cap B = \emptyset$

(iii) Joukkojen $V \cap A$ ja $V \cap B$ yhdisteenä saadaan joukko V :

$$V = (V \cap A) \cup (V \cap B)$$

Toisensa poissulkevien tapahtumien yhteenlaskusäännön mukaan:

$$(1) \quad \Pr(V) = \Pr(V \cap A) + \Pr(V \cap B)$$

Yleisen tulosäännön mukaan:

$$(2) \quad \Pr(V \cap A) = \Pr(A) \Pr(V|A)$$

$$(3) \quad \Pr(V \cap B) = \Pr(B) \Pr(V|B)$$

Sijoittamalla lausekkeet (2) ja (3) kaavaan (1) saadaan todennäköisyydeksi, että satunnaisesti poimittu ruuvi on viallinen:

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

Todennäköisyys $\Pr(V)$ on määrätty **kokonaistodennäköisyyden kaavalla**.

Ehdollisen todennäköisyyden määritelmän perusteella

$$(4) \quad \Pr(A|V) = \frac{\Pr(V \cap A)}{\Pr(V)}$$

$$(5) \quad \Pr(B|V) = \frac{\Pr(V \cap B)}{\Pr(V)}$$

Yleisen tulosäännön mukaan

$$(6) \quad \Pr(V \cap A) = \Pr(V|A) \Pr(A)$$

$$(7) \quad \Pr(V \cap B) = \Pr(V|B) \Pr(B)$$

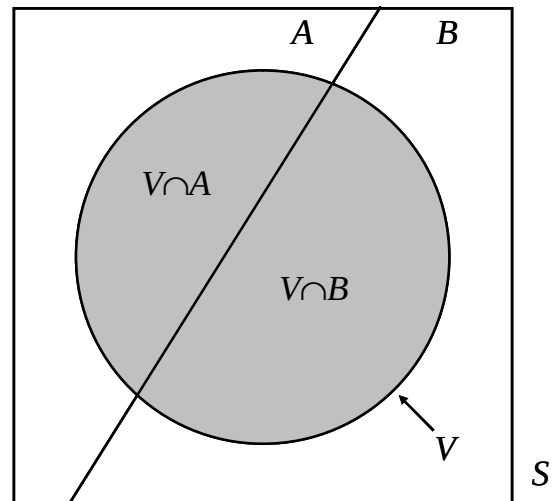
Edellä on todettu, että

$$(8) \quad \Pr(V) = \Pr(A) \Pr(V|A) + \Pr(B) \Pr(V|B)$$

Sijoittamalla lausekkeet (6) ja (8) kaavaan (4) saadaan **ehdolliseksi todennäköisyydeksi** $\Pr(A|V)$:

$$\begin{aligned} \Pr(A|V) &= \frac{\Pr(V \cap A)}{\Pr(V)} = \frac{\Pr(A) \Pr(V|A)}{\Pr(A) \Pr(V|A) + \Pr(B) \Pr(V|B)} \\ &= \frac{\frac{3}{8} \times 0.05}{\frac{3}{8} \times 0.05 + \frac{5}{8} \times 0.08} = \frac{3}{11} = 0.27 \end{aligned}$$

Ehdollinen todennäköisyys $\Pr(A|V)$ on määrätty **Bayesin kaavalla**.



Sijoittamalla lausekkeet (7) ja (8) kaavaan (5) saadaan **ehdolliseksi todennäköisyydeksi** $\Pr(B|V)$:

$$\begin{aligned}\Pr(B|V) &= \frac{\Pr(V \cap B)}{\Pr(V)} \\ &= \frac{\Pr(B) \Pr(V|B)}{\Pr(A) \Pr(V|A) + \Pr(B) \Pr(V|B)} \\ &= \frac{\frac{5}{8} \times 0.08}{\frac{3}{8} \times 0.05 + \frac{5}{8} \times 0.08} = \frac{8}{11} = 0.73\end{aligned}$$

Ehdollinen todennäköisyys $\Pr(B|V)$ on määrätty **Bayesin kaavalla**.

7.2. Ositukset

Joukon S osajoukot

$$B_1, B_2, \dots, B_n$$

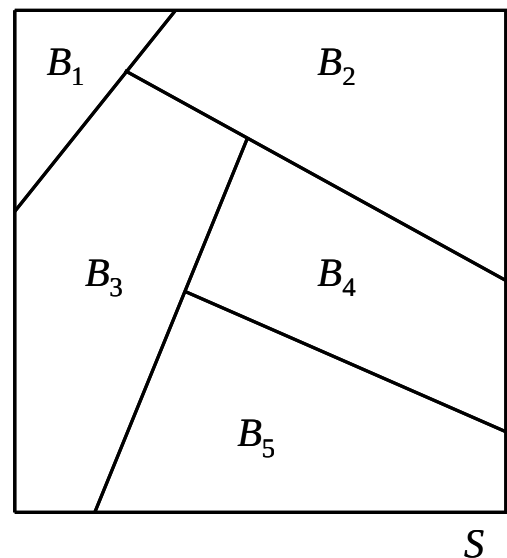
muodostavat joukon S **osituksen**, jos seuraavat ehdot pätevät:

- (i) $B_i \neq \emptyset, i = 1, 2, \dots, n$
- (ii) $B_i \cap B_j = \emptyset, i \neq j$
- (iii) $S = B_1 \cup B_2 \cup \dots \cup B_n$

Otosavaruuden S **ositus** $B_1, B_2, \dots, B_n$ muodostaa otosavaruuden S alkioiden *jaon toisensa poissulkeviin luokkiin* $B_1, B_2, \dots, B_n$, koska:

- (i) Joukot $B_1, B_2, \dots, B_n$ ovat *epätyhjiä*.
- (ii) Joukot $B_1, B_2, \dots, B_n$ ovat *pareittain pistevieraita*.
- (iii) $S = \cup B_i$

Jos tapahtumat $B_1, B_2, \dots, B_n$ muodostavat otosavaruuden S **osituksen**, *täsmälleen yksi* tapahtumista $B_1, B_2, \dots, B_n$ sattuu aina, kun se satunnaisilmiö, jonka tulostulokset otosavaruus S kuvaa, esiintyy.



7.3. Kokonaistodennäköisyyden kaava

Olkoon A epätyhjä otosavaruuden S osajoukko:

$$A \subset S, A \neq \emptyset$$

Oletetaan, että joukot

$$B_1, B_2, \dots, B_n$$

muodostavat otosavaruuden S **osituksen**. Tällöin pätee **kokonaistodennäköisyyden kaava**

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

Perustelu:

Oletetaan, että

$$A \subset S, A \neq \emptyset$$

ja että joukot

$$B_1, B_2, \dots, B_n$$

muodostavat otosavaruuden S osituksen.

Tällöin

$$(i) \quad B_i \neq \emptyset, i = 1, 2, \dots, n$$

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

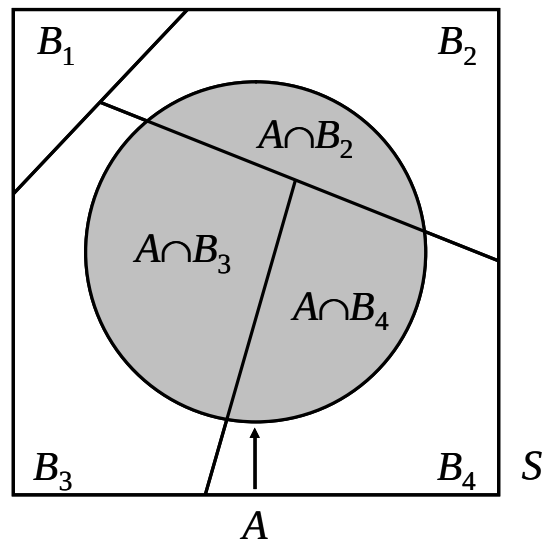
$$(iii) \quad S = B_1 \cup B_2 \cup \dots \cup B_n$$

Otosavaruuden S ositus $B_1, B_2, \dots, B_n$ **indusoi osituksen** joukkoon A :

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

ja

$$\begin{aligned} A &= A \cap S \\ &= A \cap (B_1 \cup B_2 \cup \dots \cup B_n) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n) \end{aligned}$$



Toisensa poissulkevien tapahtumien yhteenlaskusäännön perusteella

$$\Pr(A) = \Pr((A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n)) = \sum_{i=1}^n \Pr(A \cap B_i)$$

Koska yleisen tulosäännön perusteella

$$\Pr(A \cap B_i) = \Pr(B_i) \Pr(A | B_i), i = 1, 2, \dots, n$$

saamme lopulta

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

■

Kokonaistodennäköisyyden kaava ilmaisee otosavaruuden S osajoukon A todennäköisyyden $\Pr(A)$ otosavaruuden S osituksen $B_1, B_2, \dots, B_n$ määräämien todennäköisyyksien

$$\Pr(B_i), i = 1, 2, \dots, n$$

ja ehdollisten todennäköisyyksien

$$\Pr(A | B_i), i = 1, 2, \dots, n$$

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

7.4. Bayesin kaava

Olkoon A epätyhjä otosavaruuden S osajoukko:

$$A \subset S, A \neq \emptyset$$

Oletetaan, että joukot

$$B_1, B_2, \dots, B_n$$

muodostavat otosavaruuden S osituksen. Tällöin pätee *Bayesin kaava*

$$\Pr(B_i | A) = \frac{\Pr(B_i) \Pr(A | B_i)}{\sum_{i=1}^n \Pr(B_i) \Pr(A | B_i)}, i = 1, 2, \dots, n$$

Perustelu:

Oletetaan, että

$$A \subset S, A \neq \emptyset$$

ja että joukot

$$B_1, B_2, \dots, B_n$$

muodostavat otosavaruuden S osituksen.

Tällöin

$$(i) \quad B_i \neq \emptyset, i = 1, 2, \dots, n$$

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

$$(iii) \quad S = B_1 \cup B_2 \cup \dots \cup B_n$$

Otosavaruuden S ositus $B_1, B_2, \dots, B_n$ **indusoi osituksen** joukkoon A :

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

ja

$$\begin{aligned} A &= A \cap S \\ &= A \cap (B_1 \cup B_2 \cup \dots \cup B_n) \\ &= (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_n) \end{aligned}$$

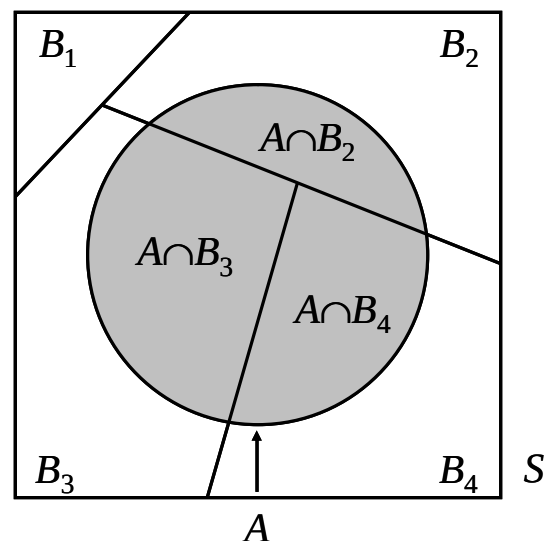
Ehdollisen todennäköisyyden määritelmän perusteella

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

Jos sovellamme tämän kaavan osoittajaan *yleistä tulosääntöä*

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

ja nimittäjään *kokonaistodennäköisyyden kaavaa*



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

niin saamme lopulta

$$\Pr(B_i | A) = \frac{\Pr(B_i) \Pr(A | B_i)}{\sum_{i=1}^n \Pr(B_i) \Pr(A | B_i)}, i = 1, 2, \dots, n$$

■

Bayesin kaavan todennäköisyyksiä

$$\Pr(B_i), i = 1, 2, \dots, n$$

on tapana kutsua **priori-todennäköisyyksiä**; *prior (lat.)*, edeltävä, aikaisempi.

Nimitys priori-todennäköisyys johtuu siitä, että todennäköisyyksille $\Pr(B_i)$ voidaan antaa seuraava tulkinta: Todennäköisyydet $\Pr(B_i)$ kuvaavat *ennakkokäsityksiämme* tapahtumien B_i todennäköisyydestä *ennen kuin käytettävissämme on tieto tapahtuman A sattumisesta*.

Bayesin kaavan todennäköisyyksiä

$$\Pr(B_i|A), i = 1, 2, \dots, n$$

on tapana kutsua **posteriori-todennäköisyyksiksi**; *posterior (lat.)*, jälkeen tuleva, myöhempi.

Nimitys posteriori-todennäköisyys johtuu siitä, että todennäköisyyksille $\Pr(B_i|A)$ voidaan antaa seuraava tulkinta: Todennäköisyydet $\Pr(B_i|A)$ kuvaavat *millaisiksi ennakkokäsityksiämme* tapahtumien B_i todennäköisyyksistä *kannattaa muuttaa sen jälkeen, kun on saamme tietää, että tapahtuma A on sattunut*.

Bayesin kaava kertoo siis miten *ennakkokäsityksiä* tapahtumien B_i todennäköisyyksistä on järkevää *korjata, sen jälkeen kun tapahtuma A on havaittu* tai miten *tietoa tapahtuman A sattumisesta* voidaan käyttää hyväksi tapahtumien B_i todennäköisyyksien arvioinnissa.

Posteriori-todennäköisyyksiä $\Pr(B_i|A)$ kutsutaan usein myös *käänteistodennäköisyyksiksi*, koska ne ovat ”käänteisiä” *tunnettuihin* todennäköisyyksiin $\Pr(A|B_i)$ nähden.

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

7.5. Kokonaistodennäköisyyden kaavan ja Bayesin kaavan systeemiteoreettinen tulkinta

Oletetaan, että **systeemissä** on *alkutila* L , *välitilat*

$$B_1, B_2, \dots, B_n$$

ja yksi sen *lopputiloista* on A . Oletetaan lisäksi, että *alkutilasta* L *voidaan päästä lopputilaan* A *vain käymällä jossakin välitiloista* $B_1, B_2, \dots, B_n$.

Olko

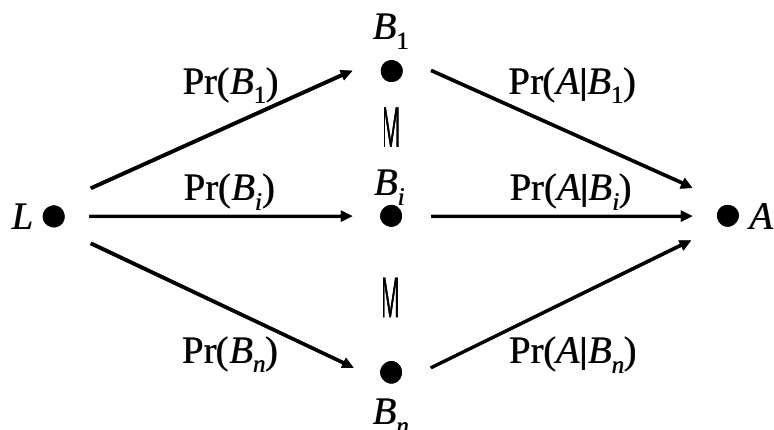
$$\Pr(B_i) = \Pr(\text{Käymme välitilassa } B_i)$$

$$\Pr(A | B_i) = \Pr(\text{Pääsemme välitilasta } B_i \text{ lopputilaan } A)$$

$$\Pr(A) = \Pr(\text{Pääsemme lopputilaan } A)$$

$$\Pr(B_i | A) = \Pr(\text{Olemme tulleet lopputilaan } A \text{ väli-tilan } B_i \text{ kautta})$$

Systeemiä voidaan havainnollistaa seuraavan **verkkodiagrammin** avulla:



Kokonaistodennäköisyyden kaavan mukaan

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

Voimme nyt antaa *kokonaistodennäköisyyden kaavalle* seuraavan *systeemiteoreettisen tulkinnan*: Todennäköisyys $\Pr(A)$ on todennäköisyys, että pääsemme lopputilaan A , kun lopputilaan A ei voi päästä käymättä jossakin välitiloista $B_1, B_2, \dots, B_n$. Huomaa, että lopputilaan A pääsemisen todennäköisyys $\Pr(A)$ saadaan *laskemalla yhteen* alkutilasta L lopputilaan A välitilojen $B_1, B_2, \dots, B_n$ kautta kulkevien *reittien* todennäköisyydet

$$\Pr(A \cap B_i) = \Pr(B_i) \Pr(A | B_i), i = 1, 2, K, n$$

Tässä sovellettu sääntö saa lisävalaistusta tutkittaessa ns. **puumaisten verkkojen** soveltamista todennäköisyyslaskentaan; lisätietoja: ks. lukua **Verkot ja todennäköisyyslaskenta** ja liitteen **Todennäköisyyslaskenta ja puumaiset verkot** kappaletta **Puut ja todennäköisyyslaskennan peruslaskusäännöt**.

Bayesin kaavan mukaan

$$\Pr(B_i | A) = \frac{\Pr(B_i) \Pr(A | B_i)}{\sum_{i=1}^n \Pr(B_i) \Pr(A | B_i)}, i = 1, 2, K, n$$

Voimme nyt antaa *Bayesin kaavalle* seuraavan *systeemiteoreettisen tulkinnan*: Todennäköisyys $\Pr(B_i | A)$ on *ehdollinen todennäköisyys* sille, että *olemme käyneet* välitilassa B_i , kun *ehto-tapahtumana* on se, että *olemme päässeet* lopputilaan A .

8. Verkot ja todennäköisyyslaskenta

8.1. Johdatteleva esimerkki

8.2. Puudiagrammit ja todennäköisyyslaskenta

8.3. Toimintaverkot ja niiden toimintatodennäköisyydet

Tarkastelemme tässä luvussa *todennäköisyyslaskennan ongelmien esittämistä ja ratkaisemista puumaisten verkkojen avulla*. Lisäksi tarkastelemme ns. **toimintaverkkojen toimintatodennäköisyyksien** määrittämistä.

Lisätietoja verkkoiteorian alkeista ja puumaisista verkoista eli puista sekä puumaisten verkkojen soveltamisesta todennäköisyyslaskentaan: ks. liitettä **Todennäköisyyslaskenta ja puudiagrammit**.

Avainsanat:

Insidenssikuvaus, Piste, Puu, Puutodennäköisyys, Päästöpuut, Reitti, Rinnan kytkentä, Sarjaan kytkentä, Silmukka, Särämä, Toimintatodennäköisyys, Toimintaverkko, Tulosääntö, Verkko, Yhteenlaskusääntö, Yhtenäisyys

8.1. Johdatteleva esimerkki

Tehdään lasten syntymisestä seuraavat (yksinkertaistavat) oletukset:

- (i) Lapset syntyvät perheisiin yksi kerrallaan.
- (ii) Syntyvän lapsen sukupuoli *ei riipu* siitä, mitä sukupuolta ovat olleet aikaisemmin perheeseen syntyneet lapset.
- (iii) Tytön todennäköisyys on sama kuin pojan todennäköisyys:

$$\Pr(\text{Syntyy tyttö}) = \Pr(\text{Syntyy poika}) = 1/2$$

Eräs aviopari haluaa oletusten (i)-(iii) pätiessä saada *tytön*, mutta *ei halua* hankkia *neljää lasta enempää*. Siksi pari päättää käyttää lasten hankkimisessa seuraavaa *strategiaa*: Lapsia hankitaan *kunnes* saadaan tyttö, mutta lapsia ei kuitenkaan hankita *neljää enempää*. Jos siis neljäskin lapsi on poika, pari on *epäonnistunut* strategiassaan. Mikä on todennäköisyys, että pari *onnistuu* strategiassaan?

Voimme rakentaa paria kohtaavista tapahtumavaihtoehtoista alla olevan **puudiagrammin**.

Puudiagrammissa:

T = Tyttö

P = Poika

Puun vasemmanpuoleiset *särmät* johtavat strategian onnistumiseen ja oikeanpuoleiset *särmät* johtavat strategian epäonnistumiseen.

Jokaisen särmän *todennäköisyys* = $1/2$.

Olkoon

A = Pari onnistuu strategiassaan

T_i = i . lapsi on tyttö

P_i = i . lapsi on poika

T = Syntyy tyttö

P = Syntyy poika

Tapahtumat T_1, T_2, T_3, T_4 muodostavat tapahtuman A osituksen:

$$A = T_1 \cup T_2 \cup T_3 \cup T_4, \quad T_i \cap T_j = \emptyset, \quad i \neq j$$

Tapahtumien T_i ja P_i todennäköisyydet

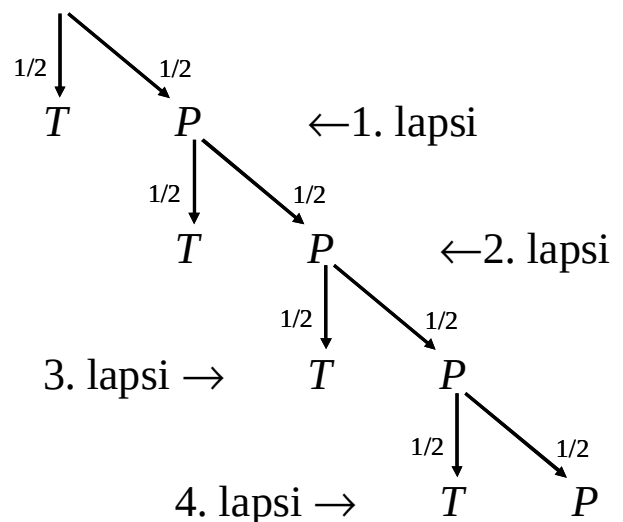
$$\Pr(T_i), \Pr(P_i), \quad i = 1, 2, 3, 4$$

voidaan määrätä *rekursiivisesti*.

Riippumattomien tapahtumien *tulosäännön* nojalla:

$$\Pr(T_1) = \Pr(T) = \frac{1}{2} = \Pr(P_1)$$

$$\Pr(T_2) = \Pr(T \cap P_1) = \Pr(T) \Pr(P_1) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = \Pr(P_2)$$



$$\Pr(T_3) = \Pr(T \cap P_2) = \Pr(T) \Pr(P_2) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} = \Pr(P_3)$$

$$\Pr(T_4) = \Pr(T \cap P_3) = \Pr(T) \Pr(P_3) = \frac{1}{2} \times \frac{1}{8} = \frac{1}{16} = \Pr(P_4)$$

Strategian onnistumisen todennäköisyydeksi saadaan toisensa poissulkevien tapahtumien yhteenlaskusäännön nojalla:

$$\begin{aligned} \Pr(A) &= \Pr(T_1 \cup T_2 \cup T_3 \cup T_4) \\ &= \Pr(T_1) + \Pr(T_2) + \Pr(T_3) + \Pr(T_4) \\ &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16} \end{aligned}$$

Voimme ajatella, että yo. laskutoimitukset on tehty seuraavalla tavalla:

(i) Todennäköisyydet

$$\Pr(T_1) = \frac{1}{2}, \Pr(T_2) = \frac{1}{4}, \Pr(T_3) = \frac{1}{8}, \Pr(T_4) = \frac{1}{16}$$

on saatu määräämällä puun loppupisteisiin T vievien reittien todennäköisyydet ja reittien todennäköisyydet on saatu reitin muodostavien puun särmien todennäköisyyksien tulona.

(ii) Strategian onnistumisen todennäköisyys

$$\Pr(A) = \frac{15}{16}$$

on saatu laskemalla loppupisteisiin T vievien reittien todennäköisyydet yhteen.

Johdattelevassa esimerkissä esitetty tekniikka voidaan yleistää ja sitä voidaan käyttää menestyksellisesti monien todennäköisyyslaskennan laskutehtävien ratkaisemiseen. Tämä perustuu siihen, että kaikkia äärellisten otosavaruuksien todennäköisyyksien laskusääntöjä voidaan havainnollistaa sopivasti konstruoitujen puudiagrammien avulla; lisätietoja: ks. liitettä **Todennäköisyyslaskenta ja puudiagrammit**.

8.2. Puudiagrammit ja todennäköisyyslaskenta

Periaatteessa jokainen alkeistodennäköisyyslaskennan tehtävä voidaan ratkaista käyttämällä apuna ns. **puudiagrammeja**. Jos tehtävän satunnaisilmiö on osattu kuvata puudiagrammilla, tehtävän ratkaisemisessa tarvittavat **puutodennäköisyydet** saadaan määrätyksi käyttämällä kahta yksinkertaista laskusääntöä, **tulosääntöä** ja **yhteenlaskusääntöä**.

Verkot

Verkko eli **graafi** koostuu **pisteistä**, **särmistä** ja **insidenssikuvauksesta**, joka kertoo *mitkä pisteet ovat särmien yhdistämiä*. Koska ajattelemme, että jokaisella särmällä on **alkupiste** ja **loppupiste**, tässä tarkasteltavat verkot ovat *suunnattuja verkkoja*. Tämä merkitsee sitä, että jokaisella särmällä on *suunta*, joka osoittaa särmän alkupisteestä särmän loppupisteeseen.

Pisteestä v on **reitti** pisteeseen w , jos on olemassa (suunnattujen) särmien jono, joka vie pisteestä v pisteeseen w niin, että jokaisen särmän alkupiste (pistettä v lukuun ottamatta) on edellisen särmän loppupiste. Reitti muodostaa **silmukan**, jos $v = w$.

Verkko on **yhtenäinen**, jos sen pisteiden joukkoa V ei voida osittaa kahdeksi epätyhjäksi osajoukoksi siten, että verkon *jokaisen* särmän alkupiste ja loppupiste kuuluvat *samaan* osajoukkoon.

Lisätietoja verkoista: ks. liitteen **Todennäköisyyslaskenta ja puudiagrammit** kappaletta **Verkot**.

Puut

Sanomme, että verkko on **puu**, jonka **juurena** on piste v_1 , jos seuraavat ehdot pätevät:

- (i) Verkko on *yhtenäinen*.
- (ii) Verkossa ei ole *silmukoita*.
- (iii) Jos $w \neq v_1$ on mielivaltainen verkon piste, pisteestä v_1 pisteeseen w pääsee *täsmälleen yhtä reittiä* pitkin.

Lisätietoja puista: ks. liitteen **Todennäköisyyslaskenta ja puudiagrammit** kappaletta **Puut**.

Puiden konstruointi

Satunnaisilmiötä voidaan kuvata **puudiagrammilla**, jos ilmiö voidaan esittää seuraavassa muodossa:

- (i) Ilmiöllä on **alkutila** ja useita vaihtoehtoisia **lopputiloja**.
- (ii) Ilmiö koostuu **tapahtumajonoista**, joissa alkutilasta siirrytään *vaiheittain* johonkin ilmiön vaihtoehtoisista lopputiloista.
- (iii) Jokaisessa ilmiön **vaiheessa** kohdataan yksi tai useampia **tapahtumavaihtoehtoja**, joista yksi *realisoituu* ja johtaa *uusin* tapahtumavaihtoehtoihin.

Satunnaisilmiötä vastaavan **puudiagrammin konstruointi**:

1. Asetetaan puun **juuri** vastaamaan ilmiön *alkutilaa*.
2. Asetetaan puun **loppupisteet** ("oksien kärjet") vastaamaan ilmiön *lopputiloja*.
3. Asetetaan puun **pisteet** ("oksien haarautumiskohdat") vastaamaan ilmiön *tapahtumia*.
4. Viedään puun jokaisesta pisteestä **särmä** ("oksa") kaikkiin sellaisiin pisteisiin, joita vastaavat *tapahtumavaihtoehdot* ovat ilmiön *siinä vaiheessa mahdollisia*.
5. Liitetään jokaiseen pisteestä *lähtevään* särmään *siinä vaiheessa* mahdollisten **tapahtumavaihtoehtojen todennäköisyydet**.

Havainnollistus:

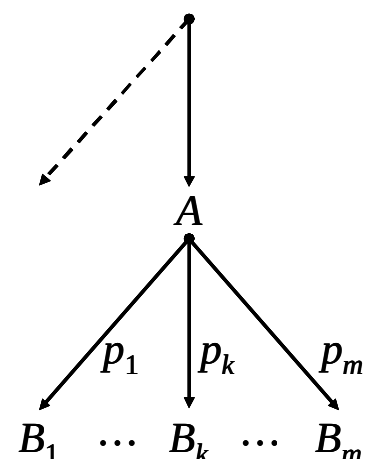
Puudiagrammin konstruointia voidaan havainnollistaa viereisellä kuviolla. Siinä tarkastellaan jotakin satunnaisilmiötä sellaisessa vaiheessa, jossa tapahtuma A on sattunut.

Olkoot tapahtuman A sattumisen jälkeen *mahdolliset tapahtumavaihtoehdot*

$$B_i, i = 1, 2, \dots, m$$

Viedään pisteestä A *särmä* jokaiseen pisteistä

$$B_i, i = 1, 2, \dots, m$$



Liitetään jokaiseen särmään

$$(A, B_i), i = 1, 2, \dots, m$$

ehdollinen todennäköisyys

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

jossa A on tapahtumajono, joka on tuonut pisteeseen A . Koska A :n sattumisen jälkeen *ei ole muita mahdollisia tapahtumavaihtoehtoja* kuin

$$B_i, i = 1, 2, \dots, m$$

niin todennäköisyyksien

$$p_i, i = 1, 2, \dots, m$$

on toteutettava ehto

$$\sum_{i=1}^m p_i = \sum_{i=1}^m \Pr(B_i | A) = 1$$

Huomautuksia:

- Puudiagrammi *piirretään* tavallisesti *joko* niin, että sen alkupiste on *ylhällä* ja loppupisteet ovat *alhaalla* tai niin, että sen alkupiste on *vasemmalla* ja loppupisteet ovat *oikealla*.
- Useat puun pisteet saattavat vastata *samaa* tapahtumaa.
- Mistä tahansa puun pisteestä *lähtevien särmien* todennäköisyyksien summa on 1.

Puutodennäköisyydet

Puutodennäköisyydellä tarkoitetaan todennäköisyyttä päästä *puun alkupisteestä yhden tai useamman muun puun pisteen määräämään yhdistettyyn tapahtumaan*. Mielivaltaisen puun *pisteen todennäköisyys* saadaan määräämällä alkupisteestä ko. pisteeseen vievän *reitin todennäköisyys*.

Reitin todennäköisyys saadaan soveltamalla reittiin kuuluvien särmien todennäköisyyksiin **tulosääntöä**. *Usean pisteen määräämän yhdistetyn tapahtuman todennäköisyys* saadaan soveltamalla ko. pisteisiin vievien reittien todennäköisyyksiin **yhteenlaskusääntöä**.

Tulosääntö puutodennäköisyyksille

Reitin todennäköisyys saadaan määräämällä reittiin kuuluvien särmien todennäköisyyksien *tulo*. Sääntöä kutsutaan **puutodennäköisyyksien tulosäännöksi**.

Perustelu:

- (1) Reitti on *tapahtumajono*, jonka muodostavat reitin pisteet.
- (2) Reitin muodostava tapahtumajono sattuu, jos *jokainen jonon tapahtumista sattuu*.
- (3) Todennäköisyyslaskennan *yleisen tulosäännön mukaan* reitin todennäköisyys saadaan määräämällä reittiin kuuluvien särmien todennäköisyyksien *tulo*.

Olkoon siis

$$L, A_1, A_2, \dots, A_k$$

jokin niistä vaihtoehtoisista *tapatumajonoista*, joista satunnaisilmiö muodostuu.

Tällöin parit

$$(L, A_1), (A_1, A_2), (A_2, A_3), \dots, (A_{k-1}, A_k)$$

muodostavat satunnaisilmiön *alkutilasta* L satunnaisilmiön (*loppu-*) tilaan A_k vievän reitin *särmät*. Liitetään reitin

$$(L, A_1), (A_1, A_2), (A_2, A_3), \dots, (A_{k-1}, A_k)$$

särmiin *todennäköisyydet* seuraavalla tavalla:

$$(L, A_1) \rightarrow \Pr(A_1) = p_1$$

$$(A_1, A_2) \rightarrow \Pr(A_2 | A_1) = p_2$$

$$(A_2, A_3) \rightarrow \Pr(A_3 | A_1 \cap A_2) = p_3$$

...

$$(A_{k-1}, A_k) \rightarrow \Pr(A_k | A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{k-1}) = p_k$$

Reitin

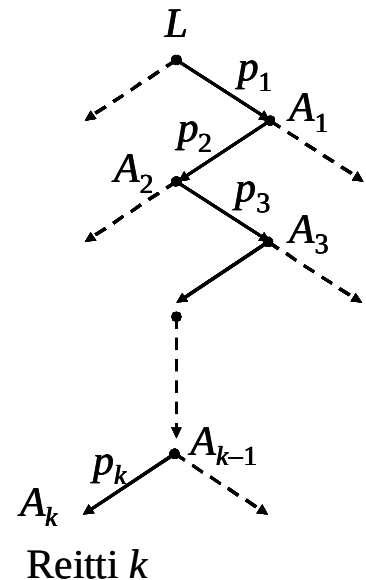
$$(L, A_1), (A_1, A_2), (A_2, A_3), \dots, (A_{k-1}, A_k)$$

todennäköisyys on **yleisen tulosäännön** nojalla:

$$\Pr(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_k)$$

$$= \Pr(A_1) \times \Pr(A_2 | A_1) \times \Pr(A_3 | A_1 \cap A_2) \times \dots \times \Pr(A_k | A_1 \cap A_2 \cap A_3 \cap \dots \cap A_{k-1})$$

$$= p_1 \times p_2 \times p_3 \times \dots \times p_k$$



■

Puutodennäköisyyksien tulosääntöä voidaan havainnollistaa yllä olevalla *puudiagrammilla*. Reitin k todennäköisyys on puutodennäköisyyksien tulosäännön mukaan

$$\Pr(\text{Reitti } k) = p_1 \times p_2 \times p_3 \times \dots \times p_k$$

Yhteenlaskusääntö puutodennäköisyyksille

Jos useita (*loppu-*) tiloja yhdistetään yhdeksi tapahtumaksi, näin saadun *yhdistetyn tapahtuman* todennäköisyys saadaan määräämällä ko. tiloihin vievien reittien todennäköisyyksien *summa*. Sääntöä kutsutaan **puutodennäköisyyksien yhteenlaskusäännöksi**.

Perustelu:

- (1) Puun eri pisteisiin vievät reitit ovat *toisensa poissulkevia*.
- (2) *Toisensa poissulkevien tapahtumien yhteenlaskusäännön mukaan* useista (*loppu-*) pisteistä yhdistämällä saatavan tapahtuman todennäköisyys saadaan määräämällä ko. pisteisiin vievien reittien todennäköisyyksien *summa*.

Yhdistetään satunnaisilmiön (*loppu-*) tilat $B_1, B_2, \dots, B_k$ yhdeksi tapahtumaksi

$$C = B_1 \cup B_2 \cup \dots \cup B_k$$

Olkoot *tiloja* $B_1, B_2, \dots, B_k$ vastaavat reitit

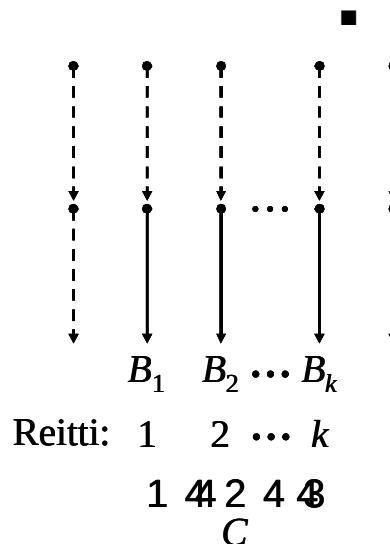
Reitti 1, Reitti 2, ..., Reitti k

Koska puun eri pisteisiin vievät reitit ovat *toisensa poissulkevia*, tapahtuman C todennäköisyys on **toisensa poissulkevien tapahtumien yhteenlaskusäännön** nojalla:

$$\begin{aligned}\Pr(C) &= \Pr(\text{Reitti 1 tai Reitti 2 tai ... tai Reitti } k) \\ &= \Pr(\text{Reitti 1}) + \Pr(\text{Reitti 2}) + \dots + \Pr(\text{Reitti } k)\end{aligned}$$

Puutodennäköisyyksien yhteenlaskusääntöä voidaan havainnollistaa viereisellä puudiagrammilla:

$$\begin{aligned}\Pr(C) &= \Pr(\text{Reitti 1}) \\ &+ \Pr(\text{Reitti 2}) \\ &\dots \\ &+ \Pr(\text{Reitti } k)\end{aligned}$$



Päätöspuut: Esimerkki

Tarkastellaan seuraavaa *päätöstilannetta*:

Munuaistaudissa potilaan munuaiset lopettavat vähitellen toimintansa, mikä johtaa potilaan kuolemaan. Oletetaan, että potilas voisi vapaasti valita hoidoksi joko *dialyysin* (munuaiskoneen) tai *munuaisensiirron*. Oletetaan edelleen, että hoitotuloksista on käytettävissä seuraavat tiedot:

Dialyysipotilaat:

- 68 % elää 5:n vuoden kuluttua
- 32 % on kuollut 5:n vuoden kuluttua

Munuaisensiirtopotilaat:

- 48 %:lla siirretty munuainen toimii normaalisti ja potilas elää 5:n vuoden kuluttua
- 43 %:lla siirretty munuainen ei toimi kunnolla ja he joutuvat dialyysiin
 - 42 % näistä potilaista elää 5:n vuoden kuluttua
 - 58 % näistä potilaista on kuollut 5:n vuoden kuluttua
- 9 % kuolee siirron aiheuttamiin komplikaatioihin

Kumpi hoidoista potilaan kannattaa valita, jos hän haluaa maksimoida todennäköisyyden olla elossa viiden vuoden kuluttua?

Merkitään:

- D = "Potilasta hoidetaan dialyysillä"
- S = "Potilaalle tehdään munuaisensiirto"
- SD = "Siirtopotilas joutuu dialyysiin"
- E = "Potilas elää 5 vuoden kuluttua"

K = ”Potilas on kuollut 5 vuoden kuluttua”

Hoitotulokset voidaan esittää seuraavina *ehdollisina todennäköisyyksinä*:

$$\Pr(E|D) = 0.68 \quad \Pr(K|D) = 0.32$$

$$\Pr(E|S) = 0.48 \quad \Pr(SD|S) = 0.43 \quad \Pr(K|S) = 0.09$$

$$\Pr(E|SD) = 0.42 \quad \Pr(K|SD) = 0.58$$

Voimme konstruoida potilaan edessä olevista vaihtoehtoisista tulevaisuuksista alla olevan *puudiagrammin*.

Jos potilas haluaa *maksimoida* todennäköisyyden olla elossa 5:n vuoden kuluttua, hänen on *verrattava* reitin 1 määräämän tapahtuman todennäköisyyttä reittien 3 ja 4 määräämän yhdistetyn tapahtuman todennäköisyyteen.

Reitin 1 määräämän tapahtuman todennäköisyys:

$$\Pr(\text{Reitti 1}) = 0.68$$

Reitin 3 määräämän tapahtuman todennäköisyys:

$$\Pr(\text{Reitti 3}) = 0.48$$

Reitin 4 määräämän tapahtuman todennäköisyys on *puutodennäköisyyksien tulosäännön* mukaan:

$$\Pr(\text{Reitti 4}) = 0.43 \times 0.42 = 0.1806$$

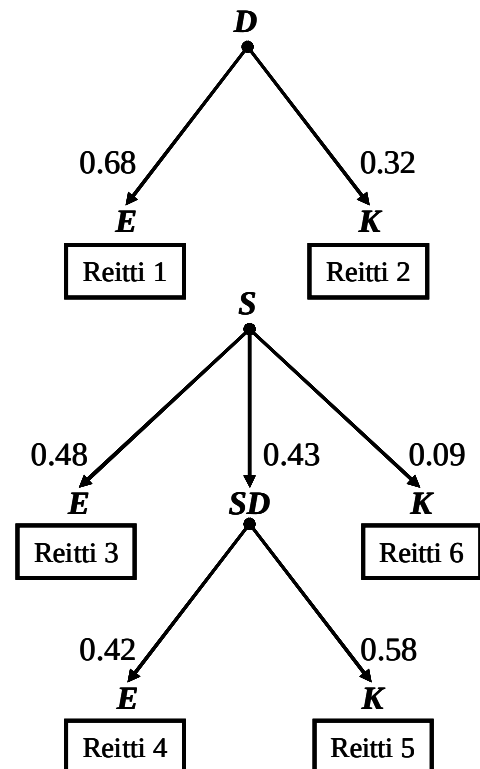
Reittien 3 ja 4 määräämän yhdistetyn tapahtuman todennäköisyys on *puutodennäköisyyksien yhteenlaskusäännön* mukaan:

$$\begin{aligned} \Pr(\text{Reitti 3 tai Reitti 4}) \\ = 0.48 + 0.1806 = 0.6606 \end{aligned}$$

Koska

$$\Pr(\text{Reitti 3 tai Reitti 4}) = 0.6606 < \Pr(\text{Reitti 1}) = 0.68$$

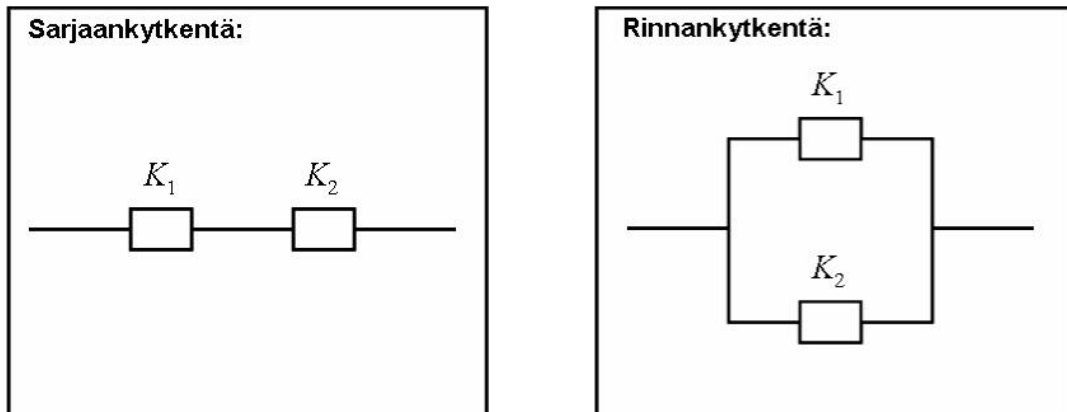
potilaan kannattaa valita dialyysi, jos hän haluaa maksimoida todennäköisyyden olla elossa viiden vuoden kuluttua.



8.3. Toimintaverkot ja niiden toimintatodennäköisyydet

Toimintaverkot

Toimintaverkko on *systeemi*, joka koostuu *komponenteista*, jotka on kytketty joko **sarjaan** tai **rinnan**. Alla olevat kytkentäkaaviot kuvaavat kahden komponentin K_1 ja K_2 muodostamia sarjaan- ja rinnankytkentöjä.



Tarkastellaan sarjaan kytkennän ja rinnankytkennän *toimintaehtoja*.

Merkitään

T = Komponentti *toimii*

F = Komponentti *ei toimi*

Komponenttien K_1 ja K_2 muodostama *sarjaan kytkentä toimii, jos komponentti K_1 toimii ja komponentti K_2 toimii*. Komponenttien K_1 ja K_2 sarjaan kytkennän *toimintaehdot* voidaan esittää seuraavana taulukkona:

K_1	K_2	K_1 ja K_2
T	T	T
T	F	F
F	T	F
F	F	F

Komponenttien K_1 ja K_2 muodostama *rinnankytkentä toimii, jos komponentti K_1 toimii tai komponentti K_2 toimii tai molemmat toimivat*. Komponenttien K_1 ja K_2 rinnankytkennän *toimintaehdot* voidaan esittää seuraavana taulukkona:

K_1	K_2	K_1 tai K_2
T	T	T
T	F	T
F	T	T
F	F	F

Sarjaan kytkennän toimintatodennäköisyys

Oletetaan, että komponentit K_1 ja K_2 on kytketty *sarjaan* ja oletetaan lisäksi, että komponentin K_1 toiminta (tai toimimattomuus) *ei riipu* komponentin K_2 toiminnasta (ja kääntäen).

Määritellään tapahtumat

$$A_1 = \text{"Komponentti } K_1 \text{ toimii"}$$

$$A_2 = \text{"Komponentti } K_2 \text{ toimii"}$$

Olkoot tapahtumien A_1 ja A_2 todennäköisyydet

$$p_1 = \Pr(A_1)$$

$$p_2 = \Pr(A_2)$$

Koska tapahtumat A_1 ja A_2 ovat oletuksen mukaan *riippumattomia*, komponenttien K_1 ja K_2 muodostama *sarjaan kytkennän toimintatodennäköisyys* on *riippumattomien tapahtumien tulosäännön* mukaan

$$\begin{aligned} & \Pr(\text{Komponentti } K_1 \text{ toimii ja komponentti } K_2 \text{ toimii}) \\ &= \Pr(A_1 \cap A_2) \\ &= \Pr(A_1)\Pr(A_2) \\ &= p_1 p_2 \\ &= \Pr(\text{Komponentti } K_1 \text{ toimii})\Pr(\text{Komponentti } K_2 \text{ toimii}) \end{aligned}$$

Rinnankytkennän toimintatodennäköisyys

Oletetaan, että komponentit K_1 ja K_2 on kytketty *rinnan* ja oletetaan lisäksi, että komponentin K_1 toiminta (tai toimimattomuus) *ei riipu* komponentin K_2 toiminnasta (ja kääntäen).

Määritellään tapahtumat

$$A_1 = \text{"Komponentti } K_1 \text{ toimii"}$$

$$A_2 = \text{"Komponentti } K_2 \text{ toimii"}$$

Olkoot tapahtumien A_1 ja A_2 todennäköisyydet

$$p_1 = \Pr(A_1)$$

$$p_2 = \Pr(A_2)$$

Koska tapahtumat A_1 ja A_2 ovat oletuksen mukaan *riippumattomia*, komponenttien K_1 ja K_2 muodostaman *rinnan kytkennän toimintatodennäköisyys* on *yleisen yhteenlaskusäännön* ja *riippumattomien tapahtumien tulosäännön* mukaan

$$\begin{aligned} & \Pr(\text{Komponentti } K_1 \text{ toimii tai komponentti } K_2 \text{ toimii}) \\ &= \Pr(A_1 \cup A_2) \\ &= \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2) \\ &= \Pr(A_1) + \Pr(A_2) - \Pr(A_1)\Pr(A_2) \\ &= p_1 + p_2 - p_1 p_2 \\ &= \Pr(\text{Komponentti } K_1 \text{ toimii}) + \Pr(\text{Komponentti } K_2 \text{ toimii}) \\ &\quad - \Pr(\text{Komponentti } K_1 \text{ toimii})\Pr(\text{Komponentti } K_2 \text{ toimii}) \end{aligned}$$

Sovellus

Viereinen kuva esittää toimintaverkkoa, joka koostuu 4:stä komponentista K_1, K_2, K_3, K_4 seuraavalla tavalla:

- (i) Komponentit K_1 ja K_2 on kytketty *sarjaan*.
- (ii) Komponentit K_3 ja K_4 on kytketty *sarjaan*.
- (iii) Komponenttipari K_1 ja K_2 on kytketty *rinnan* komponenttiparin K_3 ja K_4 kanssa.

Olkoon

$$\Pr(K_i \text{ toimii}) = p, i = 1, 2, 3, 4$$

Oletetaan lisäksi, että yhdenkään komponentin toiminta *ei riipu* muiden komponenttien toiminnasta.

Sarjaan kytkennän toimintatodennäköisyyden kaavan mukaan

$$\Pr(K_1 \text{ toimii ja } K_2 \text{ toimii}) = p \times p = p^2$$

$$\Pr(K_3 \text{ toimii ja } K_4 \text{ toimii}) = p \times p = p^2$$

Rinnan kytkennän toimintatodennäköisyyden kaavan mukaan

$$\Pr(\text{Systeemi toimii}) = p^2 + p^2 - p^2 \times p^2 = 2p^2 - p^4$$

Kuva oikealla esittää esimerkin systeemin toimintatodennäköisyyttä

$$f(p) = 2p^2 - p^4$$

yksittäisen komponentin toimintatodennäköisyyden p funktiona.

Kuvioon on piirretty myös suora

$$f(p) = p$$

Kuviosta nähdään:

- (i) $f(p)$ on p :n kasvava funktio.
- (ii) Esimerkin systeemin toimintatodennäköisyys voi olla *suurempi*, *pienempi* tai *yhtä suuri* kuin yksittäisen komponentin toimintatodennäköisyys.

