JYVÄSKYLÄN YLIOPISTO
MATEMATIIKAN JA TILASTOTIETEEN LAITOS

Pro gradu -tutkielma

Todennäköisyyslaskentaa koulumatematiikasta mittateoriaan

Sofia Länsivuori

6. kesäkuuta 2024

Tekijä
Sofia Länsivuori

Otsikko
Todennäköisyyslaskentaa koulumatematiikasta mittateoriaan

Tutkinto-ohjelma
Matematiikan aineenopettajan maisteriohjelma

Päivämäärä	**Sivumäärä**
6. kesäkuuta 2024	70

Tiivistelmä

Tässä tutkielmassa tutustutaan todennäköisyyslaskentaan ja sen opetukseen peruskoulussa ja lukiossa. Todennäköisyyslaskennan käsitettä aletaan rakentamaan alakoulusta aina yliopistomatematiikkaan asti. Tutkielmassa opitaan laskemaan erilaisia todennäköisyyksiä, kuten klassisia, geometrisia ja tilastollisia todennäköisyyksiä. Tutkielmassa käsitellään myös kombinatoriikkaa sekä diskreettiä jakaumaa. Koulumatematiikan aiheiden lisäksi opitaan muun muassa todennäköisyysavaruuden, todennäköisyysmitan ja sigma-algebran käsitteet ja käydään läpi erilaisia esimerkkejä.

Määritelmien ja lauseiden esittelyn ohessa vertaillaan myös sitä, miten näitä käsitellään eri kouluasteilla, ja mitä eroavaisuuksia määrittelyissä on. Tämän lisäksi käydään läpi erilaisia esimerkkejä ja verrataan näiden tehtävien haastavuutta peruskoulun ja lukion välillä. Vertailun apuna käytetään Sanoma Pron ja Otavan eri oppikirjojen sisältöjä sekä lukio-opetuksen ja perusopetuksen opetussuunnitelmia. Lopuksi tutkitaan lukiomatematiikan ja mittateorian eroavaisuuksia.

Tutkielman tarkoitus on antaa etenkin matematiikan aineenopettajille hyvät valmiudet opettaa todennäköisyyslaskentaa peruskoulussa ja lukiossa ja syventää omaa osaamistaan mittateorian osa-alueella. Tutkielman luettuaan lukijalla on käsitys siitä, mitä aiheita todennäköisyyslaskennasta opetetaan peruskoulussa ja lukiossa, ja mitkä aiheet jäävät koulumatematiikan ulkopuolelle.

Sisällys

Johdanto

Olet ehkä joskus miettinyt, kuinka todennäköistä on saada seitsemän oikein lotossa. Kun 40 numerosta valitaan seitsemän, on todennäköisyys voittaa hyvin pieni $\frac{1}{18643560}$ eli vain $0,0000054\%$. Nykyään erilaisia vedonlyöntiin liittyviä pelejä on monia ja niihin rahaa laittavat useat eri pelaajat. Todennäköisyyslaskentaa voidaan hyödyntää esimerkiksi Veikkauksen eri pelejä tarkastellessa, jotta ymmärretään, kuinka pienet voittomahdollisuudet todellisuudessa ovat.

Uhkapelaamiseen liittyvä todennäköisyyslaskenta ei ole mikään uusi löydös. Sanoma Pron Moodi 8 -oppikirjassa (2023) esitetään, että todennäköisyyslaskennan teoria on saanut alkunsa uhkapeleihin liittyvistä ongelmista. Girolamo Cardano pohti jo 1500-luvulla uhkapeleihin liittyvää matematiikkaa. Varsinaisesti teoria alkoi kehittyä 1650-luvulla, kun Blaise Pascal ja Pierre Fermat alkoivat etsimään ratkaisuja ritari de Meren noppapeliongelmiin. Todennäköisyyslaskenta täsmentyi teoriaksi vuonna 1933, kun Andrei Nikolajevitš Kolmogorov esitti todennäköisyyslaskennan aksioomat.

Vaikka uhkapelaamisella on merkittävä rooli todennäköisyyslaskennan historiassa, niin nykyään sitä hyödynnetään myös monella muulla osa-alueella. Moodi 8 -oppikirja (2023) kertoo, että todennäköisyyslaskentaa käytetään muun muassa vakuutustoiminnassa, sääennusteissa ja monilla eri tieteenaloilla. Tämän vuoksi on tärkeää, että todennäköisyyslaskentaa opetetaan kouluissa.

Heiton Pro Gradu -tutkielmassa (2015) mainitaan todennäköisyyslaskennan olevan haastavimmista aiheista lukiossa. Tämä on mielenkiintoista, sillä todennäköisyyslaskentaan tutustuminen aloitetaan jo alakoulussa. Aiheen käsittely jatkuu myös peruskoulussa ja lukiossa, joissa käsitellään paljon samoja asioita. Siksi on hyvä kysymys, miksi todennäköisyyslaskenta koetaan haastavaksi.

Heitto (2015) kertoo, että todennäköisyyslaskennan oppimisen haaste on etenkin virheelliset ennakkokäsitykset. Heitto painottaa, että opettajan tulisi itse hallita todennäköisyyslaskennan salat niin hyvin, että hän pystyisi tunnistamaan nämä virheelliset käsitykset ja korjaamaan ne. Tämän tutkielman tarkoituksena on tutustuttaa lukija todennäköisyyslaskentaan ja antaa etenkin juuri valmistuneille ja työnsä aloittaville matematiikan aineenopettajille hyvä pohja todennäköisyyslaskennan opettamiseen.

Lukioon pyrkiville ja siellä oleville opiskelijoille todennäköisyyslaskennan opettaminen ja oppiminen on tärkeää, sillä ylioppilaskirjoituksissa todennäköisyyslaskennalla on merkittävä rooli. Lähes joka vuosi on jokin tehtävä liittyen todennäköisyyksiin, niin pitkän kuin lyhyen oppimäärän matematiikan ylioppilaskirjoituksissa. Lyhyen oppimäärän ylioppilaskirjoituksissa tämä painottuu vielä enemmän, sillä usein ylioppilaskirjoituksissa on ollut jopa kaksi tai kolme tehtävää liittyen todennäköisyyslaskentaan.

Tässä tutkielmassa käydään läpi todennäköisyyslaskentaa koulumatematiikasta mittateoriaan. Ensimmäisessä luvussa käsitellään todennäköisyyslaskentaa peruskoulun ja lukion matematiikan opetuksessa. Jokaiselta asteelta käydään läpi, mitä opetussuunnitelman perusteissa vaaditaan opetukselta ja miten oppikirjoissa näitä aiheita käsitellään. Tämän jälkeen vertaillaan opetusta ja käydään läpi eroja peruskoulun ja lukion välillä.

Toisessa luvussa käsitellään todennäköisyyslaskentaa yliopistomatematiikassa. Luvussa tutustutaan muun muassa todennäköisyysavaruuden, sigma-algebran ja satunnaismuuttujan käsitteisiin. Luvussa käydään läpi erilaisia määritelmiä, lauseita, lauseiden todistuksia ja esimerkkejä. Samalla kun asioita esitellään, vertaillaan aiheiden käsittelyä lukion ja mittateorian välillä.

1 Koulumatematiikkaa

1.1 Alakoulu

Perusopetuksen opetussuunnitelman perusteiden (2014) mukaan todennäköisyyslaskennan opetus aloitetaan jo alakoulussa. 1-2-luokilla matematiikan opetuksen tavoitteena on tarjota erilaisia kokemuksia matemaattisten käsitteiden ja rakenteiden muodostumisen perustaksi. Näillä luokilla ei vielä tutustuta todennäköisyyksiin. 3-6-luokilla opetuksen tavoitteissa mainitaan, että todennäköisyyteen aletaan tutustumaan arkitilanteita hyödyntämällä. Opetuksen apuna käytetään päättelyä, onko joku tapahtuma mahdoton, mahdollinen vai varma. Todennäköisyyden osaamisesta ei ole mainintaa hyvän arvosanan kriteereissä.

Alakoulun matematiikan oppikirjoista käsitellään tässä tutkielmassa Sanoma Pron Milli-kirjasarjaa, sekä Otavan Tuhattaituri-kirjasarjaa. Näissä kirjasarjoissa todennäköisyyteen tutustutaan eri vuosiluokilla. Milli-kirjasarjassa todennäköisyyslaskentaa opetetaan viidennen luokan syyslukukaudella, kun taas Tuhattaituri-kirjasarjassa opetuksen on suunniteltu tapahtuvan kuudennen luokan kevätlukukaudella.

Sanoma Pron Milli 5A -oppikirjassa (2023) todennäköisyyslaskentaan tutustuminen aloitetaan arkielämän esimerkkien kautta. Otavan Tuhattaituri 6b -oppikirjassa (2018) todennäköisyyteen taas perehdytään huomattavasti syvällisemmin kuin Perusopetuksen opetussuunnitelman perusteissa (2014) vaaditaan. Toisin kuin Milli 5A -kirjassa, todennäköisyyttä aletaan opettamaan suoraan todennäköisyyden käsitteellä. Tämän lisäksi Tuhattaituri 6b -kirjassa tutustutaan geometriseen todennäköisyyteen, todennäköisyyteen prosentteina, erilaisten yhdistelmien, sekä erilaisten järjestysten lukumäärään. Opetuksessa kannattaa huomioida se, että oppilailla, joilla on ollut alakoulun matematiikassa käytössä Tuhattaituri -kirjasarja, saattaa olla paremmat valmiudet todennäköisyyden käsittelyyn.

1.2 Yläkoulu

Alakoulussa matematiikan opetuksen tavoitteena oli tarjota kokemuksia todennäköisyyksistä. 7–9-luokilla todennäköisyyteen perehdytään tarkemmin. Perusopetuksen opetussuunnitelman perusteissa (2014) yksi matematiikan opetuksen tavoitteista on ohjata oppilaita laskemaan todennäköisyyksiä. Päättöarvioinnin kriteerinä hyvälle arvosanalle on se, että oppilas kykenee määrittämään sekä klassisia, että tilastollisia todennäköisyyksiä.

Todennäköisyyslaskennan opetukseen yläkoulussa tutustutaan tutkimalla Sanoma Pron Kuutio-kirjasarjaa sekä Otavan Ääretön-kirjasarjaa. Sanoma Pron Kuutio -kirjasarjassa aletaan käymään läpi todennäköisyyslaskentaa K11 Todennäköisyyslaskenta ja kertaus -kurssissa. Tämä kurssi on sijoitettu yhdeksännen luokan keväälle, joka on puolen vuosiviikkotunnin mittainen. Otavan Ääretön -kirjasarja sisältää oman oppikirjan Ääretön Tilastot ja todennäköisyys (2023), jonka käsittely voidaan aloittaa millä tahansa yläkoulun vuosiluokista.

Yläkoulun kirjasarjoissa ei ole niin suurta eroa käsiteltävien aiheiden välillä kuin alakoulun kirjasarjoissa. Kirjat kuitenkin käsittelevät aiheita eri järjestyksessä. Seuraavaksi tutkitaan yksityiskohtaisemmin mitä aiheita ja tehtäviä oppikirjat sisältävät, sekä missä järjestyksessä aiheita käydään. Alla on molempien kirjasarjojen sisällysluettelot todennäköisyyslaskennan osalta.

1. Matemaattinen päättely
2. Kuinka monella tavalla?
3. Klassinen todennäköisyys
4. Tilastollinen todennäköisyys
5. Peräkkäiset tapahtumat
6. TEEMA: Liikennematematiikkaa
7. TEEMA: Pelimatematiikkaa
8. Kertaus

Kuva 1.1: Kuutio

1. Todennäköisyys
2. Klassinen todennäköisyys
3. Tilastollinen todennäköisyys
4. Tuloperiaate
5. Järjestysten lukumäärä
6. Peräkkäiset tapahtumat
7. Todennäköisyyslaskuja
8. Matemaattinen päättely

Kuva 1.2: Ääretön

Sanoma Pron Kuutio K11 -oppikirjassa (2023) todennäköisyyden käsittely aloitetaan matemaattisella päättelyllä. Tämä kappale sisältää oletuksen ja johtopäätöksen käsitteet. Tässä kappaleessa tutkitaan, onko jokin väite tosi vai epätosi ja lisäksi tutustutaan vastaesimerkkeihin. Otavan Ääretön -kirjassa (2023) tämä matemaattinen päättely -kappale käydään vasta viimeisenä. Kappale käsittelee pitkälti samat asiat kuin Kuutio-oppikirja.

Kuutio-kirjan toinen kappale käsittelee aihetta “Kuinka monella tavalla?”. Kappaleessa tutustutaan tuloperiaatteeseen, kertomaan sekä järjestykseen. Ääretön kirjasarjassa tuloperiaate ja järjestysten lukumäärä on jaettu kahteen eri kappaleeseen, joista neljäs käsittelee tuloperiaatteen ja viides järjestysten lukumäärän. Itse kertoman käsitettä tässä kirjassa ei esitetä, vaikka tämän tyylisiä laskuja kirjassa esiintyy.

Klassista todennäköisyyttä aletaan Kuutio-oppikirjassa käsittelemään kappaleessa kolme. Ääretön kirjassa klassinen todennäköisyys on jaettu kahteen kappaleeseen; Todennäköisyys ja Klassinen todennäköisyys. Nämä kappaleet käydään kirjassa ensimmäisenä. Molemmissa kirjoissa käydään läpi sekä todennäköisyyden, että klassisen todennäköisyyden käsitteet. Kuutio-oppikirja sisältää myös pienen osion todennäköisyyden historiaa. Ääretön-oppikirjassa taas käsitellään vastatapahtuma. Molempien kirjasarjojen seuraava kappale käsittelee tilastollista todennäköisyyttä. Kirjat sisältävät erilaisia aineistoja, joista oppilaiden tehtävänä on laskea tilastollisia todennäköisyyksiä.

Kuutio-oppikirjan viides ja viimeinen kappale ennen teema- ja kertausosioita on peräkkäiset tapahtumat. Ääretön-oppikirjassa tätä aihetta käsitellään kappaleessa 6 vasta tuloperiaatteen ja järjestysten lukumäärän jälkeen. Peräkkäisten tapahtumien määritelmän lisäksi Kuutio-oppikirjassa perehdytään myös riippumattomuuteen. Molemmat kirjat sisältävät peräkkäisiin tapahtumiin, kuten nopan heittoon, liittyviä tehtäviä.

Näiden lisäksi Ääretön-oppikirja sisältää kappaleen todennäköisyyslaskuja, jossa nimensä mukaan on todennäköisyyteen liittyviä laskuja. Kuutio-oppikirja taas sisältää Teema-osion, jossa todennäköisyyksiä lasketaan liikenne- ja pelimatematiikkaan liittyen. Kaiken kaikkiaan näiden oppikirjojen sisällössä ei siis ole suuria eroavaisuuksia. Käsiteltävät aiheet ovat samat, ainoastaan asioiden käsittelyjärjestys ja tehtävien määrä eroavat jonkin verran toisistaan.

1.3 Lukio

Lukion pitkän oppimäärän matematiikassa todennäköisyyteen tutustutaan MAA8 Tilastot ja todennäköisyys sekä MAA12 Analyysi ja jatkuva jakauma -moduuleissa. Molemmat moduulit ovat kahden opintopisteen arvoisia. Tässä tutkielmassa tarkastellaan vain todennäköisyyslaskentaan liittyviä sisältöjä ja jätetään tilastollinen osuus huomioimatta, niin lyhyen kuin pitkän oppimäärän matematiikassa.

Lukion opetussuunnitelman perusteiden (2019) mukaan MAA8-moduulin tavoitteita on perehtyä todennäköisyyden käsitteeseen ja laskusääntöihin. Lisäksi opiskelijan tulisi ymmärtää diskreetin todennäköisyysjakauman käsite, sekä oppia määrittämään jakauman odotusarvon ja tulkita sitä. Opiskelijan odotetaan myös osaavan hyödyntää ohjelmistoja eri todennäköisyyden tehtävissä. Moduulin keskeisiä sisältöjä ovat klassinen ja tilastollinen todennäköisyys, permutaatiot ja kombinaatiot, binomijakauma, sekä diskreetti todennäköisyysjakauma ja sen odotusarvo. Valtakunnallisessa valinnaisiin opintoihin kuuluvassa MAA12-moduulissa esiintyvät normaalijakauman, tiheyssekä kertymäfunktion käsitteet.

Lukion lyhyen oppimäärän matematiikassa on kahden opintopisteen moduuli MAB5 Tilastot ja todennäköisyys. Lukion opetussuunnitelman perusteissa (2019) käydään läpi moduulin tavoitteita ja keskeisiä sisältöjä. Moduulin tavoitteena on, että opiskelija perehtyy todennäköisyyden perusteihin ja sitä havainnollistaviin malleihin. Lisäksi opiskelijan tulisi hallita ohjelmistojen käyttö todennäköisyyslaskennassa. Moduulin keskeisiä sisältöjä todennäköisyyden kannalta ovat todennäköisyyden käsite, yhteen- ja kertolaskusääntö, kombinaatiot ja tuloperiaate, sekä todennäköisyyslaskennan malleja.

Todennäköisyyslaskennan opiskelu jatkuu vielä moduulissa MAB9 Tilastolliset ja todennäköisyysjakaumat, joka kuuluu valtakunnallisiin valinnaisiin opintoihin. Tässä moduulissa todennäköisyyslaskennan tavoitteena on tutustua normaali- ja binomijakaumaan. Moduulin keskeisiä sisältöjä ovat normaalijakauman odotusarvo ja keskihajonta, toistokoe, sekä binomijakauma.

Myös lukion oppikirjoista käsitellään Otavan ja Sanoma Pron kirjasarjoja. Otavalla on pitkän ja lyhyen oppimäärän matematiikkaan eri nimiset sarjat: Juuri (MAA) ja Huippu (MAB). Tilastot ja todennäköisyys -moduulin kirjat ovat Huippu 5 Tilastot ja todennäköisyys (2021) ja Juuri 8 Tilastot ja todennäköisyys (2022), mutta lyhennetään näitä tässä tutkielmassa Huippu 5 ja Juuri 8. Sanoma Prolla on sekä lyhyessä että pitkässä matematiikassa käytössä kirjasarja Moodi (MAA ja MAB, 2023). Merkitään näitä kirjoja Moodi 5 ja Moodi 8. Seuraavaksi lähdetään tutustumaan näiden kirjasarjojen sisältöihin.

1.3.1 Pitkä matematiikka

Juuri 8 ja Moodi 8 -oppikirjojen rakenne on hyvin samanlainen. Todennäköisyyslaskennan osiot alkavat molemmissa kirjoissa luvulla todennäköisyys, jonka jälkeen jatketaan aiheeseen kombinatoriikka ja viimeisenä käsitellään todennäköisyysjakaumia. Ainoa eroavaisuus näiden kirjasarjojen sisällysluetteloilla on se, että Moodi 8 -oppikirja sisältää lisäksi luvun todennäköisyyden laskusääntöjä, jonka aiheet Juuri 8 -kirjassa käydään ensimmäisessä ja toisessa luvussa.

Molempien oppikirjojen ensimmäisessä kappaleessa tutustutaan todennäköisyyteen ja todennäköisyyden eri käsitteisiin. Määritelmistä käydään läpi klassinen todennäköisyys sekä vastatapahtuma. Juuri 8 -kirjassa käydään näiden lisäksi myös komplementtisääntö. Toinen kappale on tilastollinen ja geometrinen todennäköisyys, jossa käydään läpi näiden määritelmät ja ominaisuudet. Moodi 8 -oppikirjan todennäköisyyttä käsittelevä luku päättyy tähän kappaleeseen.

Tämän jälkeen oppikirjoissa on pieni eroavaisuus, sillä Juuri 8 -kirja jatkuu Kertolaskusääntö ja riippumattomuus -kappaleella, joka Moodi 8 -kirjassa käydään vasta luvussa kolme. Sisällöllisesti nämä kappaleet eivät juurikaan eroa. Molemmat kappaleet käsittelevät riippumattomuuden määritelmän ja riippumattomien tapahtumien sekä yleisen kertolaskusäännön. Näiden lisäksi käydään myös ehdollisen todennäköisyyden määritelmä.

Kertolaskusäännön jälkeen kirjasarjat jatkavat yhteenlaskusääntöön, joissa käydään yhteenlaskusääntö erillisille tapahtumille ja yhteenlaskusäännön yleinen muoto. Juuri 8 -oppikirjan todennäköisyyttä käsittelevä luku päättyy todennäköisyyslaskennan harjoitteluun. Moodi 8 -kirjan luku kolme käsittelee vielä toistokokeen, mutta palataan sen sisältöihin myöhemmin.

Juuri 8 -oppikirja jatkuu luvulla kombinatoriikkaa ja Moodi 8 käsittelee samoja aiheita toisessa luvussaan vaihtoehtojen lukumääriä. Moodi 8 -oppikirjassa tuloperiaate ja jonojen lukumäärä on jaoteltu omiin kappaleisiinsa, kun taas Juuri 8 -kirja käsittelee molempia asioita ensimmäisessä luvussa tuloperiaate ja permutaatiot. Kappaleissa esitellään tuloperiaatteen, permutaation sekä kertoman käsitteet. Molemmat kirjat jatkavat tästä kappaleeseen kombinaatiot, joka sisältää binomikertoimen määritelmän.

Tähän päättyy Moodi 8 -kirjan vaihtoehtojen lukumääriä luku. Juuri 8 -kirjan luku kombinatoriikkaa jatkuu vielä alaluvulla todennäköisyyksiä ja kombinatoriikkaa, joka sisältää aiheeseen liittyviä esimerkkejä ja laskuja. Viimeinen alaluku tässä luvussa on toistokoe, joka Moodi 8 -kirjassa käsitellään alaluvussa 14. Molemmissa kirjoissa tämä kappale käsittelee toistokokeen todennäköisyyden eli binomitodennäköisyyden lauseen.

Oppikirjojen viimeiset luvut käsittelevät todennäköisyysjakaumia. Ensimmäisen alaluvun aiheena on molemmissa kirjoissa diskreetti jakauma, jossa käydään läpi diskreetin satunnaismuuttujan sekä satunnaismuuttujan jakauman määritelmät. Moodi 8 -kirjassa ensimmäisessä alaluvussa käsitellään satunnaismuuttujan odotusarvo ja keskihajonta, kun taas Juuri 8-kirjassa käydään läpi binomijakauma. Kirjojen viimeiset alaluvut ovat siten myös päinvastaiset, sillä Moodi 8 -kirja päättyy kappaleeseen binomijakauma ja Juuri 8-kirja päättyy kappaleeseen diskreetin jakauman odotusarvo.

Kuten huomataan, niin sisällölliset erot näiden kirjasarjojen välillä on hyvin pienet. Asioita käydään hieman eri järjestyksessä, mutta opetettavat asiat ovat hyvin samanlaiset, joitain pieniä yksityiskohtia lukuun ottamatta. Moodi-oppikirjassa opetetaan esimerkiksi ehdollinen todennäköisyys, jota ei ole Lukion opetussuunnitelman perusteiden (2019) sisällössä. Myös lauseiden todistuksissa on joitain eroja, esimerkiksi Juuri 8-oppikirjassa käydään permutaatiolauseiden todistukset, mutta Moodi 8 -oppikirjassa taas ei käydä näitä lainkaan.

1.3.2 Lyhyt matematiikka

Lyhyen ja pitkän oppimäärän matematiikan Tilastot ja todennäköisyys -moduulin kirjasarjojen sisällysluettelot todennäköisyyslaskennan osalta eivät juurikaan eroa toisistaan. Myös kirjojen Huippu 5 (2021) ja Moodi 5 (2023) sisällöt ovat hyvin samanlaiset. Kuten pitkässä oppimäärässä niin myös lyhyessä suurimmat erot tulevat opetettavien aiheiden järjestyksessä.

Huippu 5 -oppikirja alkaa luvulla vaihtoehtojen lukumääriä, joka sisältää alaluvut jonot ja joukot, tuloperiaate ja kertoma sekä kombinaatio. Moodi 5 -oppikirjassa samanniminen luku on taas sijoitettu kirjan viimeiseksi ja se pitää sisällään alaluvut 14 tuloperiaate ja 15 osajoukkojen lukumäärä. Moodi 5 -kirjan todennäköisyyslaskennan osio alkaa Todennäköisyys-luvulla. Tässä luvussa on alaluvut klassinen, tilastollinen ja geometrinen todennäköisyys. Huippu 5 -kirjassa luku 2 johdatus todennäköisyyslaskentaan sisältää samat aiheet sillä eroavaisuudella, että klassinen ja tilastollinen todennäköisyys käsitellään samassa alaluvussa.

Molemmat oppikirjat jatkavat tästä lukuun todennäköisyyden laskusääntöjä, jossa käsitellään alaluvuissa kertolasku- ja yhteenlaskusääntö sekä vastatapahtuman todennäköisyyksiä. Nämä luvut eroavat oppikirjoissa hieman toisistaan, sillä Moodi 5 -kirjassa käydään läpi vain riippumattomien tapahtumien kertolaskusääntö ja erillisten tapahtumien yhteenlaskusääntö, kun taas Huippu 5 -kirjassa käydään molemmista tapauksista myös yleinen sääntö.

Vaikka sisällöllisesti lyhyen matematiikan aiheet näyttävät hyvin samalta kuin pitkässä matematiikassa, on aiheiden käsittelyssä selviä eroja. Lyhyessä matematiikassa ei esimerkiksi käytetä lause ja määritelmä -käsitteitä, lauseiden todistuksista puhumattakaan. Monet määritelmät on esitetty helppolukuisemmin. Kuten esimerkiksi Moodi 5 -kirjassa Geometrinen todennäköisyys on esitetty seuraavasti:

Tapahtuman A geometrinen todennäköisyys on

$$P(A) = \frac{\text{tapahtumalle A suotuisan joukon mitta}}{\text{koko perusjoukon mitta}}$$

Sama määritelmä Moodi 8 -kirjassa on taas esitetty seuraavasti:
Tapahtuman A geometrinen todennäköisyys on

$$P(A) = \frac{m(A)}{m(E)},$$

missä $m(A)$ on tapahtumalle A suotuisan geometrisen kuvion mitta ja $m(E)$ perusjoukkoa E kuvaavan geometrisen kuvion mitta.

Seuraavassa alaluvussa lähdetään konkreettisesti tutkimaan, millaisia ovat eroavuudet peruskoulun ja lukion välillä. Jätetään tästä tarkastelusta kuitenkin pois lukion lyhyen oppimäärän matematiikan vertailu.

1.4 Peruskoulun ja lukion eroavuudet esimerkein

Tässä luvussa käydään läpi sitä, miten todennäköisyyslaskennan erilaiset määritelmät ja tehtävät esitetään peruskoulussa ja lukiossa pitkässä matematiikassa. Tämän lisäksi tutkitaan, miten erilaisia opetusmateriaalit ja esimerkit ovat.

Todennäköisyyslaskentaan tutustuminen aloitetaan alakoulussa arkielämän esimerkeillä. Sanoma Pron Milli 5a -oppikirjan (2023) ensimmäinen tehtävä on hyvä kuvaamaan opetussuunnitelman tavoitteita. Tehtävässä on tarkoitus pohtia, onko jokin todennäköisyys varmaa, mahdollista, vai mahdotonta.

Merkitse X sen mukaan, kuinka varmasti asia voi tapahtua.

	Varmaa	Mahdollista	Mahdotonta
a. Saan lemmikiksi elävän yksisarvisen.	☐	☐	☐
b. Ajan lentävällä autolla.	☐	☐	☐
c. Täytän ensi vuonna 4 vuotta.	☐	☐	☐
d. Huomenna sataa lunta.	☐	☐	☐
e. Näen elävän mammutin.	☐	☐	☐
f. Käyn luistelemassa kesälomalla.	☐	☐	☐
g. Äitini on vanhempi kuin minä.	☐	☐	☐
h. Koulussamme on Suomen tuleva presidentti.	☐	☐	☐
i. Joulukuussa on 31 vuorokautta.	☐	☐	☐
j. Suomessa purkautuu tulivuori.	☐	☐	☐

Kuva 1.3: Sanoma Pro Milli 5a t.1 s.186

Tuhattaituri 6b -oppikirjassa (2018) käsitellään myös todennäköisyyslaskennan muita aiheita, mutta koska ne ovat hyvin samantapaisia kuin yläkoulussa, eikä opetussuunnitelman tavoitteissa niistä ole mainintaa, niin käydään esimerkkejä läpi seuraavaksi yläkoulun puolelta.

Tämän luvun määritelmät, lauseet esimerkit ja esimerkkien ratkaisut ovat peräisin tässä tutkielmassa käytetyistä oppikirjoista. Suurin osa yläkoulun materiaaleista on Sanoma Pron Kuutio K11 -oppikirjasta (2023) ja lukion materiaalit on pääosin Otavan Juuri 8 -oppikirjasta (2022). Mikäli tehtävät ovat muista kirjoista, on siitä erillinen maininta suluissa tehtävän vieressä.

1.4.1 Tuloperiaate

Aloitetaan aiheiden käsittely tuloperiaatteesta. Yläkoulun kirjoissa ei puhuta määritelmistä eikä lauseista, mutta tässä tutkielmassa käytetään niitä selkeyden vuoksi. Yläkoulussa tuloperiaate esitetään seuraavalla tavalla:

Lause 1.1. *Jos kokonaisuus muodostetaan useassa eri vaiheessa, kokonaisuuksien lukumäärä saadaan kertomalla eri vaihtoehtojen lukumäärät keskenään*

Seuraavaksi käydään muutama esimerkki siitä, millaisia tehtäviä tähän aiheeseen liittyy. Ensimmäinen esimerkki on itse keksitty, loput kaksi ovat Kuutio K11 (2023) kirjasta.

Esimerkki 1.2. Miljalla on viisi paitaa ja kolme hametta. Kuinka monta erilaista asuyhdistelmää Milja voi valita?
Ratkaisu: Yhdelle paidalle löytyy kolme erilaista hametta ja viidelle paidalle löytyy viisi kertaa kolme erilaista vaihtoehtoa eli $5 \cdot 3 = 15$. Milja voi siis valita 15 erilaista asuyhdistelmää.

Esimerkki 1.3. Alinalla on kuusi erilaista kirjainkorttia. Kuinka monta erilaista kolmen kirjaimen yhdistelmää Alina voi niistä muodostaa?
Ratkaisu: Ensimmäisen kirjaimen paikalle on kuusi eri vaihtoehtoa. Toisen kirjaimen paikalle on enää viisi vaihtoehtoa. Ja näin kolmannen kirjaimen paikalle on enää neljä vaihtoehtoa, sillä kaksi korteista on jo käytetty aiemmin. $6 \cdot 5 \cdot 4 = 120$ eli Alina voi muodostaa 120 erilaista yhdistelmää.

Esimerkki 1.4. Biologian kokeessa on kuusi tehtävää ja jokaisessa kolme eri vaihtoehtoa. Kuinka monta erilaista vastausriviä voi tehdä?
Ratkaisu: Tuloperiaatteen mukaan kerrotaan jokaisen vaiheen vaihtoehdot keskenään.

$$3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^6 = 729.$$

Eli vastausrivejä voi tehdä 729 riviä.

Nyt kun on käyty läpi, miten tuloperiaate opetetaan yläkoulussa, voidaan esitellä sama aihe lukiossa. Lukiossa tuloperiaate on annettu seuraavana lauseena:

Lause 1.5. *Kun valitaan k kertaa ja vaiheessa i vaihtoehtoja on n_i kappaletta, mahdollisia lopputuloksia on kaikkiaan $n_1 \cdot n_2 \cdot ... \cdot n_k$ kappaletta.*

Kuten huomataan, on yläkoulussa esitetty Lause huomattavasti yksinkertaisempi. Näin on myös esimerkkien kanssa. Seuraavat esimerkit ovat hyvin samantapaisia kuin yläkoulussa, mutta vaatimustaso niissä on korkeampi. Nämä kolme esimerkkiä on Sanoma Pron Moodi 8 -oppikirjasta. Ensimmäinen esimerkki eroaa yläkoulun esimerkistä siten, että siinä on useampia eri vaatekappaleita ja kenkiä. Haastetta lisää se, että esimerkissä täytyy ensin laskea kaikki paidat, housut ja kengät erikseen yhteen, ennen kuin voidaan laskea kaikkien asuyhdistelmien määrä.

Esimerkki 1.6. Miljalla on neljä T-paitaa ja kolme toppia sekä kahdet farkkushortsit ja kolmet pitkät housut. Kenkinä hänellä on tennarit ja kahdet sandaalit.

a) Kuinka monella tavalla Milja voi valita topin ja pitkät housut?
Ratkaisu: Topin voi valita kolmella tavalla ja pitkät housut myös kolmella tavalla eli $3 \cdot 3 = 9$.

b) Kuinka monta erilaista asuyhdistelmää Milja voi muodostaa?
Ratkaisu: Paitoja voi valita seitsemällä eri tapaa, housuja viidellä eri tapaa ja kengät kolmella tapaa eli kaikkia asuyhdistelmiä on yhteensä $7 \cdot 5 \cdot 3 = 105$.

Seuraavassa esimerkissä tilanne on samanlainen kuin yläkoulussa. Yläkoulun esimerkissä etsittiin kuudesta kirjainkortista kolmen kirjaimen yhdistelmää. Lukion esimerkissä kirjaimia on enemmän eli muodostuvasta laskusta tulee haastavampi, mutta myös yläkoulun oppilailla on tarvittava tieto tehtävän ratkaisemiseksi.

Esimerkki 1.7. Tietokoneohjelma muodostaa käyttäjälle viisikirjaimisen tunnisteen aakkosten 29 kirjaimen joukosta. Kuinka monta erilaista tunnistetta voidaan muodostaa, kun sama kirjain saa esiintyä tunnisteessa vain kerran?
Ratkaisu: Koska sama kirjain ei saa toistua, niin voidaan jokaisessa vaiheessa valita yksi kirjain vähemmän eli

$$29 \cdot 28 \cdot 27 \cdot 26 \cdot 25 = 14\,250\,600.$$

Myös viimeinen esimerkki on vastaava yläkoulun kolmannen esimerkin kanssa. Kuten huomataan, niin usein lukion esimerkeissä on enemmän vaihtoehtojen määriä kuin yläkoulussa. Myös seuraavan esimerkin ymmärtäminen ja ratkaiseminen pitäisi onnistua yläkoulun oppilaalta.

Esimerkki 1.8. Vakioveikkauksessa on 13 kohdetta, joista jokaisessa valitaan yksi seuraavista vaihtoehdoista: 1 = kotijoukkue voittaa, x = tasapeli tai 2 = vierasjoukkue voittaa. Kuinka monta erilaista veikkausriviä voidaan muodostaa?
Ratkaisu: Rivin täytössä on 13 peräkkäistä vaihetta, joissa on 3 vaihtoehtoa. Erilaisten rivien määrä on siis

$$3^{13} = 159432$$

1.4.2 Permutaatio

Yläkoulussa ei vielä opita permutaation käsitettä, vaan kirjoissa puhutaan siitä, kuinka monella tavalla asiat voidaan järjestää. Permutaatio on esitetty yläkoulussa seuraavasti:

Lause 1.9.

$$n \textit{ kertoma tarkoittaa tuloa: } \quad n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1.$$

Lisäksi

n alkiota voidaan järjestää jonoon $n!$ eri tavalla.

Käydään heti perään esimerkki permutaatiosta yläkoulussa. Tämä esimerkki on itse keksitty.

Esimerkki 1.10. Penkille istuu 8 oppilasta. Kuinka monessa eri järjestyksessä he voivat istua vierekkäin?
Ratkaisu: Oppilaat voivat istua $8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot \ldots \cdot 1 = 40320$ järjestyksessä.

Lukiossa permutaatio esitetään lähes samalla tapaa kuin yläkoulussa:

Lause 1.11. *Jos joukossa on n alkiota, joukon permutaatioiden lukumäärä on*

$$n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1.$$

Toisin sanoen n alkion joukolla on $n!$ permutaatiota.

Lisäksi pitkässä matematiikassa käydään läpi k-permutaatio, joka ei tule esiin yläkoulun oppikirjoissa. Jos joukosta asetetaan jonoon vain k alkiota ja muut jätetään pois, puhutaan joukon k-permutaatiosta:

Lause 1.12. *Jos joukossa on n alkiota ja $k \leq n$, niin joukon k-permutaatioiden lukumäärä on*

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}.$$

Yläkoulussa käytävä esimerkki sopii hyvin myös lukion perustehtäväksi, ja se on samantapainen kuin seuraavan esimerkin a) -kohta. Kuitenkin lukiossa käytävät tehtävät muuttuvat haastavammaksi ja tästä hyvä esimerkki on alla olevan tehtävän b) -kohta. Tällaista tehtävää ei yläkoulun oppikirjoista löydy. Yläkoulun oppikirjoissa samantyylisessä tehtävässä on esimerkiksi annettu valmiiksi erilaisten lottorivien lukumäärän lauseke, josta pitää laskea vain lauseke auki.

Esimerkki 1.13. (Moodi 8)
Hiihtokilpailun viestijoukkueessa on kuusi kilpailijaa: neljä lukiolaista ja kaksi peruskoululaista.

a) Kuinka monta hiihtojärjestystä voidaan muodostaa?
Ratkaisu: Kuusi kilpailijaa voidaan järjestää $6! = 720$ tavalla.

b) Kuinka monta sellaista hiihtojärjestystä voidaan muodostaa, joissa lukiolaiset hiihtävät peräkkäin?
Ratkaisu: Lukiolaiset voivat hiihtää neljä ensimmäistä, neljä keskimmäistä tai neljä viimeistä osuutta. Lukiolaisten hiihtämät osuudet voidaan valita kolmella tavalla. Lisäksi neljän lukiolaisen järjestys voidaan valita $4! = 24$ eri tavalla ja peruskoululaisten järjestys $2! = 2$ tavalla. Eli hiihtojärjestyksiä on kaikkiaan

$$3 \cdot 4! \cdot 2! = 144.$$

1.4.3 Klassinen todennäköisyys

Siirrytään seuraavaksi erilaisten todennäköisyyksien laskemiseen. Ensimmäiseksi käsitellään klassisen todennäköisyyden määritelmä. Käydään läpi yläkoulussa opetettava määritelmä, sekä esimerkkejä.

Määritelmä 1.14. Tapahtuman A klassinen todennäköisyys on suotuisten ja kaikkien alkeistapausten osamäärä.

$$P(A) = \frac{\text{suotuisat alkeistapaukset}}{\text{kaikki alkeistapaukset}}$$

Esimerkki 1.15. (Ääretön Tilastot ja todennäköisyys)
Noppaa heitetään kerran. Millä todennäköisyydellä saadaan silmäluvuksi 1 tai 6?
Ratkaisu: Alkeistapauksia on kuusi. Haluttuja alkeistapauksia on kaksi.

$$P(\text{silmäluku on 1 tai 6}) = \frac{2}{6} = \frac{1}{3}.$$

Esimerkki 1.16. Millä todennäköisyydellä kahta noppaa heitettäessä silmälukujen summa on 10?
Ratkaisu: Alkeistapauksia on $6 \cdot 6 = 36$. Suotuisia tapauksia on 3.

$$P(10) = \frac{3}{36} = \frac{1}{12}.$$

Klassisen todennäköisyyden määritelmä näyttää hyvin samanlaiselta niin yläkoulussa kuin lukiossa.

Määritelmä 1.17. Kun kaikkien alkeistapausten lukumäärä on n ja tapahtuma A sisältää k alkeistapausta, niin tapahtuman A klassinen todennäköisyys on

$$P(A) = \frac{\text{suotuisien alkeistapausten lukumäärä}}{\text{kaikkien alkeistapausten lukumäärä}} = \frac{k}{n}.$$

Yläkoulussa käydyissä tehtävissä nopat ovat aina tavallisia kuusisivuisia noppia. Lukiossa haastetta on lisätty käyttämällä tehtävissä myös noppia, joissa on sekä enemmän ja vähemmän sivuja kuin tavallisessa nopassa. Alla olevasta esimerkistä huomataan, miten ero näkyy käytännössä.

Esimerkki 1.18. (Moodi 8)
Milja heittää kahta noppaa, joista toinen on tavallinen ja toinen nelisivuinen noppa. Millä todennäköisyydellä silmälukujen summa on 8?
Ratkaisu: Alkeistapauksia on $6 \cdot 4$. Suotuisia tapauksia on 3. Eli

$$P(8) = \frac{3}{24} = \frac{1}{8}.$$

1.4.4 Komplementtisääntö

Tässä alaluvussa käsitellään komplementtisääntö eli vastatapahtuman todennäköisyys. Yläkoulussa ei puhuta komplementtisäännöstä vaan pelkästään vastatapahtuman todennäköisyydestä. Yläkoulussa opetetaan, että tapahtuman ja sen vastatapahtuman todennäköisyyksien summa on aina 100%. Tapahtuman todennäköisyys saadaan, kun vähennetään 100 prosentista vastatapahtuman todennäköisyys.

Esimerkki 1.19. (Ääretön Tilastot ja Todennäköisyys) Jalankulkijan liikennevalo on punainen 75 %:n todennäköisyydellä. Saara saapuu liikennevaloihin. Millä todennäköisyydellä liikennevalo ei ole punainen?
Ratkaisu: Liikennevalo ei ole punainen todennäköisyydellä $100\% - 75\% = 25\%$.

Lukiossa tulee esiin komplementin käsite: Tapahtuman A vastatapahtuma eli komplementti $\overline{A}$ sisältää ne alkeistapaukset, joita A ei sisällä.

Lause 1.20. *Tapahtumalle A ja sen vastatapahtumalle $\overline{A}$ pätee*

$$P(A) + P(\overline{A}) = 1 \quad \textit{eli} \quad P(A) = 1 - P(\overline{A}).$$

Lukion ja yläkoulun esimerkit eroavat huomattavasti, sillä yläkoulussa vastatapahtuman todennäköisyyksiä lasketaan vain yllä olevan esimerkin kaltaisesti prosenttien avulla. Lukiossa tehtävissä taas päätellään itse mikä on tapahtuman komplementti ja lasketaan sen todennäköisyys komplementtisäännön avulla.

Esimerkki 1.21. Millä todennäköisyydellä korttipakasta satunnaisesti valittu kortti

a) on kuningas?
Ratkaisu: Koska kuninkaita on 4, suotuisia alkeistapauksia on 4. Tapahtuman $B =$ ”kortti on kuningas” todennäköisyys on

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

b) ei ole kuningas?
Ratkaisu: Edellisen kohdan mukaan tapahtuman B = ”kortti on kuningas” todennäköisyys on $P(B) = \frac{1}{13}$. Tämän vastatapahtuma on $\overline{B} =$ ”kortti ei ole kuningas”. Saadaan

$$P(\overline{B}) = 1 - P(B) = 1 - \frac{1}{13} = \frac{12}{13}.$$

1.4.5 Tilastollinen todennäköisyys

Myös tilastollisen todennäköisyyden laskemista harjoitellaan sekä yläkoulussa että lukiossa. Esitetään seuraavaksi tilastollisen todennäköisyyden määritelmä yläkoulussa ja käydään läpi kolikonheittoon liittyvä esimerkki.

Määritelmä 1.22. Tapahtuman A tilastollinen todennäköisyys saadaan jakamalla tapahtuman A esiintymislukumäärä kaikkien tapausten määrällä.

$$P(A) = \frac{\text{tapahtuman A esiintymislukumäärä}}{\text{kaikkien tapausten määrä}}$$

Esimerkki 1.23. Kolikonheitossa saatiin 231 klaavaa ja 228 kruunaa. Laske kruunan tilastollinen todennäköisyys.
Ratkaisu:

$$P(\text{kruuna}) = \frac{228}{231 + 228} = \frac{228}{459} = \frac{76}{153}.$$

Tällaisten laskujen lisäksi harjoitustehtävissä on paljon erilaisia valmiita tilastotaulukkoja, joista lasketaan tilastollisia todennäköisyyksiä. Samanlaisia tehtäviä on myös lukiossa, ja kuten alla olevasta määritelmästä huomataan, myös tilastollisen todennäköisyyden määritelmä näyttää hyvin samalta kuin yläkoulussa.

Määritelmä 1.24. Kun tilastossa kaikkien havaintojen lukumäärä on n ja tapahtumalle A suotuisien havaintojen lukumäärä k, niin tapahtuman A tilastollinen todennäköisyys on

$$P(A) = \frac{k}{n}.$$

Tilastollisen todennäköisyyden osalta lukion tehtävät eivät juurikaan eroa haastavuudeltaan yläkoulusta. Suurin eroavaisuus tulee siinä, ettei yläkoulun tehtävissä esiinny vastatapahtumien todennäköisyyksien laskemista, kuten seuraavassa esimerkissä.

Esimerkki 1.25. Koripalloilija on heittänyt kauden aikana pelaamissaan peleissä 72 vapaaheittoa, joista 66 on onnistunut.
a) Millä todennäköisyydellä hänen seuraavasta vapaaheitosta tulee kori?
Ratkaisu:

$$P(\text{tulee kori}) = \frac{66}{72} = \frac{11}{12}.$$

b) Entä ei tule koria?
Ratkaisu: Tapahtuma voidaan laskea komplementtisäännön avulla.

$$P(\text{ei tule koria}) = 1 - P(\text{tulee kori}) = 1 - \frac{11}{12} = \frac{1}{12}.$$

1.4.6 Geometrinen todennäköisyys

Yläkoulun oppikirjoissa ei käydä geometrisen todennäköisyyden määritelmää, mutta alakoulun Tuhattaituri 6b -oppikirjassa (2018) se opetetaan seuraavien esimerkkien avulla.

Esimerkki 1.26. Tikkataulun sektorit ovat keskenään yhtä suuria. Millä todennäköisyydellä tikka osuu punaiselle alueelle?

Ratkaisu:
Sektoreita on yhteensä kahdeksan kappaletta, joista punaisia sektoreita on kolme. Punaisen alueen todennäköisyys on siis $\frac{3}{8}$.

Yllä olevan esimerkin kaltaisia tehtäviä on myös yläkoulun oppikirjoissa, vaikka niistä ei erikseen puhuta geometrisenä todennäköisyytenä.

Esimerkki 1.27. Tien pituus on 100 km. Sadealueen pituus on 60 km. Millä todennäköisyydellä Ullan auto on sadealueella?
Ratkaisu:

$$\frac{60\text{km}}{100\text{km}}$$

$$\frac{60}{100} = \frac{6}{10} = \frac{3}{5}.$$

Eli Ullan auto on sadealueella todennäköisyydellä $\frac{3}{5}$.

Alakoulussa ei siis käydä läpi minkäänlaista määritelmää, joten lukioon mentäessä on suurempi harppaus kuin muissa todennäköisyyden laskuissa, ja osalle opiskelijoista aihe on aivan uusi. Tämän lisäksi lukiomatematiikassa esiintyy integraalien laskemista, jota ei peruskoulussa käsitellä lainkaan. Juuri 8 -oppikirjan määritelmä geometriselle todennäköisyydelle on seuraava:

Määritelmä 1.28. Tapahtuman A geometrinen todennäköisyys on

$$P(A) = \frac{\text{tapahtumaa } A \text{ vastaavan kuvion osan geometrinen mitta}}{\text{koko kuvion geometrinen mitta}}.$$

Sanoma Pron Moodi 8 -oppikirjassa on jotkin määritelmät esitetty eri tavalla, kuin Juuri 8 -kirjassa. Seuraava määritelmä on esitetty tarkemmin, mikä on hyvä etenkin niille, jotka jatkavat matematiikan opiskelua pidemmälle.

Määritelmä 1.29. Tapahtuman A geometrinen todennäköisyys on

$$P(A) = \frac{m(A)}{m(E)},$$

missä $m(A)$ on tapahtumalle A suotuisan geometrisen kuvion mitta ja $m(E)$ on perusjoukkoa E kuvaavan geometrisen kuvion mitta.

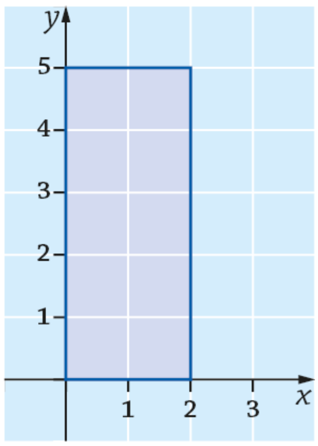
Myös esimerkkitehtävät ovat hyvin erilaisia alakoulun ja lukion välillä. Siinä missä alakoulussa tehtävät ovat hyvin yksinkertaisia, niin lukiossa tehtävät ovat vaikeustasoltaan huomattavasti haastavampia.

Esimerkki 1.30. (Moodi 8)
Metro lähtee asemalta 4 minuutin välein ja pysähtyy asemalla aina 20 sekunnin ajaksi ennen lähtöä. Matkustaja saapuu satunnaisesti asemalle tietämättä metron aikataulua. Millä todennäköisyydellä matkustaja pääsee suoraan metroon?
Ratkaisu: Metro lähtee hetkellä L_1, seuraava metro saapuu S ja lähtee 20 sekuntia myöhemmin hetkellä L_2. Jotta matkustaja pääsee suoraan metroon, hänen on saavuttava asemalle 20 sekunnin aikana, minkä metro odottaa asemalla. Suotuisan ajanjakson pituus on siis $20s$. Koko perusjoukkoa kuvaavan ajanjakson pituus on $4\text{min} = 4 \cdot 60s = 240s$.
Lasketaan todennäköisyys

$$P(\text{pääsee suoraan metroon}) = \frac{20s}{240s} = \frac{1}{12}.$$

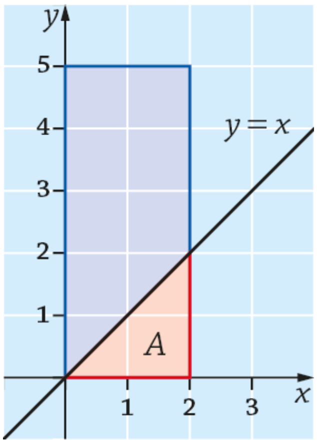
Lisäksi alakoulussa tehtävät ovat sellaisia, joissa oppilaiden ei itse tarvitse hahmotella kuvia, vaan ne on annettu valmiiksi. Lukiossa tehtävät vaativat usein kuvan hahmottamista tilanteesta, kuten seuraavassa esimerkissä.

Esimerkki 1.31. Olkoon x satunnainen luku väliltä $[0, 2]$ ja y satunnainen luku väliltä $[0, 5]$. Millä todennäköisyydellä $x > y$?
Ratkaisu: Lukuparista (x, y) muodostuu suorakulmion muotoinen alue. Suorakulmion kanta on 2 ja korkeus 5, joten suorakulmion pinta-ala eli koko kuvion geometrinen mitta on $2 \cdot 5 = 10$.



Kuva 1.4: Juuri 8 (LOPS21) s.87

Kun lisätään muodostuneeseen kuvioon suora $y = x$; suoran alapuolella $y < x$. Tapahtumaa vastaa siis se suorakulmion osa, joka on suoran $y = x$ alapuolella. Kyseessä on suorakulmainen kolmio, jonka kanta ja korkeus ovat 2. Tämän alueen pinta-ala eli vastaavan alueen geometrinen mitta on $\frac{2 \cdot 2}{2} = 2$.



Kuva 1.5: Juuri 8 (LOPS21) s.87

Näin ollen todennäköisyys

$$P(x > y) = \frac{2}{10} = \frac{1}{5}.$$

1.4.7 Kertolaskusääntö ja riippumattomuus

Tässä alaluvussa käytävää kertolaskusääntöä puhutellaan yläkoulussa peräkkäisinä tapahtumina. Todennäköisyys sille, että peräkkäiset tapahtumat toteutuvat kumpikin, saadaan kertomalla tapausten todennäköisyydet keskenään. Lisäksi opitaan riippuvuuden määritelmä, joka on hyvin yksinkertainen:

Määritelmä 1.32. Jos tapahtumat eivät vaikuta toisiinsa, ne ovat toisistaan riippumattomia. Jos edellinen tapahtuma vaikuttaa seuraavaan todennäköisyyteen ne ovat toisistaan riippuvia.

Käydään läpi kaksi esimerkkiä peräkkäisistä tapahtumista.

Esimerkki 1.33. (Ääretön Tilastot ja Todennäköisyys)
Milja heittää kolikkoa. Millä todennäköisyydellä hän saa kaksi kruunaa peräkkäin?
Ratkaisu: Todennäköisyys saada kruuna on $\frac{1}{2}$. Nyt

$$P(\text{kaksi kruunaa peräkkäin}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Esimerkki 1.34. Laatikossa on 3 sinistä, 4 punaista ja 6 keltaista palloa. Laatikosta nostetaan pallo kolme kertaa. Mikä on todennäköisyys sille, että jokaisella kerralla nostetaan punainen pallo
a) kun pallot laitetaan takaisin laatikkoon noston jälkeen?
Ratkaisu: Laatikossa on yhteensä 13 palloa. Todennäköisyys saada punainen pallo jokaisella nostolla on $\frac{4}{13}$ eli todennäköisyys

$$P(\text{kaikki punaisia}) = \frac{4}{13} \cdot \frac{4}{13} \cdot \frac{4}{13} = \frac{64}{2197}.$$

b) kun palloja ei laiteta takaisin laatikkoon noston jälkeen?
Ratkaisu: Nyt

$$P(\text{1. pallo on punainen}) = \frac{4}{13}.$$

$$P(\text{2. pallo on punainen}) = \frac{3}{12}.$$

$$P(\text{3. pallo on punainen}) = \frac{2}{11}.$$

Eli todennäköisyys

$$P(\text{kaikki punaisia}) = \frac{4}{13} \cdot \frac{1}{4} \cdot \frac{2}{11} = \frac{2}{143}.$$

Lukiossa riippumattomuuden määritelmä taas esitetään seuraavasti:

Määritelmä 1.35. Jos tapahtuman A todennäköisyys ei riipu tapahtuman B toteutumisesta eikä toisin päin, sanotaan, että tapahtumat A ja B ovat riippumattomat. Tällöin myös niiden vastatapahtumat $\overline{A}$ ja $\overline{B}$ ovat riippumattomat.

Riippumattomien tapahtumien kertolaskusäännön avulla saadaan selvitettyä ovatko tapahtumat A ja B riippumattomia:

Lause 1.36. *Tapahtumat A ja B ovat riippumattomat, kun*

$$P(A\,ja\,B) = P(A) \cdot P(B).$$

Riippumattomien tapahtumien kertolaskusäännön lisäksi Moodi 8 -kirjassa opetetaan, että riippuvuus voidaan määritellä täsmällisesti ehdollisen todennäköisyyden avulla. Ehdollisen todennäköisyyden määritelmää ei opeteta yläkoulussa, eikä Juuri 8 -kirjassakaan ole sille erikseen määritelmää. Ehdollinen todennäköisyys esiintyy kuitenkin tehtävissä useaan otteeseen. Siispä seuraavaksi esitellään ehdollinen todennäköisyys, niin kuin se on Moodi 8 -kirjassa määritelty:

Määritelmä 1.37. Olkoon $P(A) > 0$. Tapahtuman B todennäköisyys ehdolla A on

$$P(B \mid A) = \frac{P(B \,\text{ja}\, A)}{P(A)}.$$

Nyt voidaan määritellä riippuvuus seuraavasti ehdollisen todennäköisyyden avulla.

Määritelmä 1.38. Tapahtumat A ja B ovat riippumattomia, jos

$$P(B) = P(A \mid B).$$

Ensimmäinen esimerkki on jälleen samankaltainen yläkoulussa ja lukiossa. Tämän jälkeen tehtävät muuttuvat lukiossa haastavammiksi. Lukiossa esimerkiksi noppiin liittyvissä tehtävissä voidaan laskea todennäköisyyksiä siten, että nopan silmäluku on vähintään tai enintään neljä. Yläkoulussa ei tällaisia laskuja esiinny, vaan noppiin liittyvissä tehtävissä lasketaan, millä todennäköisyydellä molempien noppien silmäluvut ovat esimerkiksi kuutosia.

Esimerkki 1.39. Heitetään noppaa kolmesti peräkkäin. Millä todennäköisyydellä saadaan
a) kolme ykköstä
b) kolme paritonta silmälukua?
Ratkaisu:
a) Riippumattomien tapahtumien kertolaskusäännön mukaan

$$\begin{aligned}
P(\text{kolme ykköstä}) &= P(\text{ykkönen JA ykkönen JA ykkönen}) \\
&= P(\text{ykkönen})P(\text{ykkönen})P(\text{ykkönen}) \\
&= \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{216}.
\end{aligned}$$

b) Riippumattomien tapahtumien kertolaskusäännön mukaan

$$\begin{aligned}
P(\text{kolme paritonta}) &= P(\text{pariton JA pariton JA pariton}) \\
&= P(\text{pariton})P(\text{pariton})P(\text{pariton}) \\
&= \frac{3}{6} \cdot \frac{3}{6} \cdot \frac{3}{6} = \frac{1}{8}.
\end{aligned}$$

Yleisen kertolaskusäännön avulla opitaan myös laskemaan todennäköisyys saada seitsemän oikein tai ei yhtään oikein lotossa. Tämän lisäksi komplementtisäännön avulla voitaisiin laskea tilanteet, joissa ainakin yksi numero on oikein ja ainakin yksi on väärin. Kuten aiemmin mainittiin, komplementtisääntöön liittyviä tehtäviä ei esiinny yläkoulussa myöskään tässä aiheessa.

Esimerkki 1.40. Lotossa arvotaan seitsemän numeroa 1-40. Kuinka suurella todennäköisyydellä pelaajan lottorivin seitsemästä numerosta
a) kaikki ovat oikeita
Ratkaisu: Yleisen kertolaskusäännön mukaan

$$P(7 \text{ oikein}) = \frac{7}{40} \cdot \frac{6}{39} \cdot \frac{5}{38} \cdot \frac{4}{37} \cdot \frac{3}{36} \cdot \frac{2}{35} \cdot \frac{1}{34} = \frac{1}{18643560}.$$

b) kaikki ovat vääriä?
Ratkaisu:

$$P(7 \text{ väärin}) = \frac{33}{40} \cdot \frac{32}{39} \cdot \frac{31}{38} \cdot \frac{30}{37} \cdot \frac{29}{36} \cdot \frac{28}{35} \cdot \frac{27}{34} = 0,22914...$$

Näin saatiin käytyä määritelmiä ja esimerkkejä aiheista, joita käsitellään sekä peruskoulussa että lukiossa. Seuraavaksi siirrytään aiheisiin, joita opetetaan vain lukiossa.

1.4.8 Yhteenlaskusääntö

Aloitetaan lukioaiheiden läpikäynti yhteenlaskusäännöllä. Juuri 8 ja Moodi 8 -oppikirjoissa opetetaan sekä erillisten tapahtumien yhteenlaskusääntö että yleinen yhteenlaskusääntö. Nämä asiat esitetään kirjassa eri järjestyksessä. Juuri 8 -kirjassa käsitellään ensin yleinen yhteenlaskusääntö, kun taas Moodi 8 -kirjassa käsitellään ensin erillisten tapahtumien yhteenlaskusääntö.

Jos tapahtuma A ja B on mahdoton eli ei sisällä yhtään alkeistapausta, sanotaan, että A ja B ovat erilliset eli toisensa poissulkevat tapahtumat. Erillisille tapahtumille pätee $P(A \,\text{ja}\, B) = 0$ ja näille tapahtumille yhteenlaskusääntö on seuraava:

Aksiooma 1.41. Jos tapahtumat ovat erilliset, niin

$$P(A \,\text{tai}\, B) = P(A) + P(B)$$

Edellisen aksiooman avulla voidaan todistaa yleinen yhteenlaskusääntö. Yleisellä yhteenlaskusäännöllä saadaan laskettua todennäköisyys, että ainakin toinen tapahtumista A ja B tapahtuu:

Lause 1.42. *Todennäköisyys, että ainakin toinen tapahtumista A ja B tapahtuu, on*

$$P(A \,\textit{tai}\, B) = P(A) + P(B) - P(A \,\textit{ja}\, B)$$

Lasketaan seuraavaksi todennäköisyyksiä yhteenlaskusäännön ja riippumattomien tapahtumien kertolaskusäännön avulla.

Esimerkki 1.43. Oletetaan, että jokainen viikonpäivistä on yhtä mahdollinen syntymiselle. Kuinka suurella todennäköisyydellä eri vuosina syntyneistä veljeksistä
a) Matti on syntynyt maanantaina ja Sulevi sunnuntaina
Ratkaisu: Koska jokainen viikonpäivistä on yhtä todennäköinen, niin

$$P(\text{lapsi on syntynyt maanantaina}) = P(\text{lapsi on syntynyt sunnuntaina}) = \frac{1}{7}$$

Veljesten syntymien viikonpäivät eivät riipu toisistaan, joten

$$P(\text{Matti maanantaina ja Sulevi sunnuntaina}) = \frac{1}{7} \cdot \frac{1}{7} = \frac{1}{49}.$$

b) Vain toinen on syntynyt maanantaina
Ratkaisu: Tapahtuma koostuu kahdesta erillisestä osasta: "Matti on syntynyt maanantaina ja Sulevi ei"sekä "Sulevi on syntynyt maanantaina ja Matti ei". Yhteenlaskusäännön ja riippumattomien tapahtumien kertolaskusäännön avulla

$$\begin{aligned}
&P(\text{vain toinen maanantaina})\\
&= P(\text{Matti ma}) \cdot P(\text{Sulevi ei ma}) + P(\text{Sulevi ma}) \cdot P(\text{Matti ei ma})\\
&= \tfrac{1}{7} \cdot \left(1 - \tfrac{1}{7}\right) + \tfrac{1}{7} \cdot \left(1 - \tfrac{1}{7}\right)\\
&= \frac{12}{49}.
\end{aligned}$$

c) kumpikaan ei ole syntynyt maanantaina?
Ratkaisu: Riippumattomuuden perusteella

$$\begin{aligned}
&P(\text{kumpikaan ei ole syntynyt maanantaina})\\
&= P(\text{Matti ei ma}) \cdot P(\text{Sulevi ei ma})\\
&= \left(1 - \tfrac{1}{7}\right) \cdot \left(1 - \tfrac{1}{7}\right)\\
&= \frac{36}{49}.
\end{aligned}$$

1.4.9 Kombinaatiot

Tässä alaluvussa käsitellään kombinaatioita eli osajoukkoja. k-kombinaatioksi kutsutaan k-jäsenistä joukkoa, jossa järjestyksellä ei ole merkitystä.

Lause 1.44. *Jos joukossa on n alkiota ja $k \leq n$, niin joukon k-kombinaatioiden lukumäärä on*

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

Käytetään seuraavaksi k-kombinaatioiden lukumäärän lauseketta hyödyksi laskiessa seuraavat kaksi esimerkkiä.

Esimerkki 1.45. Kuinka monella tavalla tavallisesta 52 kortin korttipakasta voidaan nostaa viiden kortin joukko eli niin sanottu viiden kortin käsi?
Ratkaisu:

$$\binom{52}{5} = \frac{52!}{5! \cdot (52-5)!} = 2598960.$$

Esimerkki 1.46. (Moodi 8)
Hannalla on viisi pinkkiä, neljä vihreää ja kaksi harmaata Kivi-lyhtyä. Kuinka monella tapaa hän voi asettaa ne riviin, kun otetaan huomioon vain lyhtyjen väri?
Ratkaisu: Sijoitetaan ensin viisi pinkkiä lyhtyä. 11 paikasta voidaan valita viisi paikkaa $\binom{11}{5}$ tavalla. Sijoitetaan seuraavaksi neljä vihreää lyhtyä. Vapaana olevista paikoista voidaan valita neljä paikkaa $\binom{6}{4}$ tavalla. Vapaana on vain kaksi paikkaa, joten paikat voidaan valita vain yhdellä tavalla. Tuloperiaatteen mukaan lyhdyt voidaan siis asettaa riviin $\binom{11}{5} \cdot \binom{6}{4} \cdot 1 = 6930$ tavalla.

1.4.10 Diskreetti jakauma

Seuraavaksi siirrytään diskreettiin jakaumaan. Määritellään diskreetti satunnaismuuttuja:

Määritelmä 1.47. Satunnaismuuttuja tarkoittaa satunnaiskokeen tuloksesta tietyn säännön mukaan määräytyvää lukua. Satunnaismuuttuja on diskreetti, jos se saa vain yksittäisiä arvoja.

Diskreetin satunnaismuuttujan arvoihin x_i liittyviä todennäköisyyksiä $p_i = P(X = x_i)$ kutsutaan pistetodennäköisyyksiksi:

Lause 1.48. *Olkoot satunnaismuuttujan X mahdolliset arvot $x_1, x_2, \ldots, x_n$. Tällöin satunnaismuuttujan jakauman muodostavien pistetodennäköisyyksien $p_i = P(X = x_i)$ summa on 1 eli $p_1 + p_2 + \cdots + p_n = 1$.*

Esimerkki 1.49. Pelaaja nostaa korttipakasta yhden kortin. Korteilla $1-6$ hän saa yhden euron, korteilla $7-10$ hän saa kolme euroa, ja kuvakorteilla $(11-13)$ hän menettää kymmenen euroa. Satunnaismuuttuja X on rahasumma, jonka pelaaja voittaa yhdellä kortilla. Muodosta satunnaismuuttujan X jakauma.
Ratkaisu: Pelaajan voittosumman mahdolliset arvot ovat $1, 3$ ja -10 euroa.

Nämä ovat siis satunnaismuuttujan X mahdolliset arvot. Satunnaismuuttujan X jakauman muodostamiseksi tarvitaan eri arvojen todennäköisyydet.

$$P(X=1) = P(\text{kortti}\, 1-6) = \frac{4\cdot 6}{52} = \frac{6}{13}.$$

$$P(X=3) = P(\text{kortti}\, 7-10) = \frac{4\cdot 4}{52} = \frac{4}{13}.$$

$$P(X=-10) = P(\text{kuvakortti}) = \frac{4\cdot 3}{52} = \frac{3}{13}.$$

1.4.11 Diskreetin jakauman odotusarvo

Edellisessä alaluvussa saatiin käsiteltyä diskreetin satunnaismuuttujan määritelmä, joten voidaan siirtyä laskemaan diskreetin jakauman odotusarvoja ja keskihajontoja. Käydään ensin odotusarvon määritelmä ja heti perään keskihajonnan määritelmä.

Määritelmä 1.50. Olkoot satunnaismuuttujan X mahdolliset arvot $x_1, x_2, \ldots, x_n$ ja pistetodennäköisyydet $p_i = P(X = x_i), i = 1, \ldots, n$. Tällöin satunnaismuuttujan X odotusarvo $E(X)$ on luku

$$E(X) = p_1x_1 + p_2x_2 + \cdots + p_nx_n = \sum_{i=1}^{n} p_ix_i.$$

Määritelmä 1.51. Olkoot satunnaismuuttujan X mahdolliset arvot $x_1, x_2, \ldots, x_n$ ja pistetodennäköisyydet $p_i = P(X = x_i), i = 1, \ldots, n$. Tällöin satunnaismuuttujan keskihajonta D(X) on luku

$$D(X) = \sqrt{p_1(x_1-\mu)^2 + \cdots + p_n(x_n-\mu)^2} = \sqrt{\sum_{i=1}^{n} p_i(x_i-\mu)^2},$$

jossa μ on satunnaismuuttujan X odotusarvo.

Käydään seuraavaksi esimerkki, jonka avulla opitaan laskemaan odotusarvoja ja keskihajontoja.

Esimerkki 1.52. (Moodi 8)
Kolikkopelissä heitetään kahta tavallista kolikkoa. Peliin osallistuminen maksaa 2 euroa. Jos tulee kaksi kruunaa, pelaaja saa 7 euroa. Muissa tapauksissa pelaaja ei saa mitään. Olkoon satunnaismuuttuja X : "pelaajan yhdellä pelikierroksella saama voitto". Laske satunnaismuuttujan X odotusarvo ja

keskihajonta.
Ratkaisu: Satunnaismuuttujalla X on kaksi mahdollista arvoa.

$$7€ - 2€ = 5€$$

$$0€ - 2€ = -2€$$

Määritetään satunnaismuuttujan jakauma:

$$\begin{array}{rcl} 5€ & | & \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \\ -2€ & | & 1 - \frac{1}{4} = \frac{3}{4}. \end{array}$$

Lasketaan odotusarvo.

$$E(X) = \tfrac{1}{4} \cdot 5 + \tfrac{3}{4} \cdot (-2) = -\tfrac{1}{4} = -0,25\,(€)$$

Yhdellä pelikierroksella pelaajan voiton odotusarvo on $-0,25€$, eli hän häviää keskimäärin $0,25€$ jokaisella kierroksella.
Lasketaan satunnaismuuttuja X keskihajonta.

$$D(X) = \sqrt{\tfrac{1}{4} \cdot \left(5 - \left(-\tfrac{1}{4}\right)\right)^2 + \tfrac{3}{4} \cdot \left(-2 - \left(-\tfrac{1}{4}\right)\right)^2} = 3,03108...$$

Kierrosvoiton keskihajonta on siis noin $3,03€$.

1.4.12 Jatkuva jakauma

Tämän alaluvun sisällöt ovat lukion pitkän matematiikan valtakunnallisesta valinnaisesta moduulista 12. Määritelmät, lauseet, esimerkit ja esimerkkien ratkaisut ovat kirjasta Juuri 12 Analyysi ja jatkuva jakauma (2022). Aloitetaan aiheiden käsittely tiheysfunktion määritelmällä ja siihen liittyvällä esimerkillä.

Määritelmä 1.53. Funktio f on satunnaismuuttujan X tiheysfunktio, jos se on jatkuva mahdollisia yksittäisiä kohtia lukuun ottamatta ja jos

1. $f(x) \geq 0$ kaikilla x
2. $P(a \leq X \leq b) = \displaystyle\int_a^b f(x)\,dx$ aina kun $a \leq b$.

Esimerkki 1.54. Erään laitteen virheetön toiminta-aika (vuosina) on satunnaismuuttuja X, jonka tiheysfunktio on

$$f(x) = \begin{cases} \frac{1}{36}(-x^2 + 6x), & \text{kun} \quad 0 \leq x \leq 6 \\ 0, & \text{muuten.} \end{cases}$$

Millä todennäköisyydellä laite toimii virheettömästi yli kaksi vuotta?
Ratkaisu: Tapahtuman $X > 2$ todennäköisyys on epäoleellinen integraali $\int_2^\infty f(x)\,dx$. Välillä $[2, 6]$ tiheysfunktion f lauseke on $\frac{1}{36}(-x^2+6x)$, ja muuten se on 0. Niinpä

$$\begin{aligned}
P(X > 2) &= P(2 < X < \infty) \\
&= \int_2^\infty f(x)\,dx \\
&= \int_2^6 \tfrac{1}{36}(-x^2 + 6x)\,dx + \int_6^\infty 0\,dx \\
&= \frac{20}{27} = 0,740...
\end{aligned}$$

Laite toimii virheettömästi yli kaksi vuotta todennäköisyydellä $0,74$.

Seuraavaksi opitaan, miten lasketaan jatkuvasti jakautuneen satunnaismuuttujan odotusarvo ja keskihajonta.

Määritelmä 1.55. Jatkuvasti jakautuneen satunnaismuuttujan X odotusarvo on

$$E(X) = \mu = \int_{-\infty}^\infty xf(x)\,dx$$

Määritelmä 1.56. Jatkuvasti jakautuneen satunnaismuuttujan X keskihajonta on

$$D(X) = \sigma = \sqrt{\int_{-\infty}^\infty (x - \mu)^2 f(x)\,dx}.$$

Esimerkki 1.57. Erään laitteen virheetön toiminta-aika (vuosina) on satunnaismuuttuja X, jonka tiheysfunktio on

$$f(x) = \begin{cases} \frac{1}{36}(-x^2 + 6x), & \text{kun} \quad 0 \leq x \leq 6 \\ 0, & \text{muuten.} \end{cases}$$

Määritä virheettömän toiminta-ajan odotusarvo.
Ratkaisu: Lasketaan odotusarvo. Koska f eroaa nollasta vain välillä $[0, 6]$, niin

$$\begin{aligned}
E(X) &= \int_0^6 xf(x)\,dx \\
&= \int_0^6 x\frac{1}{36}(-x^2 + 6x)\,dx \\
&= 3.
\end{aligned}$$

Odotusarvon ja keskihajonnan jälkeen jatketaan määritelmien läpikäyntiä ja seuraavaksi vuorossa on satunnaismuuttujan kertymäfunktio. Tämän lisäksi opitaan laskemaan kertymäfunktio tiheysfunktion avulla.

Määritelmä 1.58. Satunnaismuuttujan X kertymäfunktio on

$$F(x) = P(X \leq x), x \in \mathbb{R}.$$

Esimerkki 1.59. Määritä satunnaismuuttujan X kertymäfunktio F, kun sen tiheysfunktio on

$$f(x) = \begin{cases} \frac{3}{4}(1-x^2), & \text{kun } -1 < x < 1 \\ 0, & \text{kun}\, x \leq -1 \,\text{tai}\, x \geq 1. \end{cases}$$

Ratkaisu: Lasketaan $F(X)$ erikseen tapauksissa $x \leq 1, -1 < x < 1, x \geq 1$.

- Kun $x \leq 1$, niin

$$F(X) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0.$$

- Kun $-1 < x < 1$, niin

$$\begin{aligned} F(X) &= \int_{-\infty}^{x} f(t)\,dt \\ &= \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{x} \tfrac{3}{4}(1-t^2)\,dt \\ &= 0 + \int_{-1}^{x} \left(\tfrac{3}{4} - \tfrac{3}{4}t^2\right)\,dt \\ &= \frac{3}{4}x - \frac{1}{4}x^3 + \frac{1}{2}. \end{aligned}$$

- Kun $x \geq 1$, niin

$$\begin{aligned} F(X) &= \int_{-\infty}^{x} f(t)\,dt \\ &= \int_{-\infty}^{-1} 0\,dt + \int_{-1}^{1} \tfrac{3}{4}(1-t^2)\,dt + \int_{1}^{x} 0\,dt \\ &= \frac{1}{2} + \frac{1}{2} = 1. \end{aligned}$$

Siis $F(x) = \begin{cases} 0, & \text{kun}\, x \leq -1 \\ \frac{3}{4}x - \frac{1}{4}x^3 + \frac{1}{2}, & \text{kun } -1 < x < 1 \\ 1, & \text{kun}\, x \geq 1. \end{cases}$

Tähän päättyy todennäköisyyslaskennan osio peruskoulussa ja lukiossa. Seuraavaksi lähdetään syventämään näitä tietoja tutustumalla todennäköisyyksiin yliopistomatematiikassa.

2 Todennäköisyyslaskenta yliopistossa

Tässä luvussa esitetyt tulokset ja todistukset ovat pääosin peräisin Introduction to probability and stochastic processes with application -kirjasta (Arunachalam ym., 2012). Käytettäessä muita lähteitä, niistä mainitaan erikseen.

2.1 Todennäköisyysavaruus

Ensimmäisessä alaluvussa aletaan rakentamaan todennäköisyysmitan käsitettä ja esitellään sen perusominaisuuksia. Kun tavallista noppaa heitetään kerran, lopputulosta ei voida ennustaa tarkasti, mutta tiedetään mahdollisten tulosten joukko. Tällaista koetta kutsutaan satunnaiseksi.

Määritelmä 2.1. Koe on satunnainen, jos sen tulosta ei voida määrittää etukäteen, mutta mahdolliset lopputulokset tunnetaan. Mahdollisten tuloksien joukkoa Ω kutsutaan otosavaruudeksi. Sen alkioita eli tulosvaihtoehtoja $\omega \in \Omega$ kutsutaan alkeistapauksiksi.

Esimerkki 2.2. Heitetään kolikkoa. Kolikon heitosta saadaan tulokseksi joko ”kruuna” $= Kr$ tai ”klaava” $= Kl$. Nyt siis $\Omega = \{Kr, Kl\}$.

Esimerkki 2.3. Heitetään noppaa kolme kertaa peräkkäin. Nopan heiton mahdolliset tulokset ovat muotoa (a, b, c) ja $a, b, c \in \{1, 2, 3, 4, 5, 6\}$. Tällöin $\Omega = \{(a, b, c) : a, b, c \in \{1, 2, 3, 4, 5, 6\}\}$.

Otosavaruuden alkiot voivat olla numeroita, vektoreita, symboleja ynnä muita sellaisia, ja ne määräytyvät tarkasteltavan tilanteen mukaan.

Määritelmä 2.4. Otosavaruutta kutsutaan diskreetiksi, jos se on joko äärellinen tai numeroituva. Satunnaiskoetta kutsutaan äärelliseksi, jos sen otosavaruus on äärellinen (diskreetti).

Tapahtumat ovat otosavaruuden osajoukkoja, mutta kaikki otosavaruuden osajoukot eivät ole tapahtumia. Tapahtumien kokoelmalla on oltava sigma-algebran rakenne joukossa Ω. Sigma-algebra on yksi mittateorian perustyökaluista. Sigma-algebrat ovat kokoelma joukkoja, jotka noudattavat tiettyjä sääntöjä. Lukion opetussuunnitelman perusteiden (2019) mukaan matematiikan opetuksessa on tarkoitus tutustuttaa opiskelija matematiikan peruskäsitteisiin, perusideoihin ja rakenteisiin. Sigma-algebran määritelmän ymmärtäminen vaatii vahvempaa matemaattista osaamista ja tämän vuoksi seuraava määritelmä on jätetty lukio-opintojen ulkopuolelle.

Määritelmä 2.5. Olkoon $\Omega \neq \emptyset$. Joukon Ω osajoukkojen kokoelmaa Γ kutsutaan σ-algebraksi joukossa Ω,

(i) jos $\Omega \in \Gamma$.

(ii) jos $A \in \Gamma$, niin $A^c \in \Gamma$.

(iii) jos $A_1, A_2, \dots \in \Gamma$, niin $\bigcup_{i=1}^{\infty} A_i \in \Gamma$.

Joukon Γ alkioita kutsutaan tapahtumiksi.

Esimerkiksi kolikon ja nopan heitossa kaikkien perusjoukon osajoukkojen muodostama kokoelma on σ-algebra. Seuraavaksi käydään esimerkit näistä tapauksista. Tämän jälkeen jatketaan erilaisiin σ-algebroihin tutustumista.

Esimerkki 2.6. Heitetään kolikkoa, eli $\Omega = \{Kr, Kl\}$. Tällöin $\Gamma = \{\emptyset, \Omega\}$ on triviaali σ-algebra joukossa Ω, koska $\Omega^c = \emptyset$ ja $\emptyset^c = \Omega$ ja koska yhdistämällä tyhjiä joukkoja saadaan edelleen tyhjä joukko ja yhdistämällä Ω itsensä kanssa saadaan Ω.

$G = \{\emptyset, Kr\}$ ei taas ole σ-algebra, sillä se ei täytä σ-algebran määritelmää. Määritelmän toinen ehto ei ole voimassa, sillä $\{Kr\}^c = \{Kl\}$ ja $\{Kl\} \notin G$.

Esimerkki 2.7. Heitetään kolmea noppaa peräkkäin, eli $\Omega = \{(a, b, c) : a, b, c \in \{1, 2, 3, 4, 5, 6\}\}$. Tällöin $\Gamma = \{\emptyset, \{(1, 2, 3)\}, \Omega \setminus \{(1, 2, 3)\}, \Omega\}$ on σ-algebra joukossa Ω. Käydään läpi σ-algebran ominaisuudet

1. $\Omega \in \Gamma$ pätee

2. Jos $A = \emptyset$, niin $A^c = \Omega \in \Gamma$. Jos taas $A = \{(1, 2, 3)\}$, tällöin $A^c = \Omega \setminus \{(1, 2, 3)\} \in \Gamma$. Jos $A = \Omega$, niin $A^c = \emptyset \in \Gamma$ eli toinen ominaisuus pätee.

3. Minkä tahansa Γ:n joukkojen yhdistäminen johtaa tuloksiin $\emptyset, \{(1, 2, 3)\}, \Omega \setminus \{(1, 2, 3)\}$ tai Ω, ja kaikki nämä kuuluvat joukkoon Γ eli myös kolmas ominaisuus pätee.

Näin ollen Γ on σ-algebra joukossa Ω. $G = \{(1, 2, 3), (1, 1, 1)\}$ ei taas ole σ-algebra joukossa Ω, sillä jos $A = \{(1, 2, 3)\}$, niin $A^c = \Omega \setminus \{(1, 2, 3)\} \notin G$.

Lause 2.8. *Jos $\Omega \neq \emptyset$ ja $\Gamma_1, \Gamma_2, \dots$ ovat σ-algebroja joukossa Ω, niin $\bigcap_{i=1}^{\infty} \Gamma_i$ on myös σ-algebra samassa joukossa.*

Todistus. Koska $\Omega \in \Gamma_j$ kaikilla j, niin $\Omega \in \bigcap_j \Gamma_j$. Olkoon $A \in \Gamma_j$, kaikilla j, jolloin $A^c \in \Gamma_j$. Siksi myös $A^c \in \bigcap_j \Gamma_j$. Olkoon $A_1, A_2, \dots \in \bigcap_j \Gamma_j$. Tällöin

$A_i \in \Gamma_j$, kaikilla i, j, eli myös $\bigcup\limits_i A_i \in \Gamma_j$ kaikilla j. Näin ollen voidaan päätellä, että $\bigcup\limits_{i=1}^{\infty} A_i \in \bigcap\limits_j \Gamma_j$. □

Huomaa, että σ-algebrojen yhdiste ei yleisesti ole σ-algebra joukossa Ω. Esimerkiksi jos $\Omega = \{1, 2, 3\}$, $\Gamma_1 = \{\emptyset, \{1\}, \{2, 3\}\}$ ja $\Gamma_2 = \{\emptyset, \Omega, \{1, 2\}, \{3\}\}$, niin Γ_1 ja Γ_2 ovat σ-algebroja joukossa Ω, mutta $\Gamma_1 \cup \Gamma_2$ ei ole.

Määritelmä 2.9. Olkoon $\Omega \neq \emptyset$ ja olkoon A joukon Ω osajoukko. Olkoon $M = \{\Gamma : \Gamma \text{ on } \sigma \text{ -algebra joukossa } \Omega \text{ sisältäen joukon } A\}$. Tällöin $\sigma(A) := \bigcap\limits_{\Gamma \in M} \Gamma$ on pienin mahdollinen joukon A sisältämä σ-algebra joukossa Ω. $\sigma(A)$ kutsutaan joukon A generoimaksi σ-algebraksi.

Esimerkki 2.10. Pienintä σ-algebraa joukossa $\mathbb{R}$, joka sisältää kaikki välit $]-\infty, a]$, missä $a \in \mathbb{R}$ kutsutaan Borelin σ-algebraksi. Jos $a, b \in \mathbb{R}$ siten, että $a < b$, niin seuraavat ovat $\mathbb{R}$:n Borel-joukkoja.

$$(a, \infty) = \mathbb{R} \setminus (-\infty, a]$$

$$(a, b\,] = (-\infty, b\,] \cap (a, \infty)$$

$$(-\infty, a) = \bigcup_{n=1}^{\infty} \left(-\infty, a - \frac{1}{n}\right]$$

$$[a, \infty) = \mathbb{R} \setminus (-\infty, a)$$

$$(a, b) = (-\infty, b) \cap (a, \infty)$$

$$[a, b] = \mathbb{R} \setminus ((-\infty, a) \cup (b, \infty))$$

$$\{a\} = [a, a]$$

$$\mathbb{N} = \bigcup_{n=0}^{\infty} \{n\}$$

$$\mathbb{Q} = \bigcup_{\substack{m,n \in \mathbb{Z} \\ n \neq 0}} \left\{\frac{m}{n}\right\}$$

$$\mathbb{Q}^c = \mathbb{R} \setminus \mathbb{Q}$$

Huomaa, että kaikki $\mathbb{R}$:n joukot eivät ole Borel-joukkoja.

Esimerkki 2.11. Olkoot $a = (a_1, \cdots, a_n)$ ja $b = (b_1, \cdots, b_n)$ joukon $\mathbb{R}^n$ alkioita, missä $a_i \leq b_i$ kaikilla $i = 1, \cdots, n$. Borelin-joukkojen σ-algebran generoivat kaikki välit, jotka ovat muotoa

$$(a, b] := \{x = (x_1, \cdots, x_n) \in \mathbb{R}^n : a_i < x_i \leq b_i, i = 1, \cdots, n\}.$$

Jo peruskoulussa opetetaan varman ja mahdottoman tapahtuman käsitteet. Todennäköisyysavaruudessa tyhjää joukkoa $\emptyset$ kutsutaan mahdottomaksi tapahtumaksi ja joukkoa Ω varmaksi tapahtumaksi. Lukiossa opitaan, että todennäköisyys on aina epänegatiivinen ja varman tapahtuman todennäköisyys on 1. Nämä kaksi aksioomaa ovat todennäköisyysmitan ehtoja. Lisäksi opitaan erillisten eli toisensa poissulkevien tapahtumien käsite, joka seuraavassa määritelmässä tulee esiin. Tätä määritelmää ei kuitenkaan opeteta lukiossa, vaikka tarvittavat tiedot siihen olisi.

Määritelmä 2.12. Olkoot $\Omega \neq \emptyset$ ja Γ σ-algebra joukossa Ω. Reaaliarvoista funktiota P kutsutaan todennäköisyysmitaksi, jos se täyttää seuraavat ehdot:

(i) $P(A) \geq 0$ kaikilla $A \in \Gamma$

(ii) $P(\Omega) = 1$

(iii) Jos $A_1, A_2, \dots$ ovat toisensa poissulkevat tapahtumat joukossa Γ, eli

$$A_i \bigcap A_j = \emptyset \quad \text{kaikilla} \quad i \neq j,$$

niin silloin

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Kolmikkoa (Ω, Γ, P) kutsutaan todennäköisyysavaruudeksi.

Esimerkki 2.13. Olkoot $\Omega = \{1, 2, 3\}$, $\Gamma = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$ ja P

$$P(A) = \begin{cases} 1, & \text{jos} \quad 3 \in A \\ 0, & \text{jos} \quad 3 \notin A. \end{cases}$$

Tällöin P on todennäköisyysmitta joukossa (Ω, Γ).

Seuraavaksi esitellään todennäköisyysmitan tärkeimpiä ominaisuuksia. Tämän Lauseen kohtia käsitellään myös lukio-opinnoissa. Lukiossa ei käytetä yhdisteen ja leikkauksen matemaattisia merkintöjä, vaan ne esitetään "ja" ja "tai" -sanoilla. Kuten huomataan, niin Lauseen toinen kohta kuvaa yhteenlaskusäännön erillisille tapahtumille. Kolmas kohta on lukiossa esiintyvän vastatapahtuman todennäköisyys. Viidennessä kohdassa käsitellään yleinen yhteenlaskusääntö. Matematiikan opettajien täytyy ymmärtää nämä ominaisuudet pystyäkseen opettamaan lukion matematiikkaa laadukkaasti.

Lause 2.14. *Olkoon* (Ω, Γ, P) *todennäköisyysavaruus. Tällöin*

1. $P(\emptyset) = 0$.

2. *Jos* $A, B \in \Gamma$ *ja* $A \cap B = \emptyset$, *niin* $P(A \cup B) = P(A) + P(B)$.

3. *Kaikille* $A \in \Gamma$ *pätee* $P(A^c) = 1 - P(A)$.

4. *Jos* $A \subseteq B$, *niin* $P(A) \leq P(B)$ *ja* $P(B \setminus A) = P(B) - P(A)$. *Erityisesti* $P(A) \leq 1$ *kaikilla* $A \in \Gamma$.

5. *Kaikille* $A, B \in \Gamma$ *pätee* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

6. *Olkoon* $(A_n)_n$ *kasvava jono joukon* Γ *alkioita, eli* $A_n \in \Gamma$ *ja* $A_n \subseteq A_{n+1}$ *kaikilla* $n = 1, 2, \dots$, *niin tällöin*

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

   missä $\lim_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} A_n$.

7. *Olkoon* $(A_n)_n$ *laskeva jono joukon* Γ *alkioita, eli* $A_n \in \Gamma$ *ja* $A_n \supseteq A_{n+1}$ *kaikilla* $n = 1, 2, \dots$, *niin tällöin*

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n)$$

   missä $\lim_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} A_n$.

Todistus. 1. Määritelmän 2.12 nojalla $1 = P(\Omega)$. Voidaan merkitä $\Omega = \Omega \cup \emptyset \cup \emptyset \cup \cdots$. Tästä saadaan, että

$$\begin{aligned}
1 &= P(\Omega) \\
&= P(\Omega \cup \emptyset \cup \emptyset \cdots) \\
&= P(\Omega) + P(\emptyset) + P(\emptyset) + \cdots \\
&= 1 + P(\emptyset) + P(\emptyset) + \cdots
\end{aligned}$$

Eli $0 \geq P(\emptyset) \geq 0$, joten täytyy olla $P(\emptyset) = 0$.

2. Määritelmän 2.12 kohdan 3. ja todistuksen 1. kohdan tuloksesta seuraa, että koska $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \cdots$, niin

$$\begin{aligned}
P(A \cup B) &= P(A \cup B \cup \emptyset \cup \emptyset \cup \cdots) \\
&= P(A \cup B) + P(\emptyset) + P(\emptyset) + \cdots \\
&= P(A) + P(B) + 0 + 0 + \cdots \\
&= P(A) + P(B).
\end{aligned}$$

3. Koska $A \cup A^c = \Omega$ ja $A \cap A^c = \emptyset$, niin määritelmän 2.12 nojalla

$$\begin{aligned} 1 &= P(\Omega) \\ &= P(A \cup A^c) \\ &= P(A) + P(A^c) \end{aligned}$$

Eli $P(A^c) = 1 - P(A)$.

4. Tiedetään $B = A \cup (B \setminus A)$. Jolloin saadaan $P(B) = P(A) + P(B \setminus A) \geq P(A)$. Eli $P(B) \geq P(A)$.
Edelleen $B = A \cup (B \setminus A)$. Koska $A \cap (B \setminus A) = \emptyset$, niin

$$P(B) = P(A) + P(B \setminus A)$$

eli

$$P(B \setminus A) = P(B) - P(A).$$

Määritelmän 2.12 ja tuloksen $P(B) \geq P(A)$ nojalla saadaan, että

$$P(A) \leq P(\Omega) = 1.$$

Eli

$$P(A) \leq 1.$$

5. Tiedetään, että $A \cup B = (A \cup B) \cap (A \cup A^c) = A \cup (B \cap A^c)$ eli

$$P(A \cup B) = P(A) + P(B \cap A^c).$$

Lisäksi $B = (B \cap A) \cup (B \cap A^c)$, joten myös

$$P(B) = P(B \cap A) + P(B \cap A^c)$$

eli

$$P(B \cap A^c) = P(B) - P(B \cap A).$$

Joten nyt yhdistämällä nämä kaksi saadaan, että

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

6. Olkoon $C_1 = A_1, C_2 = A_2 \setminus A_1, \cdots, C_n = A_n \setminus \bigcup_{i=1}^{n-1} A_i$. On selvää, että:

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n.$$

Lisäksi koska $C_i \cap C_j = \emptyset$ kaikilla $i \neq j$, niin

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P\left(\bigcup_{n=1}^{\infty} C_n\right) \\ &= \sum_{n=1}^{\infty} P(C_n) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} P(C_k) \\ &= \lim_{n\to\infty} P\left(\bigcup_{k=1}^{n} C_k\right) \\ &= \lim_{n\to\infty} P(A_n). \end{aligned}$$

7. Jätetään tämä todistus harjoitustehtäväksi.

$\square$

Huomautus 2.15. Olkoot A, B, C tapahtumia. Edellistä lausetta soveltamalla saadaan, että

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) \qquad (2.1) \\ &\quad - P(A \cap C) + P(A \cap B \cap C). \end{aligned}$$

Jos $A_1, A_2, \ldots, A_n$ ovat tapahtumia, niin

$$\begin{aligned} P(A_1 \cup A_2 \cup \cdots \cup A_n) &= \sum_{i=1}^{n} P(A_i) - \sum_{i_1<i_2} P(A_{i_1} \cap A_{i_2}) + \cdots \\ &\quad + (-1)^{r+1} \sum_{i_1<i_2<\cdots<i_r} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r}) \\ &\quad + \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n), \end{aligned}$$

missä summaan

$$\sum_{i_1<i_2<\cdots<i_r} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r})$$

on otettu kaikki joukon $\{1, 2, \ldots, n\}$ mahdolliset osajoukot.

Esimerkki 2.16. Olkoon (Ω, Γ, P) todennäköisyysavaruus, jossa:

$$\Omega = \{1, 2, 3, 4\}$$

$$\Gamma = \{\emptyset, \Omega, \{1\}, \{2, 3\}, \{4\}, \{1, 2, 3\}, \{2, 3, 4\}, \{1, 4\}\}$$

$$P(\{1\}) = \frac{1}{4},\ P(\{2,3\}) = \frac{1}{2},\ P(\{4\}) = \frac{1}{4}.$$

Tällöin:

$$\begin{aligned} P(\{1,2,3\}) &= \frac{1}{4} + \frac{1}{2} = \frac{3}{4} \\ P(\{2,3,4\}) &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4} \\ P(\{1,4\}) &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \end{aligned}$$

2.2 Klassinen todennäköisyys

Seuraavaksi käydään läpi tarkempi määritelmä klassiselle todennäköisyydelle, ja käydään sen hyödyntämisestä esimerkki. Jos alkeistapauksia on äärellinen määrä ja ne ovat yhtä todennäköisiä, puhutaan klassisesta todennäköisyydestä. Lukiossa klassinen todennäköisyys opetetaan alla olevan huomautuksen kaltaisesti.

Määritelmä 2.17. Olkoot Ω äärellinen, Γ kaikkien Omegan osajoukkojen muodostama σ-algebra ja $P(\omega) = \frac{1}{|\Omega|}$ alkeistapahtumien todennäköisyys, joka on luku väliltä $[0,1]$, kaikilla $\omega \in \Omega$. Todennäköisyysavaruutta (Ω, Γ, P), kutsutaan Laplacen todennäköisyysavaruudeksi. Todennäköisyyttä P kutsutaan klassiseksi todennäköisyydeksi.

Huomautus 2.18. Jos (Ω, Γ, P) on Laplace todennäköisyysavaruus, $A \subseteq \Omega$ mikä tahansa Ω:n osajoukko, niin

$$P(A) = P\left(\bigcup_{\omega \in A} \omega\right) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|},$$

missä $|A|$ on joukon A alkioiden lukumäärä. Toisin sanoen

$$P(A) = \frac{\text{Joukon } A \text{ suotuisat alkeistapaukset}}{\text{alkeistapausten lukumäärä}}.$$

Esimerkki 2.19. Laatikossa on 4 Tobleronea, 8 Fazerin sinistä suklaalevyä ja 5 Tupla-suklaapatukkaa. Laatikosta nostetaan satunnaisesti kuusi suklaata ilman, että niitä laitetaan takaisin. Mikä on todennäköisyys P sille, että nostetaan 2 Tobleronea, 3 Fazerin sinistä ja 1 Tupla?
Ratkaisu: Laatikosta voidaan nostaa suklaita $\binom{17}{6}$ eri tavalla. Alkeistapauksia ovat kaikki kuuden suklaan osajoukot ja näiden lukumäärä on $|\Omega| = \binom{17}{6}$.

Kaksi Tobleronea voidaan nostaa $\binom{4}{2}$ eri tavalla, Fazerin sinisiä voidaan nostaa $\binom{8}{3}$ eri tavalla ja Tuplia $\binom{5}{1}$ eri tavalla. Tuloperiaatteen mukaan suklaita voidaan nostaa $\binom{4}{2}\binom{8}{3}\binom{5}{1} = 6 \cdot 56 \cdot 5 = 1680$ eri tavalla. Suotuisten alkeistapausten lukumäärä on siis $|A| = 1680$. Eli nyt

$$P(A) = \frac{|A|}{|\Omega|} = \frac{1680}{12376} = \frac{30}{221}.$$

2.3 Ehdollinen todennäköisyys ja riippumattomuus

Tässä alaluvussa käsitellään jo lukioaiheista tuttuja ehdollista todennäköisyyttä sekä riippumattomuutta. Opitaan näiden täsmälliset määritelmät ja käydään läpi esimerkkejä aiheisiin liittyen. Kun verrataan seuraavaa määritelmää lukiossa käytävään määritelmään, huomataan sen olevan hyvin samanlainen.

Määritelmä 2.20. Olkoon (Ω, Γ, P) todennäköisyysavaruus. Jos $A, B \in \Gamma$ siten, että $P(A) > 0$, niin tapahtuman B todennäköisyys ehdolla A on määritelty seuraavanlaisesti:

$$P(B \mid A) := \frac{P(A \cap B)}{P(A)}.$$

Esimerkki 2.21. Kahta noppaa heitetään kerran. Olkoot A tapahtuma ”silmäluvut ovat erisuuret” ja B = ”vähintään yksi silmäluvuista on 6”. Nyt

$$A = \{(a, b) : a, b \in \{1, 2, 3, \ldots, 6\}, a \neq b\} \quad \text{ja}$$

$$B = \{(a, 6) : a \in \{1, 2, \ldots, 6\}\} \cup \{(6, b) : b \in \{1, 2, , \ldots, 6\}\}.$$

Lasketaan ensin $P(A)$. Ensimmäisen nopan silmäluku voi olla mikä tahansa kuudesta. Jotta silmäluvut ovat erisuuret, toisen nopan silmäluvulle on vaihtoehtoja 5. Suotuisien alkeistapausten määrä on siis $6 \cdot 5 = 30$. Kahta noppaa heittämällä alkeistapauksia on yhteensä $6 \cdot 6 = 36$. Eli

$$P(A) = \frac{30}{36} = \frac{5}{6}$$

Nyt $P(A \cap B)$ on todennäköisyys, että silmäluvut ovat erisuuret ja toinen silmäluvuista on 6. Jos ensimmäisen nopan silmäluku on kuusi, niin toisen nopan silmäluvulle on vaihtoehtoja 5. Jos toisen nopan silmäluku on kuusi, niin myös ensimmäisen nopan silmäluvulle on vaihtoehtoja 5. Yhteensä suotuisia alkeistapauksia on siis 10. Eli

$$P(A \cap B) = \frac{10}{36}.$$

Voidaan seuraavaksi laskea ehdollinen todennäköisyys.

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{10}{36}}{\frac{5}{6}} = \frac{1}{3}.$$

Seuraavassa Lauseessa käsitellään ehdollisen todennäköisyyden keskeisimpiä ominaisuuksia, joita ei käsitellä lukiossa. Matematiikan opettajien on hyödyllistä osata nämä ymmärtääkseen, miten ehdollinen todennäköisyys käyttäytyy eri tilanteissa.

Lause 2.22. *Olkoon* (Ω, Γ, P) *ja olkoon* $A \in \Gamma$ *siten, että* $P(A) > 0$*. Tällöin:*

1. $P(\cdot \mid A)$ *on todennäköisyysmitta joukossa* Ω *ja* $P(A \mid A) = 1$.
2. *Jos* $A \cap B = \emptyset$*, niin* $P(B \mid A) = 0$.
3. $P(B \cap C \mid A) = P(B \mid A \cap C)P(C \mid A)$*, jos* $P(A \cap C) > 0$.
4. *Jos* $A_1, A_2, \ldots, A_n \in \Gamma$ *siten, että* $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$*, niin tällöin*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots \\ P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Todistus. 1. Tarkistetaan todennäköisyysmitan kolme ominaisuutta:

i) Selvästi nähdään, että $P(B \mid A) \geq 0$, kaikilla $B \in \Gamma$.

ii) $P(\Omega \mid A) = \frac{P(A \cap \Omega)}{P(A)} = \frac{P(A)}{P(A)} = 1$. Siksi myös $P(A \mid A) = 1$.

iii) Olkoot $A_1, A_2, \ldots$ erillisiä tapahtumia. Tällöin:

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} A_i \mid A\right) &= \frac{P\left(A \cap \left(\bigcup_{i=1}^{\infty} A_i\right)\right)}{P(A)} \\ &= \frac{P\left(\bigcup_{i=1}^{\infty}(A_i \cap A)\right)}{P(A)} \\ &= \frac{\sum_{i=1}^{\infty} P(A_i \cap A)}{P(A)} \\ &= \sum_{i=1}^{\infty} P(A_i \mid A). \end{aligned}$$

2. Jos $A \cap B = \emptyset$, niin tällöin $P(A \cap B) = 0$ eli $P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0}{P(A)} = 0$.

3.

$$\begin{aligned}
P(B \cap C \mid A) &= \frac{P(B \cap C \cap A)}{P(A)} \\
&= \frac{P(B \cap C \cap A)}{P(C \cap A)} \times \frac{P(C \cap A)}{P(A)} \\
&= P(B \mid A \cap C) P(C \mid A).
\end{aligned}$$

4. Todistetaan väite induktiolla. Osoitetaan aluksi, että väite pätee, kun $n = 2$. Nyt siis

   $$P(A_1)P(A_2 \mid A_1) = P(A_1)\frac{P(A_2 \cap A_1)}{P(A_1)} = P(A_1 \cap A_2).$$

   Eli väite pätee, kun $n = 2$.

   Oletetaan nyt, että väite pätee kaikilla $n - 1$ tapahtumalla. Koska $P(A_1 \cap A_2 \cap \cdots \cap A_{n-2}) \geq P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, voidaan päätellä, että

   $$\begin{aligned}
   P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots \\
   P(A_{n-1} \mid A_1 \cap A_2 \cap \cdots \cap A_{n-2}).
   \end{aligned}$$

   Olkoon $B = A_1 \cap A_2 \cap \cdots \cap A_{n-1}$. Tällöin

   $$\begin{aligned}
   P(A_1 \cap A_2 \cap \cdots \cap A_n) &= P(B \cap A_n) \\
   &= P(B)P(A_n \mid B) \\
   &= P(A_1 \cap A_2 \cap \cdots \cap A_{n-1})P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \\
   &= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots \\
   &\quad P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1}).
   \end{aligned}$$

   Nyt voidaan päätellä, että väite on totta kaikilla n:n arvoilla.

$\square$

Esimerkki 2.23. Laatikossa on 12 palloa, joista 4 on mustia ja loput kahdeksan ovat valkoisia. Pelataan seuraavanlaista peliä: Nostetaan laatikosta satunnaisesti pallo ja katsotaan sen väri. Tämän jälkeen pallo palautetaan takaisin laatikkoon ja lisäksi lisätään kaksi nostetun väristä palloa laatikkoon. Laske todennäköisyys sille, että pelin kolmella ensimmäisellä kierroksella nostetut pallot ovat kaikki mustia.

Ratkaisu: Kaikille $i = 1, 2, 3$ pätee

$$A_i := \text{Musta pallo nostettiin } i\text{:nellä kierroksella.}$$

Sovelletaan Lauseen 2.22 kohtaa neljä, kun $n = 3$ ja saadaan, että

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_3 \mid A_2 \cap A_1)P(A_2 \mid A_1)P(A_1) \\ &= \frac{8}{16} \times \frac{6}{14} \times \frac{4}{12} \\ &= \frac{1}{14}. \end{aligned}$$

Esimerkki 2.24. Kolme lasta yrittää päästä aikuisille suunnattuun konserttiin. Jokaiselta lapselta kysytään henkilöllisyystodistukset porteilla ja virkailijan hylätessä lasten pääsyn konserttiin, palauttaa hän henkilöllisyystodistukset satunnaisesti. Laske todennäköisyys sille, että kukaan ei saa takaisin omaa henkilöllisyystodistusta.

Ratkaisu: Olkoot:

$$A := \text{kukaan ei saa omaa henkilötodistustaan takaisin}$$

$$B_i := i\text{:nnes lapsi saa oman henkilötodistuksena takaisin.}$$

Todennäköisyys, jota etsitään, saadaan hyödyntämällä huomautusta 2.15:

$$\begin{aligned} P(A) &= P(B_1^c \cap B_2^c \cap B_3^c) \\ &= 1 - P(B_1 \cup B_2 \cup B_3) \\ &= 1 - P(B_1) - P(B_2) - P(B_3) + P(B_1 \cap B_2) \\ &+ P(B_2 \cap B_3) + P(B_1 \cap B_3) - P(B_1 \cap B_2 \cap B_3). \end{aligned}$$

Koska on olemassa kolme erilaista vaihtoehtoa, joista vain yksi on oikea, niin

$$P(B_i) = \frac{1}{3} \quad \text{kaikilla} \quad i = 1, 2, 3.$$

Toisaalta myös kaikilla $i \neq j$ pätee

$$P(B_i \cap B_j) = P(B_i)P(B_j \mid B_i) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}.$$

Annettuaan oikean henkilötodistuksen takaisin i:nnelle lapselle, j:nnen lapsen henkilötodistukselle jää yksi oikea vaihtoehto kahdesta eli:

$$P(B_1 \cap B_2 \cap B_3) = P(B_1)P(B_2 \mid B_1)P(B_3 \mid B_1 \cap B_2) = \frac{1}{6}.$$

Siksi

$$P(A) = \frac{1}{3}.$$

Lause 2.25. *Oletetaan, että* $A_1, A_2, \ldots$ *muodostavat otosavaruuden* Ω *ositexcept eli* $A_i \cap A_j = \emptyset$, *kaikilla* $i \neq j$ *ja* $\bigcup_{i=1}^{\infty} A_i = \Omega$ *siten, että* $P(A_i) > 0$ *kaikilla* $A_i \in \Gamma$. *Tällöin kaikille* $B \in \Gamma$ *pätee:*

$$P(B) = \sum_i P(B \mid A_i) P(A_i).$$

Todistus. Nyt

$$\begin{aligned} B &= B \cap \Omega \\ &= B \cap \bigcup_{i=1}^{\infty} A_i \\ &= \bigcup_{i=1}^{\infty} (B \cap A_i) \end{aligned}$$

Koska $B \cap A_i$ ovat erilliset, niin todennäköisyysmitan määritelmän kolmatta kohtaa hyödyntämällä saadaan, että

$$\begin{aligned} P(B) &= \sum_i P(B \cap A_i) \\ &= \sum_i P(B \mid A_i) P(A_i). \end{aligned}$$

□

Esimerkki 2.26. Markku tietää, että on 40% todennäköisyys sille, että yritys, jossa hän työskentelee, avaa uuden osaston Helsinkiin. Jos uusi osasto avataan, todennäköisyys sille, että Markusta tulee uuden toimiston johtaja, on 80%. Jos uutta toimistoa ei avata todennäköisyys sille, että Markusta tulee nykyisen osaston johtaja, on 10%. Laske todennäköisyys, että Markusta tulee johtaja hänen työskentelemässään yrityksessä.
Ratkaisu: Olkoon

$$J := \text{Markusta tulee johtaja.}$$

$$U := \text{Yritys avaa uuden osaston Helsinkiin.}$$

Tällöin

$$\begin{aligned} P(J) &= P(J \mid N) P(U) + P(J \mid U^c) P(U^c) \\ &= 0,8 \times 0,4 + 0,10 \times 0,60 \\ &= 0,38. \end{aligned}$$

Edellisen Lauseen seurauksena saadaan Bayesin sääntönä tunnettu tulos.

Seuraus 2.27. *Olkoot $A_1, A_2, \ldots$ otosavaruuden Ω ositus, missä $P(A_i) > 0$ kaikilla i. Tällöin kaikilla $B \in \Gamma$, jolle $P(B) > 0$ pätee:*

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_j P(B \mid A_j)P(A_j)} \quad \textit{kaikilla } i.$$

Todistus.

$$\begin{aligned} P(A_i \mid B) &= \frac{P(A_i) \cap B)}{P(B)} \\ &= \frac{P(A_i)P(B \mid A_i)}{P(P)} \\ &= \frac{P(A_i)P(B \mid A_i)}{\sum_j P(B \mid A_j)P(A_j)}. \end{aligned}$$

□

Esimerkki 2.28. Jos esimerkin 2.26 tilanteessa Markku nimitetään osaston johtajaksi, niin mikä on todennäköisyys, että yritys on avannut uuden osaston Helsinkiin?

Ratkaisu:

Tiedetään, että $P(U) = 0,4$ ja $P(J \mid U) = 0,8$. Lisäksi esimerkissä 2.26 laskettiin, että Markusta tulee johtaja $P(J) = 0,38$ todennäköisyydellä. Seurauksen 2.27 nojalla saadaan, että

$$\begin{aligned} P(U \mid J) &= \frac{P(U)P(J \mid U)}{P(J)} \\ &= \frac{0,4 \cdot 0,8}{0,38} \\ &= 0,84211. \end{aligned}$$

Eli todennäköisyys, että yritys on avannut uuden osaston Helsinkiin, on noin $0,84$.

Joskus tapahtuma B ei vaikuta tapahtuman A todennäköisyyteen, jolloin $P(A \mid B) = P(A)$. Tiedetään, että tässä tapauksessa tapahtuma A on riippumaton tapahtumasta B, mikä vaatii sen, että $P(B) > 0$. Tämän vaatimuksen välttämiseksi, riippumattomuus määritellään seuraavalla tavalla:

Määritelmä 2.29. Tapahtumat A ja B ovat riippumattomia, jos ja vain jos:

$$P(A \cap B) = P(A)P(B).$$

Jos edellinen ehto ei täyty tapahtumat riippuvat toisistaan.

Esimerkki 2.30. Tavallista noppaa heitetään kaksi kertaa. Olkoot:

$$\begin{aligned} A &:= \text{Noppien silmälukujen summa on parillinen.} \\ B &:= \text{Toisen heiton silmäluku on parillinen} \end{aligned}$$

Tällöin:

$$P(A) = P(B) = \frac{1}{2}$$

Lisäksi $P(A \cap B) = \frac{1}{4}$. Eli tapahtumat ovat riippumattomat.

Lause 2.31. *Olkoot tapahtumat A ja B riippumattomat. Tällöin*

1. *Tapahtumat A ja B^c ovat riippumattomat (myös B ja A^c ovat)*

2. *Tapahtumat A^c ja B^c ovat riippumattomat*

Todistus. 1.

$$\begin{aligned} P(A) &= P(A \cap B^c) + P(A \cap B) \\ &= P(A \cap B^c) + P(A)P(B), \end{aligned}$$

Siksi:

$$P(A \cap B^c) = P(A)\,[1 - P(B)] = P(A)P(B^c).$$

2.

$$\begin{aligned} P(A^c \cap B^c) &= 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= [1 - P(A)]\,[1 - P(B)] \\ &= P(A^c)P(B^c). \end{aligned}$$

□

Yllä oleva riippumattomuuden määritelmä on identtinen Moodi 8 -kirjassa (2023) esitettävän määritelmän kanssa. Usein on tarpeen tarkastella kahden tai useamman tapahtuman riippumattomuuksia. Tämän vuoksi riippumattomuudelle on annettava laajempi määritelmä kuin lukiossa.

Määritelmä 2.32. Joukkoperhettä $\{A_i : i \in I\}$ sanotaan riippumattomaksi, jos

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$$

kaikilla I:n äärellisillä osajoukoilla $\emptyset \neq J$. Nämä tapahtumat ovat toisistaan riippumattomia.

Määritelmä 2.33. Joukkoperhettä $\{A_i : i \in I\}$ sanotaan pareittain 2×2 riippumattomaksi jos:

$$P(A_i \cap A_j) = P(A_i)P(A_j) \text{ kaikilla } i \neq j.$$

Esimerkki 2.34. Noppaa heitetään kaksi kertaa peräkkäin. Olkoot A, B, C seuraavia tapahtumia:

$$\begin{aligned}
A &:= \text{"Silmäluku 2 saadaan ensimmäisellä heitolla".} \\
B &:= \text{"Silmäluku 5 saadaan toisella heitolla".} \\
C &:= \text{"Noppien silmälukujen summa on 7".}
\end{aligned}$$

Nyt

$$\begin{aligned}
&P(A) = P(B) = P(C) = \frac{1}{6} \\
&P(A \cap B) = P(A \cap C) = P(B \cap C) = \frac{1}{36} \\
&P(A \cap B \cap C) = \frac{1}{36}
\end{aligned}$$

Tapahtumat ovat pareittain riippumattomat, mutta eivät ole toisistaan riippumattomia sillä:

$$P(A \cap B \cap C) \neq P(A)P(B)P(C).$$

2.4 Geometrinen todennäköisyys

Tässä alaluvussa käsitellään geometristä todennäköisyyttä ja käydään läpi esimerkki aiheeseen liittyen. Huomataan, että seuraavan määritelmän merkinnät näyttävät hyvin samalta kuin Moodi 8 -oppikirjassa (2023).

Määritelmä 2.35. Olkoot $\Omega \neq \emptyset$ ja Γ σ-algebra joukossa Ω ja oletetaan, että geometrinen mitta m on määritelty joukossa (Ω, Γ). Tällöin tapahtuman A geometrinen todennäköisyys on

$$P(A) := \frac{m(A)}{m(\Omega)}.$$

Geometrista mittaa kutsutaan Lebesguen mitaksi, jota käytetään yleisesti reaalilukujen, taso- ja avaruusjoukkojen mittaamiseen. Lebesguen mitta perustuu σ-algebroihin. Sigma-algebraan kuuluvat vain ne joukot, joiden mitta voidaan määrittää Lebesguen mitan avulla. Kaikki joukot eivät kuitenkaan ole mitallisia, eikä Lebesguen mittaa voida määrittää ei-mitallisille joukoille.

Esimerkki 2.36. Liisa ja Pekka sopivat tapaamisen kahvilaan kello 12 ja 13 välille. He voivat saapua paikalle minä ajan hetkenä hyvänsä. Olettaen, että heidän saapumisaikansa ovat riippumattomia, laske todennäköisyys, että Liisa ja Pekka tapaavat siten, että kumpikin odottaa enintään 10 minuuttia. Ratkaisu:

Olkoot X ja Y tapahtumia:

$$X := \text{"Liisan saapumisaika"}$$

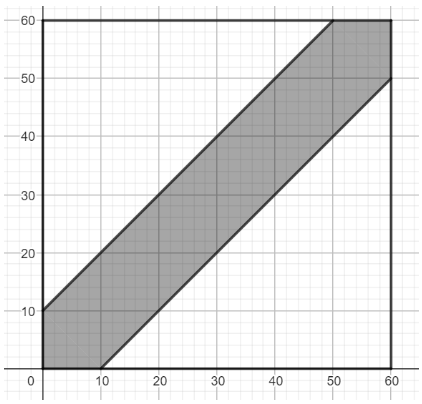
$$Y := \text{"Pekan saapumisaika"}$$

Otosavaruus

$$\Omega := \{(x, y) \in \mathbf{R}^2 : 0 \leq x \leq 60, 0 \leq y \leq 60\}.$$

Halutaan laskea tapahtuman T todennäköisyyttä, missä

$$T = \{(x, y) \in \Omega :| x - y |\leq 10\}.$$



Kuva 2.1: Esimerkki 2.36

Eli nyt

$$P(T) = \frac{T\text{:n pinta-ala}}{\Omega\text{:n pinta-ala}} = \frac{(3600 - 50 \cdot 50)}{3600} = \frac{1100}{3600} = \frac{11}{36}.$$

2.5 Satunnaismuuttuja

Tässä alaluvussa käsitellään satunnaismuuttujaan liittyviä määritelmiä ja esimerkkejä. Aihetta käsiteltiin hieman jo lukiomatematiikassa ja opittiin, mikä on satunnaismuuttuja. Lukiossa satunnaismuuttujaa ei käsitellä funktiona, vaan se selitetään vain sanallisesti. Juuri 8 -oppikirjassa (2022) kyllä mainitaan, että satunnaismuuttuja on täsmällisemmin funktio, mutta siihen ei perehdytä tarkemmin. Seuraavaksi käydään läpi satunnaismuuttujalle tarkempi määritelmä.

Määritelmä 2.37. Olkoon (Ω, Γ, P) todennäköisyysavaruus. Satunnaismuuttuja on kuvaus $X : \Omega \to \mathbb{R}$ siten, että $A \in \mathcal{B}, X^{-1}(A) \in \Gamma$, missä $\mathcal{B}$ on Borelin σ-algebra joukossa $\mathbb{R}$.

Huomautus 2.38. Olkoon (Ω, Γ, P) jokin mielivaltainen todennäköisyysavaruus. Ottaen huomioon, että joukot $(-\infty, x]$, missä $x \in \mathbb{R}$, generoivat Borelin σ-algebran, voidaan osoittaa, että funktio $X : \Omega \to \mathbb{R}$ on satunnaismuuttuja, jos ja vain jos $X^{-1}((-\infty, x]) \in \Gamma$ kaikilla $x \in \mathbb{R}$.

Esimerkki 2.39. Heitetään kahta tavallista noppaa. Tällöin $\Omega = \{(i, j), i, j = 1, 2, \cdots, 6\}$ on kaikki noppien heitoista saatavat tulokset. Olkoon Γ kaikkien osajoukkojen kokoelma $p(\Omega)$. Määritellään X seuraavasti

$$X(\omega) = i + j \quad \text{jos} \quad \omega = (i, j) \in \Omega$$

ja koska nyt $X^{-1}((-\infty, x]) \in \Gamma$, niin X on satunnaismuuttuja.

Esimerkki 2.40. Hehkulamppu valmistetaan. Valmistuksen jälkeen sen käyttöikä X testataan ja kirjataan ylös aika, joka on kulunut silloin, kun lamppu lakkaa toimimasta. Tässä tapauksessa $\Omega = \{t : t \geq 0\}$. Olkoon $\Gamma = \mathcal{B}(\mathbb{R} \cap \Omega)$ σ-algebra. Määritellään

$$X : \Omega \to \mathbb{R} \quad X(\omega) = \omega, \, \omega \in \Omega.$$

Tällöin

$$X^{-1}(-\infty, x] = \begin{cases} \emptyset, \quad \text{jos} \quad x < 0 \\ (0, x], \quad \text{jos} \quad x \geq 0. \end{cases}$$

Siksi käyttöikä X on satunnaismuuttuja.

Huomautus 2.41. Olkoon X satunnaismuuttuja todennäköisyysavaruudessa (Ω, Γ, P). Satunnaismuuttujaan liittyviä tapahtumia voidaan esittää myös joukkomerkintämuodossa:

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}, \text{ missä } B \in \mathcal{B}.$$

Lukiossa opetetaan, että satunnaismuuttujan jakauma sisältää satunnaismuuttujan arvot ja sen todennäköisyydet. Lukiossa myös satunnaismuuttujan jakauma määritellään sanallisesti, eikä käytetä funktiota P_X kuvaamaan todennäköisyysmittaa. Lukiossa ei ylipäätään opita todennäköisyysmitan käsitettä. Seuraavaksi esitetään satunnaismuuttujan jakaumalle täsmällisempi esitys.

Lause 2.42. *Oletetaan, että X on satunnaismuuttuja todennäköisyysavaruudessa (Ω, Γ, P). Funktio P_X*

$$P_X(B) := P\left(\{X \in B\}\right), \textit{ kaikilla } B \in \mathcal{B}$$

on todennäköisyysmitta joukossa $(\mathbb{R}, \mathcal{B})$, ja sitä kutsutaan satunnaismuuttujan X jakaumaksi.

Todistus. Täytyy osoittaa, että P_X täyttää todennäköisyysmitan määritelmän kolme ehtoa:

1. Kaikilla $B \in \mathcal{B}$, $P_X(B) \geq 0$.
2. $P_X(\mathbb{R}) = P(\{\omega \in \Omega : X(\omega) \in \mathbb{R}\}) = P(\Omega) = 1$.
3. Olkoot $A_1, A_2, \ldots$ joukon $\mathcal{B}$ alkioita, jolle $A_i \cap A_j = \emptyset$ kaikilla $i \neq j$. Koska tapahtumat $\{X \in A_i\}$ ovat erillisiä, niin

$$\begin{aligned} P_X\left(\bigcup_{i=1}^{\infty} A_i\right) &= P\left(\left\{X \in \bigcup_{i=1}^{\infty} A_i\right\}\right) \\ &= P\left(\bigcup_{i=1}^{\infty} \{X \in A_i\}\right) \\ &= \sum_{i=1}^{\infty} P(\{X \in A_i\}) \\ &= \sum_{i=1}^{\infty} P_X(A_i). \end{aligned}$$

□

Esimerkki 2.43. Olkoot $\Omega = \{a, b, c\}$, $\Gamma = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$ ja P siten, että:

$$\begin{aligned} P(\emptyset) &= 0 \\ P(\{a\}) &= \frac{1}{5} \\ P(\{b, c\}) &= \frac{4}{5} \\ P(\Omega) &= 1. \end{aligned}$$

Olkoot $X : \Omega \to \mathbb{R}$ siten, että

$$X(\omega) = \begin{cases} 1, & \text{jos} \quad \omega = a \\ 2 & \text{jos} \quad \omega = b \quad \text{tai} \quad \omega = c. \end{cases}$$

Tällöin satunnaismuuttujan jakauma on esimerkiksi

$$\begin{aligned} &P_X(\emptyset) = 0 \\ &P_X(\{1\}) = P(\{a\}) = \frac{1}{5} \\ &P_X(\{2\}) = P(\{b, c\}) = \frac{4}{5} \\ &P_X(\mathbb{R}) = 1. \end{aligned}$$

$P_X(B)$ voitaisiin laskea kaikille Borel-joukoille kuten Lauseessa 2.42. Jakauman arvo riippuu siitä, kuuluvatko 1 ja/tai 2 joukkoon B.

Satunnaismuuttujan jakaumasta jatketaan kertymäfunktion määritelmään, joka tuli tutuksi jo lukioaiheita käsitellessä. Kuten alta huomataan tämä määritelmä ei muutu lukio-opetuksen ja yliopisto-opetuksen välillä.

Määritelmä 2.44. Olkoon X satunnaismuuttuja. Funktiota F_X joukossa $\mathbb{R}$

$$F_X(x) := P_X(]-\infty, x]) = P(X \leq x)$$

kutsutaan satunnaismuuttujan X kertymäfunktioksi.

Käydään seuraavaksi esimerkki siitä, miten kertymäfunktio määritellään.

Esimerkki 2.45. Olkoon $\Omega = \{a, b, c\}, \Gamma = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$ ja P annettu seuraavalla tavalla:

$$\begin{aligned} &P(\emptyset) = 0 \\ &P(\{a\}) = \frac{1}{5} \\ &P(\{b, c\}) = \frac{4}{5} \\ &P(\Omega) = 1. \end{aligned}$$

Olkoon $X : \Omega \to \mathbb{R}$ siten, että

$$X(\omega) = \begin{cases} 1, & \text{jos} \quad \omega = a \\ 2, & \text{jos} \quad \omega = b \quad \text{tai} \quad \omega = c. \end{cases}$$

Satunnaismuuttujan X kertymäfunktio määritellään $F_X(x) = P(X \leq x)$. Satunnaismuuttuja X saa arvoja 1 ja 2. Muodostetaan kertymäfunktio näille arvoille: Kun $x = 1$:

$$F_X(1) = P(X \leq 1) = P(\{\omega \in \Omega : X(\omega) \leq 1\}) = P(\{a\}) = \frac{1}{5}.$$

Kun $x = 2$:

$$\begin{aligned} F_X(2) &= P(X \leq 2) \\ &= P(\{\omega \in \Omega : X(\omega) \leq 2\}) \\ &= P(\{a\}) + P(\{b, c\}) \\ &= \frac{1}{5} + \frac{4}{5} \\ &= 1. \end{aligned}$$

Satunnaismuuttujan X kertymäfunktio on siis

$$F_X(x) = \begin{cases} 0, & \text{jos} \quad x < 1 \\ \frac{1}{5}, & \text{jos} \quad 1 \leq x < 2 \\ 1, & \text{jos} \quad x \geq 2. \end{cases}$$

Satunnaismuuttujan kertymäfunktiolla on tiettyjä ominaisuuksia. Seuraavassa Lauseessa käydään näitä läpi.

Lause 2.46. *Olkoon X satunnaismuuttuja avaruudessa (Ω, Γ, P). Kertymäfunktiolle F_X pätee seuraavat ehdot:*

1. *Jos $x < y$, niin $F_X(x) \leq F_X(y)$.*
2. $F_X(x^+) := \lim_{h \to 0^+} F_X(x + h) = F_X(x)$ *kaikilla* $x \in \mathbb{R}$.
3. $\lim_{x \to \infty} F_X(x) = 1$.
4. $\lim_{x \to -\infty} F_X(x) = 0$.

Todistus. 1. Jos $x < y$, niin:

$$\{\omega \in \Omega : X(\omega) \leq x\} \subseteq \{\omega \in \Omega : X(\omega) \leq y\}$$

$$F_X(x) = P(X \leq x) \leq P(X \leq y) = F_X(y).$$

2. Olkoon $x \in \mathbb{R}$. Oletetaan, että $(x_n)_n$ on reaalilukujen vähenevä jono raja-arvolla x. Eli

   $$x_1 \geq x_2 \geq \cdots > x \quad \text{ja} \quad \lim_{n\to\infty} x_n = x.$$

   Nähdään, että

   $$\{X \leq x_1\} \supseteq \{X \leq x_2\} \supseteq \cdots$$

   ja että

   $$\bigcap_{n=1}^{\infty} \{X \leq x_n\} = \{X \leq x\}.$$

   Siksi Lauseesta 2.14 seuraa, että

   $$F_X(x^+) = \lim_{n\to\infty} F_X(x_n) = \lim_{n\to\infty} P(X \leq x_n) = P(X \leq x) = F_X(x).$$

3. On selvää, että kaikille $n \in \mathbb{N}$ pätee

   $$\{X \leq n\} \subseteq \{X \leq (n+1)\}$$

   ja

   $$\Omega = \bigcup_{n=1}^{\infty} \{X \leq n\}.$$

   Eli

   $$\lim_{x\to\infty} F_X(x) = \lim_{n\to\infty} F_X(n) = \lim_{n\to\infty} P(X \leq n) = P(\Omega) = 1.$$

4. Kaikille $n \in \mathbb{N}$ pätee

   $$\{X \leq -n\} \supseteq \{X \leq -(n+1)\}$$

   ja lisäksi

   $$\emptyset = \bigcap_{n=1}^{\infty} \{X \leq -n\}.$$

   Näin ollen:

   $$\lim_{x\to\infty} F_X(x) = \lim_{n\to\infty} F_X(-n) = \lim_{n\to\infty} P(X \leq -n) = P(\emptyset) = 0.$$

$\square$

Voidaan osoittaa, että mikä tahansa funktio $F(x)$, joka toteuttaa edellisen Lauseen ehdot, on jonkin satunnaismuuttujan X kertymäfunktio.

Seuraus 2.47. *Olkoot X satunnaismuuttuja joukossa (Ω, Γ, P), F_X satunnaismuuttujan kertymäfunktio ja $a, b \in \mathbb{R}$ siten, että $a < b$. Tällöin:*

1. $F_X(x^-) := \lim_{h \to 0^+} F_X(x - h) = P(X < x)$.
2. $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$.
3. $P(a < X \leq b) = F_X(b) - F_X(a)$.
4. $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$.
5. $P(a < X < b) = F_X(b^-) - F_X(a)$.
6. $P(X = a) = F_X(a) - F_X(a^-)$.
7. *Jos $P(a < X < b) = 0$, niin F_X on vakio väliltä (a, b).*

Todistus. 1. Olkoon $x \in \mathbb{R}$. Oletetaan, että $(x_n)_n$ on kasvava reaalilukujen jono raja-arvolla x. Eli

$$x_1 \leq x_2 \leq \cdots < x \quad \text{ja} \quad \lim_{n \to \infty} x_n = x.$$

On selvää, että

$$\{X \leq x_1\} \subseteq \{X \leq x_2\} \subseteq \cdots$$

ja

$$\bigcup_{n=1}^{\infty} \{X \leq x_n\} = \{X < x\}.$$

Lauseesta 2.14 seuraa

$$F_X(x^-) = \lim_{n \to \infty} F_X(x_n) = \lim_{n \to \infty} P(X \leq x_n) = P(X < x).$$

2.

$$\Omega = \{X < a\} \cup \{a \leq X \leq b\} \cup \{X > b\}.$$

Tästä saadaan, että

$$1 = P(X < a) + P(a \leq X \leq b) + P(X > b).$$

Eli

$$P(a \leq X \leq b) = 1 - P(X < a) - P(X > b).$$

Nyt hyödyntämällä kohtaa 1 ja koska $1 - P(X > b) = P(X \leq b)$ saadaan, että

$$P(a \leq X \leq b) = P(X \leq b) - F_X(a^-).$$

Koska Määritelmän 2.44 nojalla $P(X \leq b) = F_X(b)$, niin

$$P(a \leq X \leq b) = F_X(b) - F_X(a^-).$$

Jätetään loput todistukset harjoitustehtäväksi. $\square$

Esimerkki 2.48. Olkoon X satunnaismuuttuja, jonka kertymäfunktio F_X on

$$F_X(x) = \begin{cases} 0, & \text{jos} \quad x < 0 \\ \frac{x^2}{4}, & \text{jos} \quad 0 \leq x < 2 \\ 1, & \text{jos} \quad x \geq 2, \end{cases}$$

missä F_X on jatkuva. Tällöin myös satunnaismuuttuja X on jatkuvasti jakautunut satunnaismuuttuja.

Huomautus 2.49. Jos X on jatkuva satunnaismuuttuja todennäköisyysavaruudessa (Ω, Γ, P), niin tällöin $F(a) = F(a^-)$ ja koska $P(X = a) = F(a) - F(a^-) = 0$, niin $P(X = a) = 0$ kaikilla $a \in \mathbb{R}$.

2.6 Diskreetti satunnaismuuttuja

Satunnaismuuttujat luokitellaan niiden kertymäfunktion mukaan. Jos kertymäfunktio on porrasfunktio, sanotaan, että X on diskreetti satunnaismuuttuja. Jos taas X on absoluuttisesti jatkuva funktio, niin X on jatkuva satunnaismuuttuja. Tässä alaluvussa käsitellään näistä diskreettiä satunnaismuuttujaa.

Määritelmä 2.50. Olkoon X satunnaismuuttuja ja F_X sen kertymäfunktio. Sanotaan, että kertymäfunktio F_X on epäjatkuva pisteessä $a \in \mathbb{R}$ jos:

$$F_X(a) - F_X(a^-) \neq 0.$$

Huomautus 2.51. Jos X on jatkuva satunnaismuuttuja, niin epäjatkuvuuspisteiden joukko on tyhjä joukko.

Seuraava määritelmä käsittelee pistetodennäköisyyttä.

Määritelmä 2.52. Olkoon X diskreetti satunnaismuuttuja, jonka arvoja ovat $x_1, x_2, \ldots$. Funktiota p_X joukossa $\mathbb{R}$

$$p_X(x) = \begin{cases} P(X = x_i) & \text{jos} \quad x = x_1, x_2, \ldots \\ 0, \text{ muuten} \end{cases}$$

kutsutaan satunnaismuuttujan X pistetodennäköisyysfunktioksi.

Pistetodennäköisyysfunktiota käsitellään lukiossa Lauseena, jolle pätee seuraavan huomautuksen toinen kohta.

Huomautus 2.53. Pistetodennäköisyysfunktiolle pätee:

1. $p(x_i) \geq 0$ kaikilla i ja
2. $\sum_i p(x_i) = 1$.

Esimerkki 2.54. Olkoon X satunnaismuuttuja, jonka kertymäfunktio on

$$F_X(x) = \begin{cases} 0, & \text{jos} \quad x < -2 \\ \frac{1}{7}, & \text{jos} \quad -2 \leq x < \frac{1}{2} \\ \frac{4}{7}, & \text{jos} \quad \frac{1}{2} \leq x < \sqrt{2} \\ 1, & \text{jos} \quad x \geq \sqrt{2}. \end{cases}$$

Pistetodennäköisyysfunktio voidaan määrittää kertymäfunktion avulla seuraavasti: Seurauksen 2.47 nojalla $P(X = a) = F_X(a) - F_X(a^-)$ eli

$$\begin{aligned} P(X = -2) &= F_X(-2) - F_X(-2^-) \\ &= \frac{1}{7} - 0 \\ &= \frac{1}{7}, \end{aligned}$$

$$\begin{aligned} P(X = \frac{1}{2}) &= F_X(\frac{1}{2}) - F_X(\frac{1}{2}^-) \\ &= \frac{4}{7} - \frac{1}{7} \\ &= \frac{3}{7}, \end{aligned}$$

ja

$$\begin{aligned} P(X = \sqrt{2}) &= F_X(\sqrt{2}) - F_X(\sqrt{2}^-) \\ &= 1 - \frac{4}{7} \\ &= \frac{3}{7}. \end{aligned}$$

Satunnaismuuttujan X pistetodennäköisyysfunktio on siis

$$p_X(x) = \begin{cases} \frac{1}{7}, & \text{jos} \quad x = -2 \\ \frac{3}{7}, & \text{jos} \quad x = \frac{1}{2} \\ \frac{3}{7}, & \text{jos} \quad x = \sqrt{2} \\ 0, & \text{muulloin.} \end{cases}$$

Esimerkki 2.55. Kolikkoa heitetään kahdesti. Olkoon X saatujen klaavojen määrä.

1. Muodosta satunnaismuuttujan X pistetodennäköisyysfunktio.
2. Laske $P(0,5 < X \leq 4)$, $P(-1,5 \leq X < 1)$ ja $P(X \leq 2)$.

Ratkaisu:

1. Satunnaismuuttuja X saa arvoja $0, 1$ tai 2 eli:

$$\begin{aligned}
P(X = 0) &= P((KR, KR)) = \frac{1}{4} \\
P(X = 1) &= P((KR, KL)) + P((KL, KR)) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\
P(X = 2) &= P((KL, KL)) = \frac{1}{4}.
\end{aligned}$$

Joten pistetodennäköisyysfunktio on

$$p_X(x) = \begin{cases} \frac{1}{4}, & \text{jos} \quad x = 0 \\ \frac{1}{2}, & \text{jos} \quad x = 1 \\ \frac{1}{4}, & \text{jos} \quad x = 2. \end{cases}$$

2.

$$\begin{aligned}
&P(0,5 < X \leq 4) = P(X = 1) + P(X = 2) = \frac{3}{4} \\
&P(-1,5 \leq X < 1) = P(X = 0) = \frac{1}{4} \\
&P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 1.
\end{aligned}$$

2.7 Jatkuvat satunnaismuuttujat

Tämä alaluku käsittelee jatkuvia satunnaismuuttujia. Jatkuvat satunnaismuuttujat on määritelty seuraavalla tavalla.

Määritelmä 2.56. Olkoon X satunnaismuuttuja todennäköisyysavaruudessa (Ω, Γ, P). Satunnaismuuttuja X on absoluuttisesti jatkuva, jos ja vain jos on olemassa epänegatiivinen ja integroituva funktio f_X siten, että kaikilla $x \in \mathbb{R}$ pätee

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt.$$

Funktiota f_X kutsutaan jatkuvan satunnaismuuttujan tiheysfunktioksi.

Huomautus 2.57. Tiheysfunktiolle f pätee seuraavat ehdot:

a) $f(x) \geq 0$ melkein kaikilla x:n arvoilla

b) $\int_{-\infty}^{\infty} f(x)\,dx = 1.$

Esimerkki 2.58. Olkoon X satunnaismuuttuja, jonka kertymäfunktio on

$$F_X(x) = \begin{cases} 0, \text{ jos} & x < 0 \\ x, \text{ jos} & 0 \leq x < 1 \\ 1, \text{ jos} & x \geq 1. \end{cases}$$

Satunnaismuuttujan X eräs tiheysfunktio f_X on:

$$f_X(x) = \begin{cases} 1, & \text{jos} \quad 0 \leq x < 1 \\ 0, & \text{muulloin.} \end{cases}$$

2.8 Satunnaismuuttujan muunnokset

Oletetaan, että X on satunnaismuuttuja, joka on määritelty todennäköisyysavaruudessa (Ω, Γ, P) ja olkoon g funktio siten, että $Y = g(X)$ on satunnaismuuttuja. Seuraavaksi halutaan määrittää satunnaismuuttujan Y tiheysfunktio satunnaismuuttujan X tiheysfunktion suhteen.

Lause 2.59. *Olkoon X absoluuttisesti jatkuva satunnaismuuttuja, jonka tiheysfunktio on f_X. Jos h on aidosti monotoninen ja derivoituva funktio, niin tällöin satunnaismuuttujan $Y = h(X)$ tiheysfunktio on:*

$$f_Y(y) = \begin{cases} f_X(h^{-1}(y)) \mid \frac{d}{dy}h^{-1}(y) \mid & \textit{jos } y = h(x) \textit{ jollekin } x \\ 0 & \textit{jos } y \neq h(x) \textit{ kaikilla } x, \end{cases}$$

missä h^{-1} on h:n käänteisfunktio.

Todistus. Oletetaan, että h on aidosti kasvava funktio ja olkoon $y \in \mathbb{R}$ siten, että $y = h(x)$ jollekin x.
Tällöin:

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(h(X) \leq y) \\ &= P(X \leq h^{-1}(y)) \\ &= F_X(h^{-1}(y)). \end{aligned}$$

Derivoimalla saadaan

$$\begin{aligned} f_Y(y) &= F'_X(h^{-1}(y))\frac{d}{dy}h^{-1}(y) \\ &= f_X(h^{-1}(y)) \mid \frac{d}{dy}h^{-1}(y) \mid \end{aligned}$$

olettaen, että h:n derivaatta on positiivinen.

Olkoon h aidosti vähenevä funktio ja $y \in \mathbb{R}$ siten, että $y = h(x)$, joillakin x. Tällöin

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(h(X) \leq y) \\ &= P(X \geq h^{-1}(y)) \\ &= 1 - F_X(h^{-1}(y)). \end{aligned}$$

Derivoimalla saadaan

$$\begin{aligned} f_Y(y) &= -f_X(h^{-1}(y))\frac{d}{dy}h^{-1}(y) \\ &= f_X(h^{-1}(y)) \mid \frac{d}{dy}h^{-1}(y) \mid, \end{aligned}$$

koska tässä tapauksessa h:n derivaatta on negatiivinen. Jos $y \in \mathbb{R}$ siten, että $y \neq h(x)$, kaikilla x, niin $F_Y(y) = 0$ tai $F_Y(y) = 1$ eli $f_Y(y) = 0$. $\square$

Esimerkki 2.60. Olkoon X satunnaismuuttuja, joka noudattaa jakaumaa

$$f_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, -\infty < x < \infty.$$

Määritä $Y = e^X$ tiheysfunktio.

Ratkaisu: Kun merkitään $h(X) = e^X$, niin h on aidosti kasvava ja derivoituva ja $Y = h(X)$. Nyt:

$$y = h(x) = e^x.$$

Tällöin:

$$\begin{aligned} h^{-1}(y) = x = \ln y \\ \frac{dh^{-1}(y)}{dy} = \frac{1}{y}. \end{aligned}$$

Koska $Y = e^X$ saadaan, että:

$$f_Y(y) = f_X(h^{-1}(y))\left|\frac{dh^{-1}(y)}{dy}\right| = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\ln y)^2}\frac{1}{y}.$$

Siksi satunnaismuuttujan Y tiheysfunktio on

$$f_Y(y) = \begin{cases} \frac{1}{y\sqrt{2\pi}}e^{-\frac{1}{2}(\ln y)^2} & \text{jos} \quad y > 0 \\ 0, \quad \text{muulloin.} \end{cases}$$

2.9 Satunnaismuuttujan odotusarvo ja varianssi

Tämän alaluvun aiheena ovat satunnaismuuttujan odotusarvo ja varianssi. Lukiomatematiikassa opittiin jo laskemaan satunnaismuuttujan odotusarvoja ja keskihajontoja. Tässä alaluvussa opitaan laskemaan satunnaismuuttujien variansseja ja opitaan mikä ero on varianssilla ja keskihajonnalla. Käydään ensin läpi odotusarvon määritelmä ja siihen liittyviä esimerkkejä.

Määritelmä 2.61. Olkoon X satunnaismuuttuja todennäköisyysavaruudessa (Ω, Γ, P).

1. Jos X on diskreetti satunnaismuuttuja, jonka arvoja ovat $x_1, x_2, \dots,$ niin satunnaismuuttujalla X on odotusarvo jos:

$$\sum_{k=1}^{\infty} |x_k| P(X = x_k) < \infty.$$

   Odotusarvo määritellään seuraavasti:

$$E(X) = \sum_{k=1}^{\infty} x_k P(X = x_k).$$

2. Jos X on jatkuva satunnaismuuttuja, jolla on tiheysfunktio f_X, niin satunnaismuuttujalla X on odotusarvo jos:

$$\int_{-\infty}^{\infty} |x| f_X(x) < \infty.$$

   Tällöin odotusarvo määritellään seuraavasti:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx.$$

Huomautus 2.62. Jos X on satunnaismuuttuja, jonka arvojoukko on äärellinen, niin $E(X)$ on aina olemassa.

Esimerkki 2.63. Heitetään tavallista noppaa kerran ja olkoon satunnaismuuttuja X nopan silmäluku, joka saadaan heitosta. Tällöin

$$E(X) = \sum_{k=1}^{6} \frac{k}{6} = \frac{21}{6}.$$

Esimerkki 2.64. Olkoon X diskreetti satunnaismuuttuja, jonka pistetodennäköisyysfunktio on

$$p(x) = \begin{cases} e^{-3} \frac{3^x}{x!} & \text{jos} \quad x = 0, 1, 2, \dots \\ 0, & \text{muulloin.} \end{cases}$$

Tällöin

$$\begin{aligned}
E(X) &= \sum_{k=1}^{\infty} k e^{-3} \frac{3^k}{k!} \\
&= e^{-3} \sum_{k=1}^{\infty} \frac{3^k}{(k-1)!} \\
&= e^{-3} \sum_{j=0}^{\infty} \frac{3^{j+1}}{j!} \\
&= 3e^{-3} \sum_{j=0}^{\infty} \frac{3^j}{j!} \\
&= 3e^{-3}e^{3} = 3.
\end{aligned}$$

Seuraavan tuloksen avulla voidaan löytää absoluuttisesti jatkuvan satunnaismuuttujan odotusarvo sen kertymäfunktion avulla.

Lause 2.65. *Olkoot Y jatkuva satunnaismuuttuja ja funktio f sen tiheysfunktio. Jos $E(Y)$ on olemassa, niin tällöin:*

$$E(Y) = \int_0^{\infty} [1 - F_Y(y)]dy - \int_0^{\infty} F_Y(-y)dy.$$

Todistus. Käytetään todistuksen apuna Fubinin lausetta ja saadaan, että

$$\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty} x f(x)\, dx \\
&= \int_0^{\infty} x f(x)\, dx + \int_{-\infty}^{0} x f(x)\, dx \\
&= \int_0^{\infty} \left(\int_0^{x} dy \right) f(x)\, dx - \int_{-\infty}^{0} \left(\int_0^{x} dy \right) f(x)\, dx \\
&= \int_0^{\infty} \int_y^{\infty} f(x)\, dx dy - \int_0^{\infty} \int_{-\infty}^{-y} f(x)\, dx dy \\
&= \int_0^{\infty} P(Y > y) dy - \int_0^{\infty} P(Y \leq -y) dy \\
&= \int_0^{\infty} [1 - F_Y(y)]dy - \int_0^{\infty} F_Y(-y)dy.
\end{aligned}$$

$\square$

Seuraava Lause, joka tunnetaan nimellä ”Law of the unconscious statistician” eli LOTUS, ilmaisee satunnaismuuttujan $Y = g(X)$ odotusarvon tiheysfunktion f_X ja funktion g avulla.

Lause 2.66. *Olkoon X satunnaismuuttuja, jonka tiheysfunktio on f_X ja olkoon $g : \mathbb{R} \to \mathbb{R}$ funktio siten, että $Y = g(X)$ on satunnaismuuttuja. Tällöin*

$$E(g(X)) = \begin{cases} \sum_x g(x) p_X(x) & \text{jos } X \text{ on diskreetti satunnaismuuttuja} \\ \int_{-\infty}^{\infty} g(x) f_X(x)\, dx & \text{jos } X \text{ on absoluuttisesti jatkuva satunnaismuuttuja} \end{cases}$$

aina kun summa tai integraali konvergoi absoluuttisesti.

Todistus.

1. Olkoon X diskreetti satunnaismuuttuja, joka saa arvoja $x_1, x_2, \cdots$. Tässä tapauksessa pistetodennäköisyysfunktio on

$$p_X(x) = \begin{cases} P(X = x) & \text{jos} \quad x = x_1, x_2, \cdots \\ 0 & \text{muulloin} \end{cases}$$

Satunnaismuuttuja $Y = g(X)$ saa arvoja $g(x_1), g(x_2), \cdots$. On selvää, että jotkin näistä luvuista voivat olla samat. Oletetaan, että $y_j, j \geq 1$ esittää $g(x_i)$ eri arvoja. Keräämällä yhteen kaikki $g(x_i)$, jotka ovat samat, saadaan, että

$$\begin{aligned} \sum_i g(x_i) p_X(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i) p_X(x_i) \\ &= \sum_j y_j \sum_{i:g(x_i)=y_j} p_X(x_i) \\ &= \sum_j y_j P(g(X) = y_j) \\ &= \sum_j y_j P(Y = y_j) \\ &= E(Y) = E(g(X)). \end{aligned}$$

2. Olkoon X absoluuttisesti jatkuva satunnaismuuttuja, jonka tiheysfunktio on f_X. Oletetaan, että g on epänegatiivinen funktio, yleisen tapauksen todistus sivuutetaan. Lauseen 2.65 nojalla

$$\begin{aligned} E(g(X)) &= \int_0^{\infty} P(g(X) > y) dy - \int_0^{\infty} P(g(X) \leq -y) dy \\ &= \int_0^{\infty} P(g(X) > y) dy \\ &= \int_0^{\infty} \left(\int_B f_X(x)\, dx \right) dy \end{aligned}$$

missä $B := \{x : g(x) > y\}$. Siksi

$$\begin{aligned} E(g(X)) &= \int_0^{\infty} \int_0^{g(x)} f_X(x)\, dy dx \\ &= \int_0^{\infty} g(x) f_X(x)\, dx. \end{aligned}$$

□

Esimerkki 2.67. Olkoon X jatkuva satunnaismuuttuja, jonka tiheysfunktio on

$$f(x) = \begin{cases} 2x, & \text{jos} \quad 0 < x < 1 \\ 0, & \text{muulloin.} \end{cases}$$

Tällöin:

$$\begin{aligned} E(X^3) &= \int_{-\infty}^{\infty} x^3 f(x)\,dx \\ &= \int_0^1 2x^4\,dx \\ &= \frac{2}{5}. \end{aligned}$$

Siirrytään käsittelemään odotusarvon tärkeimpiä ominaisuuksia.

Lause 2.68. *Olkoon X satunnaismuuttuja.*

1. *Jos $P(X \geq 0) = 1$ ja $E(X)$ on olemassa, niin tällöin $E(X) \geq 0$.*
2. *$E(\alpha) = \alpha$, jokaiselle vakiolle α.*
3. *Jos X on rajoitettu, eli on olemassa vakio $M > 0$ siten, että $P(\mid X \mid \leq M) = 1$, niin $E(X)$ on olemassa.*
4. *Jos α ja β ovat vakioita ja jos g ja h ovat funktioita siten, että $g(X)$ ja $h(X)$ ovat satunnaismuuttujia, joiden odotusarvot ovat olemassa, niin $E(\alpha g(X) + \beta h(X))$ on olemassa ja:*

$$E(\alpha g(X) + \beta h(X)) = \alpha E(g(X) + \beta E(h(X)).$$

5. *Jos g ja h ovat funktioita siten, että $g(X)$ ja $h(X)$ ovat satunnaismuuttujia, joiden odotusarvot ovat olemassa ja jos $g(x) \leq h(x)$ kaikilla x, niin:*

$$E(g(X)) \leq E(h(X)).$$

Erityisesti

$$|E(X)| \leq E(|X|).$$

Todistus. 1. Oletetaan, että X on diskreetti satunnaismuuttuja, joka saa arvoja $x_1, x_2, \cdots$. Koska $P(X < 0) = 0$ saadaan, että $P(X = x_j) = 0$ kaikilla $x_j < 0$. Siksi:

$$\begin{aligned} E(X) &= \sum_i x_i P(X = x_i) \\ &= \sum_{i:x_i \geq 0} x_i P(X = x_i) \geq 0. \end{aligned}$$

Jos X on absoluuttisesti jatkuva satunnaismuuttuja, niin

$$\begin{aligned}
E(X) &= \int_0^\infty [1 - F_X(x)]\, dx - \int_0^\infty F_X(-x)dx \\
&= \int_0^\infty [1 - F_X(x)]\, dx \\
&= \int_0^\infty P(X > x)\, dx \geq 0.
\end{aligned}$$

2. $E(\alpha) = \alpha P(X = \alpha) = \alpha$.

3. Jos X on diskreetti satunnaismuuttuja, joka saa arvoja $x_1, x_2, \cdots$, niin koska $P(|X| > M) = 0$, voidaan olettaa, että $\{x_1, x_2, \cdots\} \subseteq [-M, M]$. Eli
$$\sum_i |x_i| P(X = x_i) \leq M \sum_i P(X = x_i) = M < \infty$$
Jos X on absoluuttisesti jatkuva satunnaismuuttuja, jonka tiheysfunktio on f, niin koska $P(|X| > M) = 0$, voidaan olettaa, että $f(x) = 0$ kaikilla $x \notin [-M, M]$. Siksi:
$$\int_{-\infty}^\infty |x| f(x)\, dx = \int_{-M}^M |x| f(x)\, dx \leq M \int_{-M}^M f(x)\, dx = M < \infty.$$

4. Käydään läpi jatkuvan satunnaismuuttujan tapaus ja jätetään diskreetti tapaus harjoitustehtäväksi. Jos X on absoluuttisesti jatkuva satunnaismuuttuja, jonka tiheysfunktio on f, niin:
$$\begin{aligned}
\int_{-\infty}^\infty |\alpha g(x) + \beta h(x)| f(x)\, dx &\leq \int_{-\infty}^\infty |\alpha g(x)| f(x)\, dx + \int_{-\infty}^\infty |\beta h(x)| f(x)\, dx \\
&= |\alpha| \int_{-\infty}^\infty |g(x)| f(x)\, dx \\
&\quad + |\beta| \int_{-\infty}^\infty |h(x)| f(x)\, dx < \infty
\end{aligned}$$
Odotusarvo $(\alpha g(x) + \beta h(X))$ on olemassa ja on selvää, että
$$\begin{aligned}
\int_{-\infty}^\infty (\alpha g(x) + \beta h(x)) f(x)\, dx &= \int_{-\infty}^\infty (\alpha g(x)) f(x)\, dx + \int_{-\infty}^\infty (\beta h(x)) f(x)\, dx \\
&= \alpha E(g(X)) + \beta E(h(X)).
\end{aligned}$$

5. Jätetään diskreetti tapaus harjoitustehtäväksi. Jos X on absoluuttisesti jatkuva satunnaismuuttuja, jonka tiheysfunktio on f, niin tällöin
$$E(g(X)) = \int_{-\infty}^\infty g(x) f(x)\, dx \leq \int_{-\infty}^\infty h(x) f(x)\, dx = E(h(X)).$$

Koska

$$-|X| \leq X \leq |X|$$

saadaan

$$\begin{aligned} E(-|X|) &\leq E(X) \leq E|X| \\ -E(|X|) &\leq E(X) \leq E(|X|). \end{aligned}$$

Eli

$$|E(X)| \leq E(|X|).$$

□

Satunnaismuuttujan odotusarvoa kutsutaan ensimmäiseksi momentiksi. Määritellään seuraavaksi momentit yleisemmin. Momentit ovat tunnuslukuja, jotka kuvaavat satunnaismuuttujien pistetodennäköisyysfunktioiden jakautumista.

Määritelmä 2.69. Olkoon X satunnaismuuttuja. Satunnaismuuttujan r:s momentti on

$$\mu_r^{'} := E(X^r)$$

aina kun odotusarvo on olemassa.

Lause 2.70. *Jos X on satunnaismuuttuja siten, että $\mu_r^{'}$ on olemassa, niin tällöin myös $\mu_s^{'}$ on olemassa kaikilla $s < r$.*

Todistus. Käydään läpi jatkuva tapaus ja jätetään diskreetti tapaus harjoitustehtäväksi. Koska

$$|x^s| < 1 + |x^r| \quad \text{kaikilla} \quad x \in \mathbb{R}$$

saadaan

$$\int_{-\infty}^{\infty} |x^s| f_X(x)\, dx \leq \int_{-\infty}^{\infty} (1 + |x^r|) f_X(x)\, dx = 1 + \int_{-\infty}^{\infty} |x^r| f_X(x)\, dx < \infty.$$

□

Määritelmä 2.71. Olkoon X satunnaismuuttuja. Satunnaismuuttujan r:s keskusmomentti on

$$\mu_r := E\left([X - E(X)]^r\right)$$

aina kun odotusarvo $E(X)$ on olemassa.

Lause 2.72. *Olkoon X satunnaismuuttuja, jonka odotusarvo on olemassa. Jos μ_r on olemassa, niin:*

$$\mu_r = \sum_{k=0}^{r} \binom{r}{k} \mu_k^{'} (-\mu_1^{'})^{r-k}.$$

Todistus. Lausetta 2.68 ja Newtonin binomikaavaa käyttämällä saadaan, että

$$\begin{aligned}
\mu_r &= E(X - E(X))^r \\
&= E\left(X - \mu_1^{'}\right)^r \\
&= E\left(\sum_{k=0}^{r} \binom{r}{k} X^k (-\mu_1^{'})^{r-k}\right) \\
&= \sum_{k=0}^{r} \binom{r}{k} E(X^k)(-\mu_1^{'})^{r-k} \\
&= \sum_{k=0}^{r} \binom{r}{k} \mu_k^{'}(-\mu_1^{'})^{r-k}.
\end{aligned}$$

□

Perehdytään tarkemmin toiseen keskitettyyn momenttiin eli varianssiin ja sen ominaisuuksiin. Varianssi kuvaa satunnaismuuttujan arvojen vaihtelua odotusarvosta. Keskihajonnalla σ ja varianssilla on yhteys. Keskihajonta on varianssin neliöjuuri eli varianssi on σ^2.

Lause 2.73. *Olkoon X satunnaismuuttuja, jonka odotusarvo on olemassa ja $\alpha, \beta \in \mathbb{R}$ ovat vakioita. Tällöin:*

1. $Var(X) \geq 0$
2. $Var(\alpha) = 0$
3. $Var(\alpha X) = \alpha^2 Var(X)$
4. $Var(X + \beta) = Var(X)$
5. $Var(X) = 0$, *jos ja vain jos* $P(X = E(X)) = 1$.

Todistus. 1. Todistus seuraa varianssin määritelmästä ja odotusarvon ominaisuuksista

2.
$$Var(\alpha) = E(\alpha - E(\alpha))^2 = E(0) = 0.$$

3.
$$\begin{aligned}
Var(\alpha X) &= E\left[\alpha X - E(\alpha X)\right]^2 \\
&= E(\alpha X - \alpha E(X))^2 \\
&= \alpha^2 E(X - E(X))^2 \\
&= \alpha^2 Var(X).
\end{aligned}$$

4.

$$\begin{aligned} Var(X+\beta) &= E\left[(X+\beta) - E(X+\beta)\right]^2 \\ &= E\left[X+\beta - E(X) - \beta\right]^2 \\ &= Var(X). \end{aligned}$$

5. a) Jos $X = E(X)$ todennäköisyydellä 1, niin on selvää, että

$$Var(X) = 0.$$

b) Oletetaan, että $Var(X) = 0$ ja olkoon $a := E(X)$.
Jos $P(X = a) < 1$, niin silloin on olemassa $c > 0$ siten, että

$$P((X-a)^2 > c) > 0,$$

koska

$$(x-a)^2 \geq c\mathcal{X}_{\{(x-a)^2 > c\}}.$$

Seuraavan päättelyn taustalla on Markovin epäyhtälö, josta saadaan, että

$$\begin{aligned} E(X-a)^2 &\geq E(c\mathcal{X}_{\{(x-a)^2 > c\}}) \\ Var(X) &\geq cE(\mathcal{X}_{\{(x-a)^2 > c\}}) \\ Var(X) &\geq cP((X-a)^2 > c) > 0, \end{aligned}$$

mikä on ristiriita ja siksi $P(X = E(X)) = 1$.

□

Varianssi voidaan laskea ensimmäisen ja toisen momentin avulla seuraavasti:

Lause 2.74. *Olkoon X satunnaismuuttuja, jolle $E(X^2)$ on olemassa. Tällöin*

$$Var(X) = E(X^2) - (E(X))^2.$$

Todistus.

$$\begin{aligned} Var(X) &= E(X - E(X))^2 \\ &= E(X^2 - 2XE(X) + (E(X))^2) \\ &= E(X^2) - 2(E(X)E(X) + E(E(X))^2 \\ &= E(X^2) - 2(E(X))^2 + (E(X))^2 \\ &= E(X^2) - (E(X))^2. \end{aligned}$$

□

Esimerkki 2.75. Heitetään tavallista noppaa kerran ja olkoon X satunnaismuuttuja, joka esittää saatuja tuloksia. Tiedetään, että $E(X) = \frac{21}{6}$ ja

$$E(X^2) = \sum_{k=1}^{6} \frac{1}{6}k^2 = \frac{91}{6}.$$

Tällöin

$$Var(X) = E(X^2) - (E(X))^2 = \frac{91}{6} - \left(\frac{21}{6}\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Näin saatiin käsiteltyä todennäköisyyslaskennan kehittymistä aina alakoulun matematiikasta mittateoriaan saakka. Tutkielma aloitettiin todennäköisyyslaskennan perusteista ja vähitellen päästiin käsittelemään aihetta todennäköisyysavaruudessa. Lukijalla tulisi tämän jälkeen olla kattava kuva siitä, mitä asioita milläkin luokka-asteilla käydään ja miten kaikki yhdistyy lopulta mittateoriaan.

Viitteet

[1] Asikainen, K., Kiviluoma, P., Lassila, T., Nikkinen, I., Nyrhinen, K., Perälä, P., Rantavuori, P., Rokka, P., Salminen, M., Stenberg, H., Tapiainen, T., Vehmas P.: *Tuhattaituri 6B*, Otava, 2018.

[2] Harsunkorpi, J., Heiskanen, P., Leikas, M., Saarelainen, M., Tahvanainen, J.: *Moodi MAA8 Tilastot ja todennäköisyys*, 1.-2. painos, Sanoma Pro Oy, 2023

[3] Harsunkorpi, J., Heiskanen, P., Leikas, M., Saarelainen, M., Tahvanainen, J., Viitanen, V.: *Moodi MAB5 Tilastot ja todennäköisyys*, 1.-3. painos, Sanoma Pro Oy, 2023

[4] Hassinen, S., Latva, O., Makkonen, J-P., Pirttimaa, M., Tolvanen, A.: *Kuutio 9 K11*, 13.-18. painos, Sanoma Pro Oy, 2023.

[5] Heitto, J.: *Todennäköisyyslaskennan haasteet lukio-opetuksessa*, Pro Gradu -tutkielma, Helsingin yliopisto, Helda, 2015. http://urn.fi/URN:NBN:fi-fe2017112251947

[6] Herranen, J., Loimisto, J., Mäkimarttunen, T., Pusa, J., Rahila, S., Tapiainen, T., Tikka, T.: *Ääretön Tilastot ja todennäköisyys digikirja*, Otava, 2023.

[7] Hähkiöniemi, M., Juhala, S., Juutinen, P., Laitinen, A., Louhikallio-Fomin Luoma-aho, S., E., Raittila, T., Tikka, T.: *Juuri 12 (LOPS21) digikirja*, Kustannusosakeyhtiö Otava, 2022.

[8] Häkkinen, K., Hänninen, L., Malinen, K., Ranta, P., Sohlman, L., Vallo, L.: *Milli 5A*, 1-4. painos, Sanoma Pro Oy, 2023.

[9] Juhala, S., Juutinen, P., Laitinen, A., Luoma-aho, E., Raittila, T., Tikka, T.: *Juuri 8 (LOPS21) digikirja*, Kustannusosakeyhtiö Otava, 2022.

[10] Viswanathan Arunachalam, Liliana Blanco Castañeda, Selvamuthu Dharmaraja: *Introduction to probability and stochastic processes with application.* Wiley, 2012.

[11] Opetushallitus: *Lukion opetussuunnitelman perusteet.* Opetushallitus, 20.

[12] Opetushallitus: *Perusopetuksen opetussuunnitelman perusteet.* 4. painos, Opetushallitus, 2014.

[13] Ottelin, J., Santavuori, T., Tauriainen, T., Vallineva S.: *Huippu 5 (LOPS21) digikirja*, Kustannusosakeyhtiö Otava, 2021