# Uskottavuusperiaate päättelyssä

Markus Laaksonen

LuK-tutkielma
Toukokuu 2026

MATEMATIIKAN JA TILASTOTIETEEN LAITOS

**Tarkastajat:**
FT Pekka Nieminen

TURUN YLIOPISTO, Matematiikan ja tilastotieteen laitos

LuK-tutkielma
**Pääaine:** Tilastotiede
**Tekijä:** Markus Laaksonen
**Otsikko:** Uskottavuusperiaate päättelyssä
**Ohjaaja:** FT Pekka Nieminen
**Sivumäärä:** 24 sivua
**Aika:** Toukokuu 2026

---

Tässä tutkielmassa käsitellään uskottavuusperiaatetta, jonka mukaan uskottavuusfunktio sisältää kaiken oleellisen tiedon tilastollisen mallin parametreista. Periaatteesta tarkastellaan erikseen sen heikkoa ja vahvaa muotoa. Merkittävä osa tutkielmaa on myös Birnbaumin lause, jonka mukaan vahva uskottavuusperiaate seuraa kahdesta tärkeästä tilastollisen päättelyn periaatteesta, eli tyhjentävyys- ja ehdollistamisperiaatteesta. Edellisistä varsinkin vahva uskottavuusperiaate ja Birnbaumin lause ovat olleet kiistanalaisia (frekventististen) tilastotieteilijöiden keskuudessa, eikä niiden käytöstä tai oikeellisuudesta ole päästy yksimielisyyteen.

Aluksi tutkielmassa esitellään yleisesti tilastollista päättelyä ja varsinkin uskottavuusfunktiota. Tämän jälkeen tarkastellaan erityisesti tyhjentävän tunnusluvun käsitettä sekä siihen perustuvaa tyhjentävyysperiaatetta. Sitten perehdytään hieman ehdollistamisperiaatteeseen, minkä jälkeen esitellään sekä todistetaan Birnbaumin lause. Lopuksi tarkastellaan kahta Birnbaumin lausetta kritisoivaa artikkelia.

Asiasanat: Uskottavuusperiaate, tyhjentävyysperiaate, ehdollistamisperiaate, Birnbaumin lause, uskottavuusfunktio

# Sisällys

# 1 Johdanto

Tilastollisessa analyysissa tavoitteena on useimmiten tutkia kiinnostuksen kohteena olevaa satunnaisilmiötä. Yleinen toimintatapa on kerätä satunnaisotannan avulla aineisto ilmiöstä, koska koko ilmiötä ei ole mahdollista tutkia. Tutkijaa tarkasteltavassa ilmiössä usein kiinnostaa erityisesti sen noudattama todennäköisyysjakauma; saamalla tietoa ilmiön jakaumasta opitaan lisää koko ilmiöstä, eikä vain kerätystä aineistosta.

Tällaisessa kontekstissa tutkijaa tarkemmin kiinnostaa todennäköisyysjakauman parametrien arvot, sillä ne ovat yleensä ainoa tuntematon tekijä tarkasteltavan satunnaisilmiön jakauman lausekkeessa. Esimerkkejä tällaisista parametreista voisivat olla normaalijakauman kohdalla sen odotusarvo ja varianssi. Tilanteessa, jossa tutkittavan ilmiön tiedetään noudattavan (likimain) normaalijakaumaa, pyrittäisiin siis tekemään päätelmiä näistä kahdesta parametrista ja sitä kautta myös ilmiön jakaumasta. Mutta mihin näiden päätelmien tulisi tarkalleen perustua? Toisin sanoen, minkä avulla päätelmiä tehtäisiin? Näihin kysymyksiin vastaa uskottavuusperiaate.

Uskottavuusperiaate on erityisen merkityksellinen, joskin kiistanalainen, periaate tilastollisessa päättelyssä. Periaatteen mukaan parametreja koskeva päättely tulee tehdä kokonaan aineistosta lasketun uskottavuusfunktion perusteella. Uskottavuusperiaatteen perusteella uskottavuusfunktio sisältää siis kaiken keskeisen informaation mallin parametreista. Tässä työssä keskitytäänkin suurimmalta osin käsittelemään uskottavuusperiaatteen heikkoa sekä vahvaa versiota ja niihin liittyviä tuloksia.

Yksi tärkeimmistä tuloksista on Allan Birnbaumin vuonna 1962 artikkelissaan *On the Foundations of Statistical Inference* esittelemä lause, joka on nimetty hänen mukaan [8]. Kyseisen Birnbaumin lauseen mukaan uskottavuusperiaatteen vahva versio seuraa kahdesta yleisestä tilastollisen päättelyn periaatteesta, eli tyhjentävyys- ja ehdollistamisperiaatteesta. Tällä tuloksella on ollut kauaskantoisia seurauksia. Lause on myös aiheuttanut runsasta keskustelua aina julkaisustaan saakka, ja itse lausetta sekä siihen kohdistunutta kritiikkiä käsitellään myös tutkielmassa.

Tutkielman luvussa 2 esitellään tutkielmassa käytettävät merkinnät ja oletukset lyhyesti sekä tarkastellaan erityisesti uskottavuusfunktiota. Uskottavuusfunktion tarkempi käsittely on perusteltua, jotta voidaan ymmärtää uskottavuusperiaate sekä sen sanoma. Luvussa 3 jatketaan varsinkin tyhjentävien tunnuslukujen ja tyhjentävyysperiaatteen määrittelemisellä. Seuraavaksi luvussa 4 määritellään heikko ja vahva uskottavuusperiaate sekä käsitellään periaatteeseen liittyviä ongelmakohtia. Luvussa 5 määritellään ehdollistamisperiaate. Tämän jälkeen luvussa 6 esitellään Birnbaumin lause ja todistetaan se. Lopuksi tarkastellaan Birnbaumin lauseen ja sen todistuksen kritiikkiä Evansin [6] ja Mayon [7] artikkelien pohjalta.

Tärkeimpinä lähteinä työssä on käytetty Yudi Pawitanin teosta *In All Likelihood: Statistical Modelling and Inference Using Likelihood* [3] ja Casellan ja Bergerin kirjaa *Statistical Inference* [4]. Näitä on käytetty erityisesti uskottavuusperiaatetta sekä Birnbaumin lausetta koskevissa osuuksissa. Tyhjentävyyden käsittelyssä tärkein lähde edellisten ohella on Pekka Niemisen *Tilastollisen päättelyn teoria* -kurssin luentomoniste [2].

Tutkielmassa lähestytään tilastollista päättelyä frekventistisestä näkökulmasta, joten bayesläinen näkemys asioihin jää käsittelemättä. Tämä on oleellinen huomio esimerkiksi uskottavuusfunktion ja parametrin $\theta$ määrittelemisessä. Tutkielmassa myös otoskoko $n$ ajatellaan kiinteäksi luvuksi, joka annetaan tilastollista mallia määriteltäessä, eikä se siis riipu aineistosta.

Tämän tutkielman kirjoituksessa on hyödynnetty ChatGPT-tekoälytyökalua kielenhuoltoon.

# 2 Tilastollisesta päättelystä

Aluksi pohjustetaan lyhyesti sitä viitekehystä, jota tarvitaan uskottavuusperiaatteen ymmärtämiseksi. Määritellään myös *uskottavuusfunktio* tarkemmin. Luku perustuu suurelta osin luentomonisteeseen *Tilastollinen päättely I ja II* [1, luvut 1–5]. Tämän lisäksi käytetyt merkinnät ovat yhdenmukaisia kyseisen luentomonisteen kanssa. Lähteenä uskottavuusfunktion käsittelyssä on käytetty myös [3, luku 2.1].

## 2.1 Tilastollinen malli

Käydään aluksi lyhyesti läpi, millaisia oletuksia tutkielmassa käytetään ja mistä lähtökohdista tilastollista päättelyä tarkastellaan.

Tutkielmassa oletetaan tilastollisen mallin ja vastaavan yhteispistetodennäköisyys/yhteistiheysfunktion eli yptf/ytf:än $f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)$ olevan tunnettu parametria $\theta$ lukuun ottamatta. Parametri oletetaan kiinteäksi mutta tuntemattomaksi. Parametrin kaikki mahdolliset arvot muodostavat parametriavaruuden, jonka merkintä on $\Omega \subset \mathbb{R}$. Tutkielmassa useimmiten oletetaan parametrin olevan yksiulotteinen.

Satunnaisvektorilla $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ kuvataan kiinnostuksen kohteena olevaa satunnaisilmiötä. Satunnaisvektorin realisaatiosta eli havaitusta aineistosta käytetään merkintää $\boldsymbol{y}$. Yleisenä oletuksena työssä on, että $\boldsymbol{Y}$:n alkiot ovat riippumattomia eli $Y_1, \ldots, Y_n \perp\!\!\!\perp$. Tällöin todennäköisyyslaskennan nojalla vastaava yptf/ytf voidaan muodostaa seuraavasti:

$$f_{\boldsymbol{Y}}(\boldsymbol{y};\theta) = f_{Y_1}(y_1;\theta)\cdots f_{Y_n}(y_n;\theta) = \prod_{i=1}^{n} f_{Y_i}(y_i;\theta). \tag{1}$$

## 2.2 Uskottavuusfunktio

Uskottavuusfunktio on yksi tilastollisen päättelyn keskeisimmistä työkaluista, jonka avulla voidaan tehdä päätelmiä parametrista $\theta$. Helpoin tapa lähestyä tätä käsitettä on klassisen esimerkin avulla.

**Esimerkki 1.** Heitetään kolikkoa 10 kertaa. Mahdollisia tulosvaihtoehtoja yhdellä heitolla on kaksi, joko kruuna tai klaava. Tavoitteena on tässä tilanteessa mallintaa todennäköisyyttä saada kruuna, eli sen ajatellaan olevan onnistuminen. Klaavan saaminen mielletään siis epäonnistumiseksi. Käytetään onnistumisen todennäköisyydestä merkintää $\theta$.

Oletetaan, että toistokokeiden seurauksena lopputulos on 3 kruunaa ja 7 klaavaa. Merkitään onnistumisten lukumäärää kuvaavaa satunnaismuuttujaa kirjaimella $K$

ja realisoitunutta määrää kirjaimella $k$. Tällöin havaittujen onnistumisten määrä on 3, eli $k = 3$.

Mitä tällöin voidaan aineiston perusteella sanoa todennäköisyydestä saada kolikonheitosta kruuna, eli mikä voisi olla tuntemattoman parametrin $\theta$ arvo? Oletetaan siis, että parametrin arvosta tai satunnaisilmiöstä ei ole ennalta tiedossa mitään ja että kaikki päätelmät tehdään aineiston perusteella. Varmuudella voidaan ainoastaan sanoa, että $0 < \theta < 1$, mutta tämän tarkempia arvioita ei pystytä tekemään. Tätä varten tarvitaankin siis työkaluksi uskottavuusfunktio, jolla pystytään arvioimaan parametrin $\theta$ eri arvoja ja vertaamaan, kuinka uskottavia ne ovat aineistoon suhteutettuna. [3, luku 2.1]

### 2.2.1 Varsinainen määritelmä

Uskottavuusfunktio itsessään määritellään seuraavalla tavalla:

**Määritelmä 1.** (Uskottavuusfunktio): Olkoon $\theta$ parametri, joka kuuluu parametriavaruuteen $\Omega \subset \mathbb{R}$. Tarkasteltavana on tilastollinen malli, jonka yptf/ytf on $f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta)$. Tällöin aineistoon $\boldsymbol{y}$ liittyvä uskottavuusfunktio on

$$L(\theta) = L(\theta; \boldsymbol{y}) = f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta), \qquad \theta \in \Omega.$$

Kyseessä on siis käytännössä sama yptf/ytf, joka määrittelee mallin. Tässä kuitenkin mielletään aineisto $\boldsymbol{y}$ kiinteäksi ja tuntematon parametri $\theta$ muuttujaksi, joka käy läpi kaikki sen mahdolliset arvot, kun taas yptf/ytf:än kohdalla toimitaan päinvastaisesti. Tätä havainnollistaa erityisesti seuraava yleinen merkintätapa, jota on käytetty esimerkiksi kirjassa *Statistical Inference* [4]:

$$L(\theta \mid \boldsymbol{y}) = f_{\boldsymbol{Y}}(\boldsymbol{y} \mid \theta).$$

Eli uskottavuusfunktiossa on ehdollistettu aineiston $\boldsymbol{y}$ suhteen, kun yptf/ytf:ssä taas on muodollisesti "ehdollistettu" parametrin $\theta$ suhteen. Kun tietyn muuttujan suhteen ehdollistetaan, ajatellaan sen olevan silloin kiinteä. Varsinaisesti ei ole kuitenkaan mahdollista ehdollistaa parametrin $\theta$ suhteen, sillä se ei ole satunnaismuuttuja. Tuossa tapauksessa kyseessä on ennemmin näkökulman vaihdosta kuin todellisesta ehdollistamisesta.

### 2.2.2 Uskottavuusfunktio käytännössä

Jatketaan sitten esimerkkiä 1 käyttäen tätä funktiota ja mietitään, mikä tässä tilanteessa voisi olla kyseinen uskottavuusfunktio. Tarkastelu tehdään luentomonisteen [1, esim. 5.2] esitystapaa seuraten. Diskreetissä tapauksessa, jossa on vain yksi havainto $k$, uskottavuusfunktion esitys on

$$L(\theta; k) = P_\theta(K = k).$$

Uskottavuusfunktion voi siis ajatella tarkoittavan funktiota sille, kuinka todennäköisesti satunnaismuuttuja $K$ on realisoitunut juuri täksi aineistoksi $k$ kullakin arvolla $\theta$.

Seuraavaksi täytyy pohtia, mikä tilastollinen malli tilanteeseen sopisi eli mitä jakaumaa satunnaismuuttuja $K$ noudattaa. Voidaan nähdä, että tähän tilanteeseen parhaiten soveltuu binomimalli, koska tässä on $n = 10$ kappaletta Bernoullikokeita. Koe siis koostuu kymmenestä riippumattomasta toistokokeesta, joista kukin joko onnistuu tai epäonnistuu. Tämä tarkoittaa, että $K$ noudattaa binomijakaumaa parametreilla $n$ ja $\theta$, eli $K \sim \text{Bin}(n, \theta)$. Määritelmän 1 nojalla uskottavuusfunktion yleinen muoto on

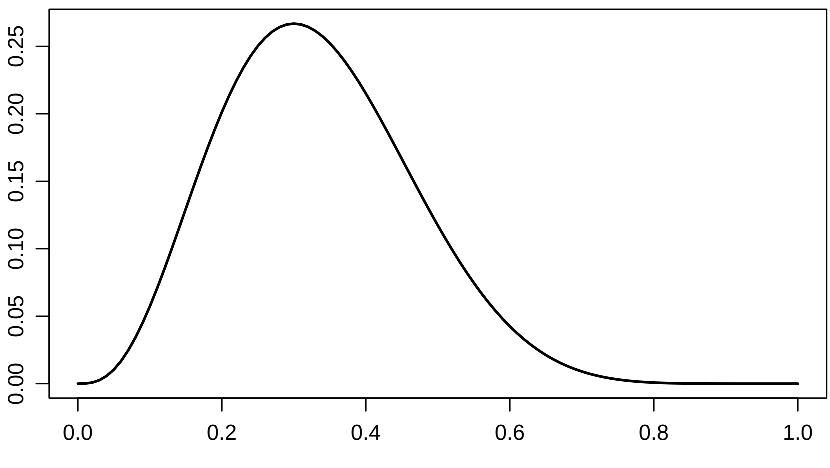
$$f_K(k; \theta) = P_\theta(K = k) = L(\theta; k) = \binom{n}{k} \theta^k (1 - \theta)^{n-k}, \qquad 0 < \theta < 1.$$

Korvaamalla $k$ ja $n$ esimerkin vastaavilla määrillä, eli $k = 3$ ja $n = 10$, saadaan funktiolle esitys

$$L(\theta; 3) = P_\theta(K = 3) = 120 \cdot \theta^3 (1 - \theta)^7, \qquad 0 < \theta < 1.$$

Miten tätä funktiota sitten voitaisiin varsinaisesti käyttää, ja mitä se kertoo? Funktion avulla pystytään vertailemaan erilaisia parametrin $\theta$ arvoja, ja arvioimaan, kuinka *uskottavia* ne ovat aineiston perusteella. Mitä suuremman arvon uskottavuusfunktio saa jollain arvolla $\theta$, sitä uskottavampi kyseinen $\theta$:n arvo on. Uskottavuusfunktion yhteydessä puhutaan nimenomaan uskottavuudesta (likelihood) eikä todennäköisyydestä (probability), koska parametri $\theta$ on tässä kiinteä, jolloin siihen ei liitetä todennäköisyysjakaumaa.

Matemaattisesti uskottavuusfunktion idean voisi ilmaista seuraavalla tavalla: Olkoot $\theta_1$ ja $\theta_2$ parametriavaruuden $\Omega$ pisteitä. Jos $L(\theta_1) = 2L(\theta_2)$, voidaan sanoa, että $\theta_1$ on aineiston perusteella kaksi kertaa uskottavampi kuin $\theta_2$. Helpoin tapa havainnollistaa tätä ideaa on esimerkin uskottavuusfunktion piirtäminen.




Kuva 1: Binomimallin uskottavuusfunktio, jossa $k = 3$ ja $n = 10$

Kuvasta 1 nähdään, mitkä arvot ovat aineiston valossa uskottavia ja mitkä taas eivät. Huomiot pätevät siis pelkästään tähän kyseiseen aineistoon. Kuvaajan perusteella tällä aineistolla uskottavin parametriarvo on $\theta = 0{,}3$, kuten saattaa päätellä heittojen $n = 10$ ja havaittujen klaavojen $k = 3$ määrästä. Samoin sen lähellä olevat 0,2 ja 0,4 sekä niiden välillä olevat arvot ovat lähes yhtä uskottavia tässä tilanteessa. Epäuskottavimmat arvot aineiston perusteella taas ovat 0,8 ja sitä suuremmat arvot, jossa funktion arvo on käytännössä nollaa. Tämähän on johdonmukainen ajatus, sillä jos havaittiin ainoastaan $k = 3$ kruunaa, ei ole uskottavaa, että todennäköisyys saada kruuna kolikonheitosta olisi suurempi kuin 0,8.

Aiemmin mainittua parametrin arvoa, jolla uskottavuusfunktio saa suurimman arvonsa, kutsutaan *suurimman uskottavuuden estimaatiksi*, lyhyemmin *su-estimaatiksi*. Tässä tutkielmassa ei esitellä käsitettä tämän tarkemmin, mutta olennaista on, että useimmiten uskottavuusfunktiota käytetään työkaluna nimenomaan selvittämään, mikä on parametrin su-estimaatti [3, luku 2.1]. Uskottavuusfunktio kuitenkin sisältää kaiken muunkin tiedon parametrista $\theta$, ei pelkästään uskottavuusfunktion maksimikohtaa. Huomionarvoista onkin usein, mitkä arvot eivät funktion eli aineiston perusteella ole uskottavia, sekä sellaiset arvot, jotka ovat lähes yhtä uskottavia kuin su-estimaatti. Tällaisten lähes yhtä uskottavien parametriarvojen selvittämiseen hyvä keino onkin uskottavuusvälin tai -alueen laskeminen.

Esitellään vielä yksi uskottavuusfunktioon liittyvä konsepti, joka on hyödyllinen myöhemmin uskottavuusperiaatteen määrittelyssä. Uskottavuusfunktion oleellisin tarkoitus on vertailla erilaisia parametriarvoja, joten funktion absoluuttisilla arvoilla ei niinkään ole merkitystä. Täten uskottavuusfunktio voidaan yksinkertaistettuna esittää myös ilman siinä esiintyviä vakiokertoimia. Vakiokertoimiksi lasketaan kaikki sellaiset tekijät uskottavuusfunktion lausekkeessa, joissa ei ole parametria $\theta$, eli parametrista riippumattomat tekijät. Tällaisessa tilanteessa käytetään merkintää

$$L(\theta) = L(\theta; \boldsymbol{y}) = c(\boldsymbol{y}) f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta) \propto f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta), \qquad \theta \in \Omega. \tag{2}$$

Tässä $\propto$-merkki tarkoittaa siis funktion lausekkeen merkitsemistä ilman siinä olevia vakiokertoimia, joista taas käytetään merkintää $c(\boldsymbol{y})$. Muuttuja $\boldsymbol{y}$ voi olla tuon lausekkeen $c$ sisällä myös joku muu, eli se voisi olla esimerkiksi muotoa $c(\boldsymbol{x}, \boldsymbol{y})$. Jos kaksi eri uskottavuusfunktiota ovat vakiota vaille samat, eli $L(\theta; \boldsymbol{y}) \propto L(\theta; \boldsymbol{x})$, niiden sanotaan olevan verrannolliset.

# 3 Tyhjentävyydestä

Tarkastellaan seuraavaksi tyhjentävyyden käsitettä, jonka ymmärtäminen helpottaa myöhempää uskottavuusperiaatteen määrittelemistä. Tärkeitä käsitteitä ovat *tyhjentävyysperiaate* (sufficiency principle) ja tämän määrittelyyn vaadittava *tyhjentävä tunnusluku* (sufficient statistic). Tyhjentävyyden käsittelyssä tärkeimpinä lähteinä on *Tilastollisen päättelyn teoria* -luentomoniste [2, luvut 1–4] sekä kirja *Statistical Inference* [4, luvut 6.1–6.2].

## 3.1 Tunnusluku ja tyhjentävä tunnusluku

Jos päättelyssä tarkasteltavan aineiston $\boldsymbol{y} = (y_1, ..., y_n)$ otoskoko $n$ on suuri, koko aineistoa saattaa olla vaikea tulkita. Tällöin usein pyritään tiivistämään aineistoa jollain tapaa ja tuomaan siitä kussakin tilanteessa kiinnostavia piirteitä esiin. Tällaista tiivistämistä tehdään yleensä laskemalla aineistosta *tunnuslukuja*. Tunnusluvut ovat aineiston perusteella laskettuja suureita eli toisin sanoen aineiston funktiota. Esimerkkejä tunnusluvuista on aineiston pienin ja suurin arvo $\min(\boldsymbol{y})$ ja $\max(\boldsymbol{y})$, sekä otoskeskiarvo $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ ja -varianssi $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$. Esimerkiksi pienimmästä ja suurimmasta arvosta voitaisiin myös muodostaa vektori $(\min(\boldsymbol{y}), \max(\boldsymbol{y}))$, joka olisi myös tunnusluku. Tunnusluvut voivat siis olla joko reaaliarvoisia tai vektoreita.

Käytetään tunnusluvun lasketusta arvosta merkintää $\boldsymbol{t} = \boldsymbol{t}(\boldsymbol{y})$ ja sitä vastaavasta satunnaismuuttujasta tai -vektorista merkintää $\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{Y})$. Tunnuslukujen laskeminen on keino tiivistää aineistoa. Tällöin, jos kahdelle aineistolle $\boldsymbol{x}$ ja $\boldsymbol{y}$ pätee ehto $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y})$, tunnusluvun perusteella nämä kaksi aineistoa ovat samanarvoisia, vaikka $\boldsymbol{x} \neq \boldsymbol{y}$. Käyttäen pelkkää tunnusluvun arvoa menetetään siis tietoa, joka on koko aineistossa jäljellä. *Tyhjentävän tunnusluvun* käsite vastaakin kysymykseen, milloin tunnusluku sisältää tarpeeksi tietoa parametrista $\theta$, jotta päättely voidaan perustaa ainoastaan tunnusluvun arvoon koko aineiston sijaan.

**Määritelmä 2.** (Tyhjentävä tunnusluku): Tunnusluku $\boldsymbol{T}(\boldsymbol{Y})$ on tyhjentävä tunnusluku parametrille $\theta$, jos satunnaisvektorin $\boldsymbol{Y}$ ehdollinen jakauma $\boldsymbol{Y} \mid \boldsymbol{T} = \boldsymbol{t}$ ei sisällä parametria $\theta$, eli se on siitä riippumaton.

Käytännössä tämä siis tarkoittaa sitä, että ehdollinen yptf/ytf eli $f_{\boldsymbol{Y}|\boldsymbol{T}}(\boldsymbol{y} \mid \boldsymbol{t}; \theta)$ ei sisällä parametria $\theta$. Määritelmän perusteella, jos tyhjentävän tunnusluvun $\boldsymbol{T}$ arvo on $\boldsymbol{t}$, se siis yksinään riittää, eikä $\boldsymbol{y}$ tuo enää mitään oleellista lisätietoa parametrista, mitä ei jo saataisi pelkästä tyhjentävästä tunnusluvusta. Kaikki tärkeä tieto on siis jo sisällytettynä tunnusluvussa $\boldsymbol{t}$.

**Esimerkki 2.** Jotta käsitteen merkitys ja tulkinta olisi selkeämpi, käydään läpi tämä vielä esimerkin kautta. Esimerkissä on vastaava tilanne kuin luvun 2.2.2 esimerkissä 1. Tässä kuitenkin lähestytään asiaa yleisesti siitä näkökulmasta, että on $n$ kappaletta toisistaan riippumattomia Bernoulli-jakautuneita satunnaismuuttujia, eikä kyseessä tarvitse olla nimenomaan kolikonheitto. Esimerkki perustuu pääosin luentomonisteeseen [2, esim. 1.3], mutta myös osittain kirjaan [4, esim. 6.2.3]

Olkoon $0 < \theta < 1$ ja aineisto $\boldsymbol{y} = (y_1, ..., y_n)$. Jokaiselle $y_i$ pätee

$$y_i = \begin{cases} 0, & \text{jos koe } i \text{ epäonnistuu,} \\ 1, & \text{jos koe } i \text{ onnistuu.} \end{cases}$$

Näitä alkioita vastaaville satunnaismuuttujille $Y_i$ pätee, että ne ovat riippumattomia ja noudattavat Bernoulli-jakaumaa parametrilla $\theta$, eli $Y_1, ..., Y_n \sim \mathrm{Ber}(\theta) \perp\!\!\!\perp$. Yksittäisen satunnaismuuttujan $Y_i$ pistetodennäköisyysfunktio on $f_{Y_i}(y_i; \theta) = \theta^{y_i}(1 - \theta)^{1-y_i}$, joten satunnaisvektorin $\boldsymbol{Y}$ yhteispistetodennäköisyysfunktio on yhtälön (1) perusteella

$$f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta) = P_\theta(\boldsymbol{Y} = \boldsymbol{y}) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \theta^{\sum y_i}(1 - \theta)^{n - \sum y_i},$$

jossa on käytetty hyödyksi seuraavia tuloksia: $\prod \theta^{y_i} = \theta^{\sum y_i}$ sekä $\sum(1 - y_i) = n - \sum y_i$.

Käytetään tässä tunnuslukuna onnistumisten lukumäärää $k = \sum_{i=1}^{n} y_i$. Perustelu tälle summaesitykselle on tietenkin se, että ainoat nollasta poikkeavat alkiot aineistossa ovat onnistumiset, joten summaamalla kaikki vektorin $\boldsymbol{y}$ alkiot saadaan yhteensä onnistumisten lukumäärä. Onnistumisten lukumäärää vastaava satunnaismuuttuja taas on $K$, ja sille pätee tunnetusti seuraava tulos: $K \sim \mathrm{Bin}(n, \theta)$. Tätä tulosta käytettiinkin jo luvussa 2.2.2, ja tuon luvun perusteella satunnaismuuttujan $K$ pistetodennäköisyysfunktio on $f_K(k; \theta) = P(K = k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}$.

Muodostetaan seuraavaksi yhteispistetodennäköisyysfunktio ehdolliselle jakaumalle $\boldsymbol{Y} \mid K$. Tässä käytetään ehdollisen todennäköisyyden kaavaa, jolla päästään muotoon

$$f_{\boldsymbol{Y}|K}(\boldsymbol{y} \mid k; \theta) = \frac{P(\boldsymbol{Y} = \boldsymbol{y}, K = k)}{P(K = k)}.$$

Käytetään oletuksena tässä, että $k = \sum y_i$. Jos tällaista oletusta ei käytettäisi, todennäköisyys $P(\boldsymbol{Y} = \boldsymbol{y}, K = k)$ olisi aina nolla, jolloin $f_{\boldsymbol{Y}|K}(\boldsymbol{y} \mid k; \theta)$ olisi nolla. Tämä ei ole mielenkiintoinen tapaus, joten käytetään kyseistä oletusta. Oletuksen ollessa voimassa tapahtumaan $\{K = k\}$ sisältyy tapahtuma $\{\boldsymbol{Y} = \boldsymbol{y}\}$. Täten ehdolliselle yhteispistetodennäköisyysfunktiolle saadaan seuraava muoto:

$$f_{\boldsymbol{Y}|K}(\boldsymbol{y} \mid k; \theta) = \frac{P(\boldsymbol{Y} = \boldsymbol{y})}{P(K = k)} = \frac{\theta^{\sum y_i}(1-\theta)^{n-\sum y_i}}{\binom{n}{k}\theta^k(1-\theta)^{n-k}} = \frac{1}{\binom{n}{k}} = \frac{1}{\binom{n}{\sum y_i}} \tag{3}$$

Tässä yhtälössä (3) käytettiin oletusta $k = \sum y_i$, jolloin osoittajassa oleva lauseke saatiin muotoon $\theta^{\sum y_i}(1-\theta)^{n-\sum y_i} = \theta^k(1-\theta)^{n-k}$, ja täten koko ehdollisen yhteispistetodennäköisyysfunktion lauseke saatiin sievennettyä lopulliseen muotoon $\frac{1}{\binom{n}{k}}$. Kuten tästä nähdään, ehdollinen jakauma ei sisällä enää parametria $\theta$, joten kyseinen tunnusluku $K$ on täten tyhjentävä tunnusluku.

### 3.1.1 Faktorointikriteeri

Seuraavaksi tarkastellaan tyhjentävyyteen liittyvää faktorointikriteeriä.

**Lause 1.** *(Faktorointikriteeri): Tunnusluku $\boldsymbol{T}(\boldsymbol{Y})$ on tyhjentävä tunnusluku parametrille $\theta$ jos ja vain jos mallin f yptf/ytf $f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta)$ voidaan esittää muodossa*

$$f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta) = h(\boldsymbol{y})g(\boldsymbol{t}(\boldsymbol{y}); \theta),$$

*jossa $h(\boldsymbol{y})$ ei sisällä parametria $\theta$, kaikilla $\boldsymbol{y}$ ja $\theta \in \Omega$.*

Faktorointikriteeri on hyödyllinen työkalu tyhjentävien tunnuslukujen etsimiseen. Sitä ei tässä erikseen todisteta, mutta todistus on kuitenkin luentomonisteessa [2, luku 2.5] ja kirjassa [4, lause 6.2.6]. Lauseen avulla nähdään merkittävä asia, nimittäin yhteys tyhjentävän tunnusluvun ja uskottavuusfunktion välillä. Uskottavuusfunktiohan alunperin määriteltiin seuraavalla tavalla: $L(\boldsymbol{\theta}; \boldsymbol{y}) = f_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\theta})$. Edellinen määritelmä käytännössä pätee siis myös uskottavuusfunktiolle, tai lauseessa voidaan korvata yptf/ytf $f_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\theta})$ uskottavuusfunktiolla $L(\boldsymbol{\theta}; \boldsymbol{y})$.

Käytetään seuraavaksi lausetta ja etsitään sen avulla toinen esimerkki tyhjentävästä tunnusluvusta. Esimerkki on hieman muokattuna monisteesta [2, esim. 2.3].

**Esimerkki 3.** Olkoon $Y_1, \ldots, Y_n$ otos normaalijakaumasta $\mathrm{N}(\mu, \sigma^2)$. Osoitetaan, että otoskeskiarvon $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ ja -varianssin $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ pari $\boldsymbol{t}(\boldsymbol{y}) = (\bar{y}, s^2)$ on tyhjentävä tunnusluku parametrille $\boldsymbol{\theta} = (\mu, \sigma^2)$. Parametri on tässä kaksiulotteinen, vastoin tutkielman yleistä oletusta. Muodostetaan parametrin $(\mu, \sigma^2)$ uskottavuusfunktion eräs hajotelma, joka on esitetty muun muassa [1, esim. 5.4]

$$L(\mu, \sigma^2; \boldsymbol{y}) = \frac{1}{(\sigma^2)^{n/2}} \exp\left\{-\frac{(n-1)s^2 + n(\bar{y} - \mu)^2}{2\sigma^2}\right\}.$$

Uskottavuusfunktio voidaan siis esittää muodossa $L(\mu, \sigma^2; \boldsymbol{y}) = h(\boldsymbol{y})g(\boldsymbol{t}(\boldsymbol{y}); \boldsymbol{\theta})$, jossa $g(\boldsymbol{t}(\boldsymbol{y}); \boldsymbol{\theta}) = \frac{1}{(\sigma^2)^{n/2}} \exp\left\{-\frac{(n-1)s^2+n(\bar{y}-\mu)^2}{2\sigma^2}\right\}$ riippuu aineistosta ainoastaan siitä laskettujen tunnuslukujen $\bar{y}$ ja $s^2$ kautta, ja $h(\boldsymbol{y}) = 1$ ei sisällä parametria $(\mu, \sigma^2)$. Tästä seuraa faktorointikriteerin eli lauseen 1 perusteella, että pari $(\bar{y}, s^2)$ on tyhjentävä tunnusluku parametrille $(\mu, \sigma^2)$.

Lauseen avulla voidaan myös osoittaa, että uskottavuusfunktio on itsessään tyhjentävä tunnusluku, mikä tehdäänkin seuraavaksi.

### 3.1.2 Uskottavuusfunktio ja tyhjentävyys

Aiheen käsittely ja lause 2 sekä siitä saatu tulos pohjautuu kirjaan [3, luku 3.2]. Osoitetaan ihan aluksi uskottavuusfunktion olevan tyhjentävä tunnusluku. Määritellään tunnusluku

$$\boldsymbol{t}(\boldsymbol{y}) = \frac{L(\theta; \boldsymbol{y})}{L(\theta_0; \boldsymbol{y})},$$

jossa $L(\theta; \boldsymbol{y})$ on koko uskottavuusfunktio kaikkien mahdollisten arvojen $\theta$ yli, ja $L(\theta_0; \boldsymbol{y})$ on uskottavuusfunktio mielivaltaisella arvolla $\theta_0$. Käyttäen faktorointikriteeriä voidaan osoittaa tämän olevan tyhjentävä, määrittelemällä $h(\boldsymbol{y}) = L(\theta_0; \boldsymbol{y})$ ja $g(\boldsymbol{t}(\boldsymbol{y}); \theta) = \frac{L(\theta; \boldsymbol{y})}{L(\theta_0; \boldsymbol{y})}$, jolloin

$$h(\boldsymbol{y})g(\boldsymbol{t}(\boldsymbol{y}); \theta) = L(\theta_0; \boldsymbol{y}) \cdot \frac{L(\theta; \boldsymbol{y})}{L(\theta_0; \boldsymbol{y})} = L(\theta; \boldsymbol{y}) = f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta).$$

Tällöin lauseen 1 nojalla $\boldsymbol{t}(\boldsymbol{y})$ on tyhjentävä tunnusluku. Tärkeänä huomiona määritelty $h(\boldsymbol{y}) = L(\theta_0; \boldsymbol{y})$ täyttää faktorintikriteerin ehdon, sillä se ei sisällä parametria $\theta$ muuttujana vaan ainoastaan sen yksittäisen arvon, joka on käytännössä pelkkä vakiokerroin.

Siirrytään tästä kuitenkin seuraavaksi tärkeään uskottavuusfunktion ominaisuuteen, joka pätee kaikille tyhjentäville tunnusluvuille.

**Lause 2.** *Oletetaan, että $\boldsymbol{y}$ on mallista $f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta)$ havaittu aineisto ja $\boldsymbol{t}(\boldsymbol{y})$ siitä laskettu tyhjentävä tunnusluku parametrille $\theta$. Tällöin koko aineiston $\boldsymbol{y}$ pohjalta laskettu uskottavuusfunktio on verrannollinen pelkästään tyhjentävän tunnusluvun avulla laskettuun uskottavuusfunktioon, eli $L(\theta; \boldsymbol{y}) \propto L(\theta; \boldsymbol{t}(\boldsymbol{y}))$.*

*Todistus.* [3, s. 56]. Muodostettaessa uskottavuusfunktiota ja siirryttäessä aineistosta $\boldsymbol{y}$ pelkästään tunnuslukuun $\boldsymbol{t}(\boldsymbol{y})$, siirrytään myös eri malliin. Eli siirryttäessä $\boldsymbol{y}$:n mallista $\boldsymbol{y}$:n ja $\boldsymbol{t}$:n yhteiseen malliin ja lopulta pelkkään $\boldsymbol{t}$:n malliin, tarvitaan siis ehdollisten jakaumien ketjusääntöä. Tämä tulee ilmi seuraavassa yhtälössä, jolla lause todistetaan:

$$\begin{aligned} L(\theta;\boldsymbol{y}) &= L(\theta;\boldsymbol{y},\boldsymbol{t}) = f_{\boldsymbol{Y},\boldsymbol{T}}(\boldsymbol{y},\boldsymbol{t};\theta) = f_{\boldsymbol{T}}(\boldsymbol{t};\theta) f_{\boldsymbol{Y}|\boldsymbol{T}}(\boldsymbol{y}\mid\boldsymbol{t}) \\ &= c(\boldsymbol{y})\cdot f_{\boldsymbol{T}}(\boldsymbol{t};\theta) \\ &= c(\boldsymbol{y})\cdot L(\theta;\boldsymbol{t}) \propto L(\theta;\boldsymbol{t}), \end{aligned} \tag{4}$$

jossa $c(\boldsymbol{y}) = f_{\boldsymbol{Y}|\boldsymbol{T}}(\boldsymbol{y}\mid\boldsymbol{t})$. Perustelu tälle: koska $\boldsymbol{T}(\boldsymbol{Y})$ määriteltiin tyhjentäväksi tunnusluvuksi, on ehdollinen yptf/ytf $f_{\boldsymbol{Y}|\boldsymbol{T}}(\boldsymbol{y}\mid\boldsymbol{t})$ määritelmän 2 nojalla parametrista $\theta$ riippumaton, jolloin yptf/ytf voidaan ajatella parametrin suhteen vakioksi. □

Lause tarkoittaa, että $L(\theta;\boldsymbol{y}) \propto L(\theta;\boldsymbol{t})$, joten ne ovat vakiota vaille samat, eli pelkän tyhjentävän tunnusluvun $\boldsymbol{t}$ avulla saadaan johdettua aineiston $\boldsymbol{y}$ uskottavuusfunktio vakiota vaille. Lauseesta saadaan johdettua seuraava tulos:

**Seuraus 1.** *Jos tunnusluku $T(\boldsymbol{Y})$ on tyhjentävä tunnusluku ja $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y})$, niin aineistojen $\boldsymbol{x}$ ja $\boldsymbol{y}$ uskottavuusfunktiot ovat verrannolliset. Toisin sanoen*

$$L(\theta;\boldsymbol{y}) = c(\boldsymbol{x},\boldsymbol{y})L(\theta;\boldsymbol{x}) \propto L(\theta;\boldsymbol{x})$$

Perustellaan vielä tämä seuraus. Otetaan lähtökohdaksi lauseen 2 tulos, ja tämän lisäksi vielä tarkasteltavaksi ehto $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y})$. Ehdosta seuraa, että

$$L(\theta;\boldsymbol{t}(\boldsymbol{x})) = L(\theta;\boldsymbol{t}(\boldsymbol{y})). \tag{5}$$

Yhtälöiden (4) ja (5) perusteella nähdään, että tällöin $L(\theta;\boldsymbol{y}) \propto L(\theta;\boldsymbol{x})$, eli aineistojen $\boldsymbol{x}$ ja $\boldsymbol{y}$ uskottavuusfunktiot ovat verrannolliset, jos tunnusluku $\boldsymbol{T}(\boldsymbol{Y})$ on tyhjentävä tunnusluku ja $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y})$. Tulosta tarvitaan myöhemmässä todistuksessa.

## 3.2 Tyhjentävyysperiaate

Ennen kuin määritellään tyhjentävyysperiaate, tarkennetaan mitä tutkielmassa tarkoitetaan sanalla *näyttö*. Sanaa käytetään tämän ja myöhemmin käsiteltävien periaatteiden määrittelemisessä.

Näyttö-sanan tilalla voisi myös käyttää sanaa *johtopäätös*. Tutkielmassa käytetään sanaa näyttö, koska englanninkielisessä kirjallisuudessa muun muassa tyhjentävyysperiaate sekä uskottavuusperiaatteen eri versiot ilmaistaan usein käyttäen näyttöä ja näyttöfunktiota (evidence function). *Johtopäätös* ja *näyttö* ei myöskään tarkoita aivan samaa asiaa; johtopäätöksiä tehdään saadun näytön perusteella, joten niiden tulkinta ei aina ole täysin sama. Toki jos näyttö on kahdessa tilanteessa sama, niin yleensä tulisi myös johtopäätösten olla samat.

Mainittu näyttöfunktio on tutkielman notaatioilla tilastollisen mallin $f_{\boldsymbol{Y}}$ funktio, eli $\mathrm{Ev}(f_{\boldsymbol{Y}},\boldsymbol{y})$, jossa $\boldsymbol{y}$ on havaittu aineisto ja Ev (evidence) merkintä itse näyttöfunktiolle. Määrittely kirjasta [3, luku 7.4]. Funktiolla merkitään siis mallin/kokeen perusteella saatua näyttöä parametrista $\theta$. Välttämättä näyttöfunktiota ei tarvitse määritellä eikä sille ole mitään yleistä määritelmää, jolloin melkein mikä tahansa järkevä muotoilu käy. Kuitenkin esimerkki tällaisesta funktiosta voisi olla itse uskottavuusfunktio. Vaihtoehtoisesti esimerkiksi normaalijakaumasta $\mathrm{N}(\mu, \sigma_0^2)$ tunnetulla varianssilla havaitusta riippumattomasta otoksesta $y_1, \ldots, y_n$ voitaisiin muodostaa näyttöfunktio $\mathrm{Ev}(f_{\boldsymbol{Y}}, \boldsymbol{y}) = (\bar{y}, \frac{\sigma_0}{\sqrt{n}})$ [4, esimerkki 6.3.4]. Funktio antaa näyttöä siitä, kuinka hyvä estimaattori otoskeskiarvo $\bar{y}$ on odotusarvolle $\mu$. Siirrytään seuraavaksi kuitenkin varsinaisesti määrittelemään tyhjentävyysperiaate käyttäen näitä merkintöjä.

Tyhjentävän tunnusluvun idea oli tiivistää kaikki oleellinen informaatio mallin parametrista, joten johdonmukaisesti päättely voidaan tehdä tällaisiin tunnuslukuihin perustuen. Tästä ajatuksesta voidaan johtaa seuraavanlainen yleinen päättelyn periaate:

**Määritelmä 3.** (Tyhjentävyysperiaate): Olkoon $\boldsymbol{T(Y)}$ tyhjentävä tunnusluku. Tilastollisessa mallissa $f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta)$ kaiken päättelyn parametrista $\theta$ tulisi pohjautua pelkästään tyhjentävään tunnuslukuun. Toisin sanoen, jos mallista $f$ havaituille aineistoille $\boldsymbol{x}$ ja $\boldsymbol{y}$ pätee tyhjentävän tunnusluvun $\boldsymbol{T}(\boldsymbol{Y})$ havaitulle arvolle ehto $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y})$, niin näytön parametrista $\theta$ tulee olla sama, eli

$$\mathrm{Ev}(f_{\boldsymbol{Y}}, \boldsymbol{x}) = \mathrm{Ev}(f_{\boldsymbol{Y}}, \boldsymbol{y}). \tag{6}$$

Tämä vaikuttaa ainakin pintapuolisesti järkevältä periaatteelta, sillä tyhjentävä tunnuslukuhan sisältää kaiken oleellisen informaation parametrista $\theta$. Näin ollen, vaikka aineistot olisivatkin erilaiset, jos tyhjentävän tunnusluvun arvo on sama, näytön parametrista $\theta$ tulee myös olla sama.

Yleisesti voidaan todeta, että malli vaikuttaa tyhjentävyyden käsitteeseen, joten mallidiagnostiikkaa ei välttämättä voida tehdä pelkästään tyhjentäviin tunnuslukuihin perustuen [2, luku 4.1]. Esimerkiksi residuaalien käyttö mallidiagnostiikassa rikkoo heti tyhjentävyysperiaatetta, koska ne eivät perustu tyhjentäviin tunnuslukuihin [4, s. 295]. Tästä huomataan, että vaikka periaate onkin hyödyllinen työkalu, ei se ole aina sopiva muiden tilastollisen päättelyn käytänteiden tai periaatteiden kanssa. Yleisesti kuitenkin voidaan sanoa, että frekventististen tilastotieteilijöiden keskuudessa tyhjentävyysperiaate vaikuttaa luontaiselta eikä periaatteen oikeellisuudesta ole käyty ainakaan runsasta keskustelua.

# 4 Uskottavuusperiaate

Tutkielmassa erityisesti tästä eteenpäin sanaa *tilastollinen malli* tai vain *malli* käytetään usein sanan *koe* (experiment) synonyymina. Englanninkielisessä kirjallisuudessa puhutaan tässä yhteydessä useimmiten kokeesta, siitä käytetään merkintää $E$ (experiment), ja se määritellään tutkielman notaatioilla seuraavanlaiseksi joukoksi: $\{\boldsymbol{Y}, \theta, f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta)\}$. Käytännössä kyse on siis samasta asiasta kuin mitä työssä tarkoitetaan, kun puhutaan tilastollisesta mallista. Koe-sanaa käytetään kuitenkin tutkielmassa yleensä keinona havainnollistaa, että usein mallien ollessa erilaiset on kyse myös erilaisista kokeista tai koeasetelmista.

Nyt kun on määritelty tyhjentävyys sekä siihen liittyvät käsitteet, päästään viimein itse uskottavuusperiaatteeseen sekä sen määrittelyyn. Käsittely perustuu suurimmalta osin kirjoihin [3, luvut 7.1 – 7.4] ja [4, luku 6.3]. Uskottavuusperiaatteita on yleisesti olemassa kaksi erilaista, niin sanottu *heikko uskottavuusperiaate* (weak likelihood principle) sekä *vahva uskottavuusperiaate* (strong likelihood principle). Tarkastellaan ensin heikkoa uskottavuusperiaatetta.

## 4.1 Heikko uskottavuusperiaate

Luvussa 3.1.2 osoitettiin uskottavuusfunktion olevan tyhjentävä tunnusluku, joten heikkoa uskottavuusperiaatetta voikin kutsua tyhjentävyysperiaatteen uskottavuuspohjaiseksi tulkinnaksi ja ne muistuttavatkin sisällöltään pitkälti toisiaan.

**Määritelmä 4.** (Heikko uskottavuusperiaate): Tilastollisessa mallissa $f_{\boldsymbol{Y}}(\boldsymbol{y};\theta)$ parametria $\theta$ koskevan päättelyn tulisi perustua pelkästään uskottavuusfunktioon. Siis, jos havaitut aineistot $\boldsymbol{x}$ ja $\boldsymbol{y}$ ovat sellaiset, että niitä vastaavat uskottavuusfunktiot ovat verrannolliset, eli

$$L(\theta;\boldsymbol{y}) = c(\boldsymbol{x},\boldsymbol{y})L(\theta;\boldsymbol{x}) \tag{7}$$

kaikilla $\theta \in \Omega$, niin molemmissa tilanteissa näytön parametrista $\theta$ tulee olla sama. Toisin sanoen tällöin

$$\mathrm{Ev}(f_{\boldsymbol{Y}},\boldsymbol{x}) = \mathrm{Ev}(f_{\boldsymbol{Y}},\boldsymbol{y}). \tag{8}$$

Heikon uskottavuusperiaatteen mukaan riittää, että eri aineistojen pohjalta lasketut uskottavuusfunktiot ovat verrannollisia. Perustelu tälle on yhtälön (2) esitys uskottavuusfunktioille; molempia aineistoja vastaavat uskottavuusfunktiot voidaan siis esittää ilman niissä esiintyviä parametrista riippumattomia vakioita, jolloin $L(\theta;\boldsymbol{x}) \propto L(\theta;\boldsymbol{y})$.

Tarkastellaan heikkoa uskottavuusperiaatetta seuraavaksi lyhyen esimerkin avulla [4, s. 291].

**Esimerkki 4.** Uskottavuusfunktion tehtävä yleisesti on arvioida erilaisten parametriarvojen uskottavuutta. Oletetaan, että verrataan kahta parametriarvoa, $\theta_1$ ja $\theta_2$, ja näille uskottavuusfunktion arvoille pätee $L(\theta_2;\boldsymbol{x}) = 2L(\theta_1;\boldsymbol{x})$. Voidaan sanoa, että funktion perusteella $\theta_2$ on kaksi kertaa niin uskottava kuin $\theta_1$. Jos tässä tilanteessa pätee myös yhtälö (7), niin $L(\theta_2;\boldsymbol{y}) = 2L(\theta_1;\boldsymbol{y})$. Myös tämän perusteella $\theta_2$ on kaksi kertaa niin uskottava kuin $\theta_1$. Täten, riippumatta aineistosta, molemmista tilanteista saatiin sama näyttö koskien parametria $\theta$.

## 4.2 Vahva uskottavuusperiaate

Uskottavuusperiaatteelle on olemassa myös vahvempi versio, joka käydään seuraavaksi läpi. Määritelmä on *In All Likelihood* -kirjasta [3, määritelmä 7.3].

**Määritelmä 5.** (Vahva uskottavuusperiaate): Olkoon $f_{\boldsymbol{Y}_1}(\boldsymbol{y}_1;\theta)$ ja $f_{\boldsymbol{Y}_2}(\boldsymbol{y}_2;\theta)$ kaksi erilaista tilastollista mallia, joilla on yhteisenä parametri $\theta$. Jos mallista $f_{\boldsymbol{Y}_1}$ on havaittu aineisto $\boldsymbol{y}_1$ ja mallista $f_{\boldsymbol{Y}_2}$ aineisto $\boldsymbol{y}_2$ ja niiden uskottavuusfunktiot ovat verrannolliset, eli

$$L_1(\theta;\boldsymbol{y}_1) = c(\boldsymbol{y}_1,\boldsymbol{y}_2)L_2(\theta;\boldsymbol{y}_2)$$

kaikilla $\theta \in \Omega$, niin näytön parametrista $\theta$ tulee olla molemmissa tilanteissa sama, eli

$$\mathrm{Ev}(f_{\boldsymbol{Y}_1},\boldsymbol{y}_1) = \mathrm{Ev}(f_{\boldsymbol{Y}_2},\boldsymbol{y}_2). \tag{9}$$

Merkittävimpänä erona tämän ja heikon uskottavuusperiaatteen välillä on siis niiden suhtautuminen tilastollisiin malleihin; heikossa uskottavuusperiaatteessa saattoi olla kaksi aineistoa, mutta niiden piti olla samasta mallista, kun taas vahvassa versiossa mallitkin voivat olla erilaiset. Vahva uskottavuusperiaate siis erityisesti asettaa suurimpaan rooliin päättelyssä itse uskottavuusfunktion, eli sen mukaan mallin rooli voidaan päättelyssä sivuuttaa. Vahva uskottavuusperiaate myös sivuuttaa käytetyn otantasuunnitelman (sampling scheme), tai sen mukaan otatantasuunitelmalla ei tulisi olla merkitystä, jos lopulta aineistojen pohjalta lasketut uskottavuusfunktiot ovat verrannolliset [3, luku 7.2]. Tarkastellaankin tätä seuraavaksi esimerkin avulla [3, esimerkki 7.1].

**Esimerkki 5.** Tarkastellaan yleistä binomikoe ja negatiivinen binomikoe -esimerkkiä, joka on tyypillisin tapa havainnollistaa tätä periaatetta. Kyseessä on siis toistokoe, esimerkiksi kolikonheitto. Ensimmäisessä tapauksessa on vastaava tilanne kuin aiemmin luvussa 2.2.2, eli mallina on binomimalli. Tällöin mallinnetaan onnistumisia, joiden lukumäärää vastaava satunnaismuuttuja on $K_1$, ja suoritettavia kokeita on kymmenen eli $n = 10$. Toisessa tapauksessa kyseessä on negatiivinen binomikoe, jossa kokeiden suorittaminen lopetetaan, kun havaitaan kaksi epäonnistumista. Negatiivisesta binomimallista käytetään siis versiota, jossa mallinnetaan onnistumisten lukumäärää, kunnes havaitaan kiinteä määrä epäonnistumisia. Tällöin pistetodennäköisyysfunktio on yleisesti muotoa $P(K = k) = \binom{k+r-1}{r-1}\theta^k(1-\theta)^r$. Kaavassa $k$ on onnistumisten lukumäärä ja $r$ epäonnistumisten. Negatiivisen binomikokeen onnistumisia vastaavasta satunnaismuuttujasta käytetään merkitään $K_2$ ja sen realisaatiosta merkintää $k_2$.

Binomikokeessa ennalta määritettynä on siis kokeiden lukumäärä $n$, kun taas negatiivisessa binomikokeessa epäonnistumisten lukumäärä $r$. Tällöin siis aineistonkeruumenetelmät ovat näissä tilanteissa erilaiset. Molemmissa tapauksissa parametri $\theta$ on kuitenkin sama, eli parametri määrää yksittäisen kokeen onnistumistodennäköisyyden. Täten, jos uskottavuusfunktiot ovat verrannolliset, näyttö parametrista on samaa vahvan uskottavuusperiaatteen mukaan.

Oletetaan, että binomikokeessa $n = 10$ kokeesta onnistumisia oli $k_1 = 8$. Negatiivisessa binomikokeessa taas tehtiin myös lopulta 10 koetta, eli ennen $r = 2$ epäonnistumista havaittiin $k_2 = 8$ onnistumista. Merkitään ensimmäistä tapausta ja mallia indeksillä 1 ja toista indeksillä 2. Tällöin vastaavat uskottavuusfunktiot ovat

$$L_1(\theta; k_1) = \binom{10}{8}\theta^8(1-\theta)^2 \propto \theta^8(1-\theta)^2 \tag{10}$$

ja

$$L_2(\theta; k_2) = \binom{8+2-1}{2-1}\theta^8(1-\theta)^2 \propto \theta^8(1-\theta)^2. \tag{11}$$

Kuten huomataan, lasketut uskottavuusfunktiot ovat verrannolliset. Täten vahvan uskottavuusperiaatteen nojalla näytön parametrista $\theta$ tulee olla sama molempien kokeiden osalta, vaikka koeasetelmat olivatkin erilaiset. Yleisesti vahva uskottavuusperiaate vaikuttaa tässä johdonmukaiselta, sillä molemmissa kokeissa on sama parametri ja havaittiin vielä sama määrä onnistumisia ja epäonnistumisia, joten ei ole ilmeistä syytä, miksei näyttö parametrista olisi sama.

Jatketaan vielä lyhyesti esimerkkiä tekemällä tilastollinen testi, jossa $H_0 : \theta = 0{,}5$ ja vastahypoteesi on yksisuuntainen $H_1 : \theta > 0{,}5$. Lasketaan molempien mallien osalta p-arvot. Negatiivisen binomimallin p-arvo $p_2$ perustellaan yksityiskohtaisemmin, sillä tulos ei ole täysin triviaali.

$$p_1 = P(K_1 \geq 8 \mid \theta = 0{,}5) = \sum_{k_1=8}^{10} \binom{10}{k_1} 0{,}5^{10} = 0{,}055$$

$$p_2 = P(K_2 \geq 8 \mid \theta = 0{,}5) = \sum_{k_2=8}^{\infty} (k_2+1) \left(\frac{1}{2}\right)^{k_2+2} = \frac{1}{4} \sum_{k_2=8}^{\infty} (k_2+1) \left(\frac{1}{2}\right)^{k_2}$$

Jos merkitään $n = k_2 - 1$, saadaan

$$\frac{1}{4} \sum_{n=9}^{\infty} n \cdot \left(\frac{1}{2}\right)^{n-1} = \frac{1}{4} \left( \sum_{n=1}^{\infty} n \cdot \left(\frac{1}{2}\right)^{n-1} - \sum_{n=1}^{8} n \cdot \left(\frac{1}{2}\right)^{n-1} \right).$$

Koska $\left|\frac{1}{2}\right| < 1$, voidaan käyttää geometristen sarjojen tulosta

$$\sum_{n=1}^{\infty} n \cdot \left(\frac{1}{2}\right)^{n-1} = \frac{1}{\left(1 - \frac{1}{2}\right)^2} = 4.$$

Nyt voidaan laskea sarjan lopullinen tulos, joka on

$$p_2 = \frac{1}{4} \left( 4 - \sum_{n=1}^{8} n \cdot \left(\frac{1}{2}\right)^{n-1} \right) = 0{,}020.$$

Huomataan, että $p_1 \neq p_2$, eli p-arvot eroavat toisistaan. Tässä voidaan havaita ristiriita; vaikka vahvan uskottavuusperiaatteen nojalla näytön parametrista tulee olla sama, tavallisen tilastollisen testaamisen menetelmin näyttö ei kuitenkaan ole aivan identtistä näissä tilanteissa. Käyttämällä p-arvoja päättelyssä "rikotaan" siis vahvaa uskottavuusperiaatetta. Kovin suurta eroa p-arvojen välillä ei ole, mutta jos asiaa tarkasteltaisiin päätösteoreettisella lähestymistavalla ja merkitsevyystasoksi olisi kiinnitetty $\alpha = 0{,}05$, niin $p_1 > \alpha$ ja $p_2 < \alpha$. Toisessa tapauksessa olisi siis hylätty nollahypoteesi ja toisessa se olisi jäänyt voimaan. Yleisesti ei kuitenkaan toimita ihan näin suoraviivaisesti. Tässä kuitenkin havaitaan yksi ristiriita tai ongelma vahvassa uskottavuusperiaatteessa, ja näiden tarkastelulla jatketaankin seuraavaksi.

## 4.3 Ongelmia uskottavuusperiaatteessa

Edellisestä esimerkistä huomattiin, että vaikka kahdella mallilla olisikin verrannolliset uskottavuusfunktiot, ei näyttö välttämättä ole niistä sama. Yleisesti ristiriitaisuus tilastollisessa päättelyssä frekventististen, eli toistetun aineistonkeruun ideaan perustuvien menetelmien kanssa on yksi merkittävimmistä ongelmista uskottavuusperiaatteessa. Ristiriita ilmenee muun muassa p-arvojen sekä luottamusvälien kanssa. Uskottavuusperiaatteen kanssa yhteensopimattomia ovatkin usein havaitun aineiston ulkopuoliset näkökannat, joihin toistettu aineistonkeruu perustuu. Tällaiset menetelmät ovatkin tärkeitä päättelyn kannalta, sillä niiden avulla pystytään arvioimaan päätelmien epävarmuutta. Tästä syystä tämä ristiriitaisuus on merkittävä ongelma. Usein tilastotieteilijät suhtautuvatkin uskottavuusperiaatteeseen varauksella eivätkä kaikki periaatetta hyväksy. Seuraavaksi esitellään tiettyjä ongelmia, joita uskottavuusperiaatteessa on. Ongelmat ovat kirjasta [3, luku 7.1].

- Peräkkäisotannalla on mahdollista luoda harhainen aineisto. Uskottavuusfunktio ei kuitenkaan millään tavalla ota huomioon käytettyä otantamenetelmää, jolloin voidaan saada satunnaiskokeista näennäisiä (spurious) tuloksia, joita ei saada toistettua. Tästä aiheesta tarkemmin [3, luku 7.5].
- Uskottavuusfunktion laskeminen vaatii todennäköisyysmallin, minkä voi nähdä uskottavuusfunktion määritelmästä. Vaikka uskottavuusfunktio sisältää kaiken tiedon mallin parametrista, mallin sopivuudesta ei voida välttämättä olla täysin varmoja. Tästä seuraa ongelma, sillä uskottavuusfunktio ei voi ilmaista epävarmuutta mallista, johon se perustuu. Mallin sopivuus kuitenkin vaikuttaa vahvasti siihen, ovatko tehdyt johtopäätökset oikeita vai ei. Uskottavuusperiaatteessa myös oletetaan mallin oikeellisuus, eikä sitä voida soveltaa, jos malli ei sovikaan.
- Yksi ongelma liittyy moniparametrisiin malleihin. Ilman frekventistisen näkökulman ottamista huomioon uskottavuuspohjainen päättely mallin kaikista parametreista voi johtaa huonosti toistettaviin otantamenetelmiin [3, luku 3.5]. Yleisesti ottaen, jos parametriavaruus on liian suuri suhteessa aineistoon, uskottavuusfunktio voi tuottaa näennäisiä tuloksia. Tällaisessa tilanteessa havaittu aineisto pysyy kiinteänä, mutta parametriavaruutta laajennetaan, jotta löydetään parhaiten aineistoa selittävä malli. Lopputuloksena saadaan malli, joka on ylisovittunut aineistoon. Ongelma tässä on siis se, ettei uskottavuusfunktion avulla ole mahdollista saada tietää tästä ongelmasta [3, luku 7.7].

Listaus ei ole tyhjentävä, mutta erään kuvan periaatteen ongelmista tästä kuitenkin saa. Siitä huolimatta, että periaatteessa on ongelmia, se on silti merkittävä periaate tilastollisessa päättelyssä. Myöhemmin käsitelläänkin periaatteeseen liittyvää tärkeää tulosta eli Birnbaumin lausetta. Sitä ennen määritellään kuitenkin *ehdollistamisperiaate.*

# 5 Ehdollistamisperiaate

Ehdollistamisperiaate on tärkeä yleinen tilastollisen päättelyn periaate, joka on olennainen käsite Birnbaumin lauseessa. Jotta päästään aiheen ytimeen, tarkastellaan sitä ensin esimerkin avulla. Esimerkki on pitkälti sama kuin kirjassa [3, esimerkki 7.2].

## 5.1 Johdatteleva esimerkki

**Esimerkki 6.** Olkoon kiinnostuksen kohteena satunnaismuuttuja $X$, joka on tietyn aineen pitoisuus näytteessä. Satunnaismuuttujan odotusarvo $\mu$ kuvaa aineen todellista pitoisuutta, kun taas mitattu pitoisuus $X$ sisältää myös mittausvirheen. Näyte voidaan lähettää joko laboratorioon A tai B. Laboratorioiden mittauksille pätee $X_A \sim \mathrm{N}(\mu, 1)$ ja $X_B \sim \mathrm{N}(\mu, 4)$. Huomataan, että toisella näistä on suurempi varianssi, jolloin siellä mittausvirhe on suurempi. Päätös laboratorioiden välillä tehdään heittäen kolikkoa, ja sanotaan, että sattumalta valikoituu laboratorio B. Kysymys kuuluu, pitääkö johtopäätöksiä tehdessä ottaa huomioon, että valinta laboratorioiden välillä tehtiin satunnaisesti? Vai onko merkitystä ainoastaan sillä, kumpi kahdesta vaihtoehdosta tuli lopulta valituksi? Palataan tähän lopuksi.

Havaitaan arvo $x = 3$, siis aineen mitattu pitoisuus on 3. Tällä havainnolla halutaan testata hypoteesia $H_0 : \mu = 0$. Vastahypoteesi on yksisuuntainen $H_1 : \mu > 0$. Jos ei tiedetä kummasta laboratoriosta mittaustulos on, mitatun pitoisuuden satunnaismuuttujalle $X$ pätee

$$X \sim \frac{1}{2}\mathrm{N}(\mu, 1) + \frac{1}{2}\mathrm{N}(\mu, 4),$$

jolloin tästä laskettu p-arvo olisi

$$p_1 = \frac{1}{2}P\left(Z > 3\right) + \frac{1}{2}P\left(Z > \frac{3}{2}\right) = 0{,}034 < 0{,}05,$$

jossa $Z \sim \mathrm{N}(0, 1)$. Jos taas otetaan p-arvon laskemisessa huomioon, että valittu laboratorio lopulta oli B, saadaan p-arvoksi

$$p_2 = P\left(Z > \frac{3}{2}\right) = 0{,}067 > 0{,}05.$$

Ensimmäisen p-arvon $p_1$ voi sanoa olevan *ehdollistamaton* ja toisen p-arvon eli $p_2$ taas olevan *ehdollistettu*, eli $p_2$ ehdollistettiin kolikonheiton tuloksen suhteen.

Kumpaa näistä p-arvoista, ehdollistettua vai ehdollistamatonta, tulee käyttää, eli tuleeko ottaa huomioon, miten laboratorion valinta tehtiin vai huomioida vain valituksi tullut laboratorio? Yleisesti ehdollistettu vaikuttaa tässä tilanteessa järkevämmältä vaihtoehdolta; tilanteessa oli kuitenkin tiedossa kummassa laboratoriossa näyte oltiin analysoitu, joten on johdonmukaista, että tietoa käytetään. Sillä ei myöskään tulisi olla merkitystä, että laboratorio A oli myös mahdollinen vaihtoehto, jos lopulta kuitenkin valinta osui laboratorioon B. Tästä ideasta päästään ehdollistamisperiaatteen ytimeen.

## 5.2 Periaatteen määritelmä

Ennen itse määritelmää täytyy määritellä siinä käytetty termi *sekoitekoe* (mixture experiment). Ilmaus tarkoittaa esimerkin 6 kaltaista tilannetta, jossa on kaksi erilaista koetta tai mallia, joista valitaan yksi satunnaisesti. Valinta voidaan tehdä esimerkiksi heittämällä kolikkoa tai vaikka satunnaislukugeneraattorilla. Valinnan jälkeen suoritetaan mallia vastaava koe. Määritellään formaalimmin vielä tämä käsite perustuen kirjaan [3, luku 7.4].

Oletetaan, että on kaksi mallia, $f_{\boldsymbol{Y}_1}(\boldsymbol{y}_1; \theta)$ ja $f_{\boldsymbol{Y}_2}(\boldsymbol{y}_2; \theta)$, joilla on sama parametri. Valinta mallien välillä tehdään satunnaismuuttujan $J$ realisaation eli arvon $j$ perusteella, jolloin valitaan malli $f_{\boldsymbol{Y}_j}$. Tätä merkintää käytetään silloin, kun sekoitekokeessa on saatu $J$:n realisoitunut arvo $j$. Satunnaismuuttujalle $J$ pätee $P(J = 1) = P(J = 2) = \frac{1}{2}$. Tilanteessa, jossa ei ole vielä valittu mallia, eli itse sekoitekokeessa, käytetään mallista merkintää $f^* = f_{\boldsymbol{Y}^*}(\boldsymbol{y}^*; \theta)$, jossa $\boldsymbol{Y}^* = (J, \boldsymbol{Y}_J)$ ja $f_{\boldsymbol{Y}^*}(j, \boldsymbol{y}_j; \theta) = \frac{1}{2} f_{\boldsymbol{Y}j}(\boldsymbol{y}_j; \theta)$.

**Määritelmä 6.** (Ehdollistamisperiaate): Olkoot $f_{\mathbf{Y}_1}(\boldsymbol{y}_1; \theta)$ ja $f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta)$ kaksi tilastollista mallia, joilla on yhteisenä parametri $\theta$. Jos yhdessä tapauksessa tehdään sekoitekoe $f^*$, jossa valitaan satunnaisesti malli $f_{\mathbf{Y}_j}$, ja toisessa tapauksessa valitaan suoraan malli $f_{\mathbf{Y}_j}$ ilman satunnaisuutta, näytön tulee olla molemmissa tilanteissa sama. Näyttö riippuu siis ainoastaan valitusta mallista, eli

$$\mathrm{Ev}(f^*, \boldsymbol{y}^*) = \mathrm{Ev}(f_{\mathbf{Y}_j}, \boldsymbol{y}_j). \tag{12}$$

Jos tätä periaatetta sovellettaisiin aiempaan esimerkkiin 6, p-arvon laskemisessa otetaan huomioon pelkästään valittu laboratorio eikä valinnan satunnaisuudella ole merkitystä eikä sitä oteta huomioon. Täten käytetään päättelyssä ehdollistettua p-arvoa.

Ehdollistamisperiaatetta voidaan myös käyttää esimerkin 5 kaltaisessa tilanteessa. Sanotaan, että valitaan satunnaisesti esimerkin kahdesta mallista, eli binomimallista ja negatiivisesta binomimallista. Jos sitten satunnaisesti valitaan ja suoritetaan esimerkiksi binomallia vastaava koe ja saadaan aineisto $k_1 = 8$, näyttö parametrista $\theta$ riippuu ainoastaan valitusta mallista. Sama näyttö olisi siis saatu, jos binomimallia vastaava koe olisi ei-satunnaisesti päätetty tehdä ja havaittu sama aineisto $k_1 = 8$.

Ehdollistamisperiaate tarkoittaa yleisesti sitä, että *ehdollistetaan* valitun mallin suhteen, eli sekoitekokeen satunnaismuuttujan $J$ havaitun arvon mukaan. Periaatteelle onkin olemassa myös yleisempi esitystapa. Tutkielmassa määriteltyä ehdollistamisperiaatetta kutsutaan usein *heikoksi* ehdollistamisperiaatteeksi, mutta on myös *vahva* ehdollistamisperiaate, jota kutsutaan lisäksi *aputunnuslukuperiaatteeksi* (ancillarity principle). Aputunnusluku on satunnaismuuttuja, jonka jakauma ei riipu parametrista $\theta$. Esimerkki tällaisesta on sekoitekokeen satunnaismuuttuja $J$. Aputunnuslukuperiaatteen mukaan ehdollistetaan aputunnusluvun havaitun arvon suhteen eli tulee käyttää ehdollista jakaumaa. Tämän työn ehdollistamisperiaate voidaankin nähdä aputunnuslukuperiaatteen erikoistapauksena. Tutkielmassa kuitenkin tarkoitetaan ehdollistamisperiaatteella nimenomaan määritelmän 6 periaatetta. [5, luku 1]

Yleisesti myös tämä periaate, kuten tyhjentävyysperiaatekin, vaikuttaa ainakin pintapuolisesti järkevältä. Johdonmukaisestihan sekoitekokeessa ainoastaan valitulla mallilla tulisi olla merkitystä, eikä muilla malleilla mitä olisi voinut tulla valituksi, mutta ei kuitenkaan tullut. Kuten todetaan kirjassa [4, luku 6.3.2], se, että tietty malli tuli valituksi jonkun toiseen sijaan, ei ole kasvattanut, vähentänyt tai muuttanut tietoa parametrista $\theta$. Ehdollistamisperiaate onkin tyhjentävyysperiaatteen ohella periaate, jota frekventistiset tilastotieteilijät pitävät varsin luonnollisena.

Nyt kun ehdollistamisperiaate on määritelty, voidaan siirtyä Birnbaumin lauseen käsittelyyn.

# 6 Birnbaumin lause

Luvussa esitellään Birnbaumin lause, joka yhdistää edellä esitellyt kaksi merkittävää tilastollisen päättelyn periaatetta.

## 6.1 Birnbaumin lause ja sen todistus

Tutkielmassa on määritelty aiemmin kaksi periaatetta, tyhjentävyysperiaate sekä ehdollistamisperiaate. Molemmat niistä ovat olleet ainakin näennäisesti järkeviä sekä johdonmukaisia. Tutkielmassa on myös esitetty uskottavuusperiaate, ja tarkemmin sen heikko ja vahva muoto. Seuraava lause ja sen todistus näyttää yhteyden vahvan uskottavuusperiaatteen sekä tyhjentävyys- ja ehdollistamisperiaatteiden välillä.

**Lause 3.** *(Birnbaumin lause): Tyhjentävyysperiaate ja ehdollistamisperiaate ovat yhdessä yhtäpitävät vahvan uskottavuusperiaatteen kanssa.*

Toisin sanoin lause voidaan ilmaista seuraavasti: vahva uskottavuusperiaate on voimassa, jos ja vain jos sekä tyhjentävyysperiaate että ehdollistamisperiaate ovat voimassa. Täten tulos pätee kumpaankin suuntaan. Tulos on merkittävä, sillä sen mukaan jos tilastotieteilijä hyväksyy tyhjentävyys- sekä ehdollistamisperiaatteen, täytyy hänen myös hyväksyä vahva uskottavuusperiaate ja toisinpäin.

Todistetaan lause seuraavaksi huolellisesti. Todistus mukailee lähteen [3] s. 198 alkavaa todistusta Birnbaumin lauseelle, mutta se on samankaltainen kuin alkuperäinen Birnbaumin [8] oma todistus vuodelta 1962.

*Todistus.* Aloitetaan osoittamalla väitteen ensimmäinen suunta, eli osoitetaan ensin tyhjentävyys- ja ehdollistamisperiaatteen yhdessä implikoivan vahvan uskottavuusperiaatteen. Aloitetaan vahvan uskottavuusperiaatteen lähtökohdasta. Olkoot $\boldsymbol{y}_1$ ja $\boldsymbol{y}_2$ aineistoja, joilla on verrannolliset uskottavuusfunktiot $L_1(\theta; \boldsymbol{y}_1) = c \cdot L_2(\theta; \boldsymbol{y}_2)$. Toisin sanoen vastaaville yptf/ytf:ille pätee $f_{\mathbf{Y}_1}(\boldsymbol{y}_1; \theta) = c \cdot f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta)$. Nyt tarkastellaan sekoitekoetta luvun 5 määrittelemällä tavalla, siis mallia $f^*$. Kun havaitaan mikä tahansa $\boldsymbol{y}^* = (j, \boldsymbol{y}_j)$, niin ehdollistamisperiaatteen nojalla voidaan todeta, että

$$\mathrm{Ev}(f^*, \boldsymbol{y}^*) = \mathrm{Ev}(f_{\mathbf{Y}_j}, \boldsymbol{y}_j). \tag{13}$$

Määritellään sitten tunnusluku $\boldsymbol{T}(\boldsymbol{Y}^*)$ kaavalla

$$\boldsymbol{T}(J, \boldsymbol{Y}_J) = \begin{cases} (1, \boldsymbol{y}_1), & \text{jos } J = 2, \boldsymbol{Y}_2 = \boldsymbol{y}_2, \\ (J, \boldsymbol{Y}_J), & \text{muulloin.} \end{cases}$$

Määrittelyn perusteella, jos havaitaan $(J, \boldsymbol{Y}_J) = (2, \boldsymbol{y}_2)$, tämä muutetaan muotoon $(1, \boldsymbol{y}_1)$. Täten voidaan todeta, että jos havaitaan $\boldsymbol{T} = \boldsymbol{t} = (1, \boldsymbol{y}_1)$, ei ole tietoa siitä, kumpi malli oli alun perin valittu, eli havaittiinko $(1, \boldsymbol{y}_1)$ vai $(2, \boldsymbol{y}_2)$. Tietoa siis tiivistettiin. Todistuksessa ideana on osoittaa, että tällä tiivistämisellä ei ole merkitystä. Toisin sanoen osoitetaan, että $\boldsymbol{T}(J, \boldsymbol{Y}_J)$ on tyhjentävä tunnusluku. Tätä varten tutkitaan ehdollista jakaumaa $\boldsymbol{Y}^* \mid \boldsymbol{T} = \boldsymbol{t}$, ja siihen liittyviä todennäköisyyksiä. Näiden osoitetaan olevan riippumattomia parametrista $\theta$. Käsitellään erikseen tapaukset, joissa $\boldsymbol{t} \neq (1, \boldsymbol{y_1})$ ja $\boldsymbol{t} = (1, \boldsymbol{y_1})$. Nämä kattavat kaikki mahdolliset tapaukset. Ensimmäinen todennäköisyyslauseke tapauksessa $\boldsymbol{t} \neq (1, \boldsymbol{y_1})$ on

$$P\{\boldsymbol{Y}^* = (j, \boldsymbol{y}_j) \mid \boldsymbol{T} = \boldsymbol{t}\} = \begin{cases} 1, & \text{jos } (j, \boldsymbol{y_j}) = \boldsymbol{t}, \\ 0, & \text{muulloin.} \end{cases} \tag{14}$$

Perustellaan yhtälön tulos huolellisesti. Tässä tarvitaan ehdollisen todennäköisyyden kaavaa, joka on $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$. Tässä tapauksessa $A = \{\boldsymbol{Y}^* = (j, \boldsymbol{y}_j)\}$ ja $B = \{\boldsymbol{T} \neq (1, \boldsymbol{y_1})\}$, jolloin $A \cap B$ on $\{\boldsymbol{Y}^* = (j, \boldsymbol{y}_j), \boldsymbol{T} \neq (1, \boldsymbol{y_1})\}$. Tunnusluvulle $\boldsymbol{t}$ on kaksi mahdollista tapausta: joko $\boldsymbol{t} = (j, \boldsymbol{y}_j)$ tai $\boldsymbol{t} \neq (j, \boldsymbol{y}_j)$, joista jälkimmäinen on yhtälön (14) tapaus ”muulloin”. Samaan aikaan, koska $\boldsymbol{t} \neq (1, \boldsymbol{y_1})$, tunnusluvun $\boldsymbol{t}$ arvo määräytyy suoraan $\boldsymbol{Y}^*$:n perusteella. Siis jos $\boldsymbol{t} = (j, \boldsymbol{y}_j)$, niin $\boldsymbol{Y}^* = (j, \boldsymbol{y_j}) = \boldsymbol{t}$ ja leikkauksen $A \cap B$ tapahtumat ovat sama tapahtuma. Tällöin ehdollinen todennäköisyys on siis

$$\frac{P(\boldsymbol{Y}^* = (j, \boldsymbol{y}_j))}{P(\boldsymbol{Y}^* = (j, \boldsymbol{y}_j))},$$

joka saa arvoksi 1. Toisessa tapauksessa, jolloin $\boldsymbol{t} \neq (j, \boldsymbol{y}_j)$, leikkaus $A \cap B$ on tyhjä joukko, koska tapahtumat ovat toisensa poissulkevia. Tästä syystä todennäköisyys on siinä tapauksessa 0.

Siirrytään sitten toiseen ehdollisen todennäköisyyden tapaukseen ehdolla $\boldsymbol{t} = (1, \boldsymbol{y}_1)$. Todennäköisyyslauseke tällöin on

$$\begin{aligned} P\{\boldsymbol{Y}^* = (1, \boldsymbol{y}_1) \mid \boldsymbol{T} = \boldsymbol{t}\} &= \frac{P(\boldsymbol{Y}^* = (1, \boldsymbol{y}_1))}{P(\boldsymbol{Y}^* = (1, \boldsymbol{y}_1)) + P(\boldsymbol{Y}^* = (2, \boldsymbol{y}_2))} \\ &= \frac{\frac{1}{2} f_{\mathbf{Y}_1}(\boldsymbol{y}_1; \theta)}{\frac{1}{2} f_{\mathbf{Y}_1}(\boldsymbol{y}_1; \theta) + \frac{1}{2} f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta)} \\ &= \frac{\frac{1}{2} c \cdot f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta)}{\frac{1}{2} c \cdot f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta) + \frac{1}{2} f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta)} \\ &= \frac{c}{c+1}. \end{aligned} \tag{15}$$

Tässä käytettiin yhtälöä $f_{\mathbf{Y}_1}(\boldsymbol{y}_1; \theta) = c \cdot f_{\mathbf{Y}_2}(\boldsymbol{y}_2; \theta)$, sekoitekokeen määritelmää, jonka perusteella $P(\boldsymbol{Y}^* = (j, \boldsymbol{y}_j)) = \frac{1}{2} f_{\mathbf{Y}_j}(\boldsymbol{y}_j; \theta)$ ja aiempaa ehdollisen todennäköisyyden kaavaa, jossa nyt $A = \{\boldsymbol{Y}^* = (1, \boldsymbol{y}_1)\}$ ja $B = \{\boldsymbol{T} = (1, \boldsymbol{y}_1)\}$. Tapahtuma $B$ on näiden kahden tapahtuman unioni: $\boldsymbol{Y}^* = (1, \boldsymbol{y}_1)$ ja $\boldsymbol{Y}^* = (2, \boldsymbol{y}_2)$. Leikkauksessa $A \cap B$ ainoa sekä $A$:lle että $B$:lle yhteinen tapahtuma on $\boldsymbol{Y}^* = (1, \boldsymbol{y}_1)$. Tästä syystä osoittajassa on pelkästään tämä lauseke.

Yhtälöissä (14) ja (15) on käsitelty ehdollisen jakauman $\boldsymbol{Y^*} \mid \boldsymbol{T} = \boldsymbol{t}$ kaikki mahdolliset tapaukset, ja näissä yhtälöissä olevat lopulliset lausekkeet eivät enää sisällä parametria $\theta$, joten ne ovat siitä riippumattomia. Täten määritelmän 2 nojalla $\boldsymbol{T}$ on tyhjentävä tunnusluku. Tällöin soveltamalla tyhjentävyysperiaatetta voidaan todeta, että

$$\mathrm{Ev}\{f^*, (1, \boldsymbol{y}_1)\} = \mathrm{Ev}\{f^*, (2, \boldsymbol{y}_2)\}. \tag{16}$$

Yhtälöiden (13) ja (16) perusteella taas voidaan todeta, että $\mathrm{Ev}(f_{\mathbf{Y}_1}, \boldsymbol{y}_1) = \mathrm{Ev}(f_{\mathbf{Y}_2}, \boldsymbol{y}_2)$, mikä on vahva uskottavuusperiaate, määritelmän 3 nojalla.

Implikaatio vasemmalta oikealle on täten osoitettu. Siirrytään sitten toiseen vaiheeseen ja osoitetaan päinvastainen. Lähtökohdaksi otetaan siis vahva uskottavuusperiaate, ja osoitetaan, että siitä seuraa ehdollistamis- ja tyhjentävyysperiaate.

Aloitetaan parametrin $\theta$ uskottavuusfunktiosta perustuen havaintoon $\boldsymbol{y}^* = (j, \boldsymbol{y_j})$ mallista $f^*$, joka siis on

$$L^*(\theta; \boldsymbol{y}^*) = \frac{1}{2} f_{\boldsymbol{Y}_j}(\theta; \boldsymbol{y}_j).$$

Funktion huomataan olevan verrannollinen havaintoon $\boldsymbol{y}_j$ perustuvaan uskottavuusfunktioon mallissa $f_j$. Koska ne ovat verrannollisia, vahvan uskottavuusperiaatteen nojalla voidaan todeta, että

$$\mathrm{Ev}(f^*, \boldsymbol{y}^*) = \mathrm{Ev}(f_{\boldsymbol{Y}_j}, \boldsymbol{y}_j).$$

Tämä taas on määritelmän 6 nojalla sama kuin ehdollistamisperiaate. Luvun 3.1 seurauksessa 1 osoitettiin, että jos tunnusluku $\boldsymbol{T}(\boldsymbol{Y})$ on tyhjentävä, ja $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{y})$, niin aineistojen $\boldsymbol{x}$ ja $\boldsymbol{y}$ uskottavusfunktiot ovat verrannolliset. Täten, koska tunnusluku $\boldsymbol{T}(J, \boldsymbol{Y}_J)$ on tyhjentävä, vahvan uskottavuusperiaatteen mukaan

$$\mathrm{Ev}(f_{\boldsymbol{Y}}, \boldsymbol{x}) = \mathrm{Ev}(f_{\boldsymbol{Y}}, \boldsymbol{y}),$$

joka taas on tyhjentävyysperiaate määritelmän 3 nojalla. Ollaan siis osoitettu vahvasta uskottavuusperiaatteesta seuraavan tyhjentävyys- sekä ehdollistamisperiaate. Väitteen molemmat suunnat on täten osoitettu, joten lause on todistettu. $\square$

Birnbaumin lause on tilastollisen päättelyn saralla merkittävä tulos, mutta lausetta on myös kritisoitu julkaisustaan lähtien [7]. Seuraavaksi tarkastellaan osaa ajan saatossa lausetta vastaan esitetystä kritiikistä.

## 6.2 Kritiikkiä Birnbaumin lauseesta

Vaikka tässä alaluvussa käsitelläänkin tarkemmin Birnbaumin lauseeseen kohdistuvaa kritiikkiä, on oleellista huomioida myös vahvan uskottavuusperiaatteen kiistanalaisuus, jota on käsitelty aiemmin tutkielmassa erityisesti luvussa 4.3. Yleisesti, jos joudutaan esimerkin 5 kaltaiseen tilanteeseen, jossa ”rikotaan” vahvaa uskottavuusperiaatetta, ”rikotaan” Birnbaumin lauseen mukaan joko tyhjentävyys- tai ehdollistamisperiaatetta. Siis Birnbaumin lauseen mukaan, jos hylätään vahva uskottavuusperiaate, täytyy myös hylätä tyhjentävyysperiaate, ehdollistamisperiaate tai mahdollisesti molemmat. Tätä seurausta ei kuitenkaan yleensä olla valmiita hyväksymään, sillä nämä kaksi periaatetta eivät ole kovinkaan kiistanalaisia. Yksi tärkeimmistä kiistanaiheista Birnbaumin lauseessa siis on, että tyhjentävyys- ja ehdollistamisperiaate vaikuttavat monille varsin luontaisilta periaatteilta, kun taas vahva uskottavuusperiaate on verrattain kiistelty [6, luku 1]. Seuraavaksi tarkastellaan kahta artikkelia, joissa pohjimmiltaan kritisoidaan Birnbaumin lausetta nimenomaan tästä syystä.

Kaksi tarkemmin käsiteltävää julkaisua ovat Evansin (2013) [6] ja Mayon (2014) [7] artikkelit. Ne ovat keskeisiä Birnbaumin lausetta kritisoivia julkaisuja, joiden sisällöstä on käyty aikanaan paljon keskustelua suuntaan ja toiseen. Tähän keskusteluun on osallistunut muun muassa Berger ja Penã (2017) [5] omalla artikkelillaan, jossa arvioidaan Evansin ja Mayon artikkelien sisältöä. Tutkielmassa näiden analyysi perustuu pääosin Bergerin ja Penãn artikkeliin, mutta erityisesti Evansin osalta myös hänen alkuperäiseen artikkeliinsa.

Seuraavaksi käydään läpi yksi kerrallaan, mitä kumpikin näistä alkuperäisistä artikkeleista todistuksessa kritisoi, ja miten tähän kritiikkiin voi mahdollisesti vastata Bergerin ja Penãn mukaan. Vastausta kritiikkiin tai itse kritiikkiä ei tarkastella yhtä perusteellisesti kuin alkuperäisissä artikkeleissa. Tutkielmassa tässä osuudessa käytetyt merkinnät ovat kuitenkin pitkälti samoja kuin Bergerin ja Penãn artikkelissa, eivätkä siis välttämättä täysin samoja kuin alkuperäisissä Evansin tai Mayon julkaisuissa.

Yleisimmin käytetty merkintä on $(f_{\boldsymbol{Y}}, \boldsymbol{y})$, jota Berger ja Penã kutsuu päättelyn perustaksi (inference base), lyhyemmin perusta. Tämä on käytännössä tapa merkitä tilastollinen malli ja siitä havaittu aineisto. Perustan $(f, \boldsymbol{y})$ avulla tehdään siis päätelmiä parametrista $\theta$. Mallista käytetään seuraavassa osiossa lyhyempää merkintää $f$, mutta tuossa muodossa sillä tarkoitetaan käytännössä notaatiota $f_{\boldsymbol{Y}}$. Lyhyempi merkintätapa selkeyttää artikkelin käsittelyä. Tämän lisäksi merkinnällä $\mathrm{Ev}(f, \boldsymbol{y})$ tarkoitetaan Bergerin ja Penãn tavoin päätelmää tai johtopäätöstä, joka tehdään mallin ja aineiston perusteella. Merkitys on siis käytännössä sama kuin aiemmin tutkielmassa.

### 6.2.1 Evansin artikkeli

Evans ei määritelmissään käytä näyttöfunktiota, jota Birnbaumin todistuksessa taas käytettiin, vaan kaikki vaadittavat periaatteet on määritelty käyttäen joukkoja ja niiden välisiä relaatioita. Evansin mielestä usein ongelmat Birnbaumin lauseessa liittyvät siihen, miten Birnbaum määrittelee tyhjentävyys- ja ehdollistamisperiaatteen. Tästä syystä seuraavaksi tarkastellaan, miten Evans on itse määritellyt Birnbaumin lauseen periaatteet. Seuraavissa merkinnöissä $S$ tarkoittaa tyhjentävyysperiaatetta, $L$ vahvaa uskottavuusperiaatetta, ja $C$ ehdollistamisperiaatetta.

- $S$: $(f, \boldsymbol{y}) \sim_S (f, \boldsymbol{y}')$ jos ja vain jos on olemassa tyhjentävä tunnusluku $\boldsymbol{T}$ parametrille $\theta$, jolle $\boldsymbol{T}(\boldsymbol{y}) = \boldsymbol{T}(\boldsymbol{y}')$,
- $L$: $(f, \boldsymbol{y}) \sim_L (f', \boldsymbol{y}')$ jos ja vain jos $f_{\boldsymbol{Y}}(\boldsymbol{y}; \theta) = c \cdot f_{\boldsymbol{Y}'}(\boldsymbol{y}'; \theta)$ vakiolle $c > 0$, joka ei riipu parametrista $\theta$.
- $C$: $(f, \boldsymbol{y}) \sim_C (f', \boldsymbol{y}')$ jos ja vain jos perusta $(f, \boldsymbol{y})$ on sekoitekokeen tilanne ja $(f', \boldsymbol{y}')$ on sekoitekokeessa valittu perusta (malli). Toisin sanoen sekoitekokeen merkinnöin siis $f = f^*$, $\boldsymbol{y} = (j, \boldsymbol{y}_j)$, $f' = f_j$ ja $\boldsymbol{y}' = \boldsymbol{y}_j$.

Huomataan, että muodoltaan nämä muistuttavat vahvasti aiempia määrittelyjä näille periaatteille, mutta tässä ne tehdään joukkojen (parien) relaatioina. Tarkastellaan seuraavaksi, mihin lopputulokseen Evans päätyy Birnbaumin todistuksesta käyttämällä näitä määritelmiä, perustuen lähteeseen [6, luvut 2–4].

Evansin mukaan Birnbaumin todistus ei osoita, että tyhjentävyysperiaatteen $S$ ja ehdollistamisperiaatten $C$ relaatiot johtavat vahvan uskottavuusperiaatteen $L$ määrittelemään relaatioon. Sen sijaan se osoittaa kahden perustan, joilla on verrannolliset uskottavuusfunktiot, olevan päättelyn kannalta keskenään yhtäpitäviä, jos sekä ehdollistamisperiaatetta $C$ että tyhjentävyysperiaatetta $S$ sovelletaan toistuvasti, mutta ne eivät ole yhtäpitäviä pelkästään periaatteiden $C$ tai $S$ perusteella. Tarkemmin $L$ ei ole sama kuin $S \cup C$, vaan se on pienin ekvivalenssirelaatio, jonka generoi $S \cup C$. Näiden periaatteiden soveltaminen peräkkäin käytännössä tarkoittaa, että perustaan $(f, \boldsymbol{y})$ sovelletaan toistuvasti periaatteita $C$ ja $S$, ja siten osoitetaan sen olevan ekvivalentti perustan $(f', \boldsymbol{y}')$ kanssa. Tämän ketjuttamisen lopputulos on pienin unionin $S \cup C$ generoima ekvivalenssirelaatio.

Sanallisesti Evansin mukaan tämän tuloksen voisi ilmaista näin: jos hyväksytään relaatio $S$ ja hyväksytään relaatio $C$ *sekä* hyväksytään kaikki periaatteiden $S$ ja $C$ yhdessä generoimat ekvivalenssit, niin tällöin tulee hyväksyä myös periaate $L$. Viimeinen on Evansin mukaan oleellinen lisäoletus, jota ei mainita Birnbaumin lauseessa. Evansin mielestä näistä ongelmista päästään siis eroon hyväksymällä, että lisäoletuksia vaaditaan, jotta Birnbaumin tulos pätee. Kuitenkin hänestä tuloksen merkitys vähenee, jos nämä oletukset esitetään oikein.

Seuraavaksi tarkastellaan, miten Berger ja Penã [5, luvut 2 ja 4] ovat vastanneet Evansin kritiikkiin. Vaikkei Evans näyttöfunktiota määritellytkään, se voidaan Bergerin ja Penãn mukaan tässäkin tilanteessa määritellä seuraavalla tavalla: $\mathrm{Ev}(f, \boldsymbol{y}) = \mathrm{Ev}(f', \boldsymbol{y}')$ jos ja vain jos $(f, \boldsymbol{y}) \sim_C (f', \boldsymbol{y}')$ tai $(f, \boldsymbol{y}) \sim_S (f', \boldsymbol{y}')$. Käyttämällä tällaista määritelmää näyttöfunktion Ev generoima ekvivalenssirelaatio on täsmälleen pienin unionin $S \cup C$ generoima ekvivalenssirelaatio, joka tässä tilanteessa on $L$. Tämän tuloksen avulla heidän mukaansa Birnbaumin tulos seuraa. Näin ollen määrittelemällä näyttöfunktio tällä tavalla Evansin argumentti ei Bergerin ja Penãn mukaan siis ole enää pitävä. Heistä Evans ei osoita, ettei käyttäen Birnbaumin määritelmiä tyhjentävyys- ja ehdollistamisperiaatteesta seuraisi vahvaa uskottavuusperiaatetta. Käyttämällä erilaista määritellyä kritiikki ei täten tavallaan kohdistu Birnbaumin todistuksiin tai määritelmiin, vaan tähän uuteen viitekehykseen, joka ei kuitenkaan ole sama kuin Birnbaumilla.

Heidän mielestä ei ole myöskään yllättävää, etteivät tyhjentävyysperiaatteen $S$ ja ehdollistamisperiaatten $C$ relaatiot johda vahvan uskottavuusperiaatteen $L$ määrittelemään relaatioon. Jos tämä olisi totta, perustat, joilla on verrannolliset uskottavuusfunktiot, olisivat päättelyn kannalta yhtäpitävät kumman tahansa periaatteen $C$ tai $S$ perusteella, mikä on heistä selvästi väärin. Totta on heidän mukaan kuitenkin sama mitä myös Evans totesi, eli unionin $S \cup C$ generoima pienin mahdollinen ekvivalenssirelaatio on yhtä suuri kuin $L$.

### 6.2.2 Mayon artikkeli

Luku perustuu Bergerin ja Penãn [5, luku 3] näkemykseen Mayon artikkelista, eikä niinkään Mayon omaan artikkeliin. Kaikkia Mayon artikkelissa esitettyjä argumentteja ei siis tässä käsitellä.

Yleisesti ottaen Bergerin ja Penãn mukaan Mayon kritiikki Birnbaumin lauseesta kumpuaa erilaisesta määritelmästä tyhjentävyysperiaatteesta kuin esimerkiksi määritelmässä 3 . Mayo määrittelee ehdollistamisperiaatteen ja tyhjentävyysperiaatteen hyödyntämällä seuraavia notaatioita:

- $\mathrm{Infr}_f[\boldsymbol{y}]$: tunnetun tai kiinteän perustan $(f, \boldsymbol{y})$ pohjalta tehty päätelmä parametrista $\theta$.

- $(f, \boldsymbol{y}) \Rightarrow \mathrm{Infr}_f[\boldsymbol{y}]$: perustan $(f, \boldsymbol{y})$ pohjalta tehty päätelmä parametrista $\theta$ tulee suorittaa menetelmällä $\mathrm{Infr}_f[\boldsymbol{y}]$

Vaikka se ei aivan eksplisiittistä näiden perusteella ole, Bergerin ja Penãn mukaan Mayo tekee eron menetelmien tuloksen ja niiden perusteella tutkijan tekemän päätelmän välillä. Tämä ero voidaan heidän mukaan selvemmin ilmaista näin:

- $\mathbf{M}(f, \boldsymbol{y})$: lopputulos kun menetelmää $\mathbf{M}$ on sovellettu perustaan $(f, \boldsymbol{y})$

- $\mathbf{Ev}(f, \boldsymbol{y})$: päätelmä, jonka tutkija tekee ottaen huomioon perustan $(f, \boldsymbol{y})$.

Voidaan tulkita, että $\mathbf{M}$ vastaa merkintää $\mathrm{Infr}_f[\boldsymbol{y}]$ ja $\mathbf{Ev}$ taas merkintää $(f, \boldsymbol{y}) \Rightarrow \mathrm{Infr}_f[\boldsymbol{y}]$. Käytännössä menetelmän $\mathbf{M}$ voi ajatella tarkoittavan esimerkiksi testisuureen arvoa tai p-arvoa, kun taas $\mathbf{Ev}$ on sen perusteella tehty päätelmä. Mayon määritelmissä käytetty $\mathrm{Infr}_f[\boldsymbol{y}]$ ja siinä oleva "Infr" antaa Bergerin ja Penãn mukaan ymmärtää, että $\mathrm{Infr}_f[\boldsymbol{y}] = \mathbf{Ev}(f, \boldsymbol{y})$. Tämä ei kuitenkaan heidän tulkintansa mukaan ole totta, sillä $\mathrm{Infr}_f[\boldsymbol{y}]$ kuvaa enemmin menetelmän lopputulosta eikä sen perusteella tehtyä päätelmää $\mathbf{Ev}(f, \boldsymbol{y})$. Merkintä $(f', \boldsymbol{y}') \Rightarrow \mathrm{Infr}_f[\boldsymbol{y}]$ heidän mukaansa tarkoittaa, että $\mathrm{Infr}_f[\boldsymbol{y}]$ ei tarvitse olla sama kuin päätelmä $\mathbf{Ev}(f, \boldsymbol{y})$.

Näitä merkintöjä voidaan tulkita seuraavanlaisesti: ottaen huomioon perustan $(f, \boldsymbol{y})$, tutkija tekee päätelmiä $\mathbf{Ev}(f, \boldsymbol{y})$ menetelmän $\mathbf{M}(f', \boldsymbol{y}')$ avulla erään perustan $(f', \boldsymbol{y}')$ avulla, mikä ei välttämättä ole sama kuin $(f, \boldsymbol{y})$. Merkintä $\mathbf{M}(f, \boldsymbol{y}) = \mathbf{M}(f', \boldsymbol{y}')$ tarkoittaa, että lopputulos on sama sen jälkeen, kun menetelmää $\mathbf{M}$ on sovellettu perustoihin $(f, \boldsymbol{y})$ ja $(f', \boldsymbol{y}')$. $\mathbf{Ev}(f, \boldsymbol{y}) = \mathbf{Ev}(f', \boldsymbol{y}')$ taas tarkoittaa, että tutkija tekee saman tulkinnan perustojen $(f, \boldsymbol{y})$ ja $(f', \boldsymbol{y}')$ pohjalta.

Tarkastellaan sitten, miten Mayo on merkinnöillään määritellyt ehdollistamis- ja tyhjentävyysperiaatteen.

- **CP:** Jos otetaan huomioon perusta $(f^*, (j, \boldsymbol{y}_j))$, ehdollistetaan perustuen valittuun malliin $f_j$, josta seuraa tulos $(f^*, (j, \boldsymbol{y}_j) \Rightarrow \mathrm{Infr}_{f_j}[\boldsymbol{y}]$. Ei tule käyttää ehdollistamatonta muotoa: $(f^*, (j, \boldsymbol{y}_j)) \not\Rightarrow \mathrm{Infr}_{f^*}[\boldsymbol{y}]$.

Edellinen on käytännössä sama kuin määritelmässä 6. Seuraavaksi kuitenkin Mayo määrittelee käyttäen omia merkintöjään tyhjentävyysperiaatteen seuraavalla tavalla:

- **SP2**: Jos on olemassa tyhjentävä tunnusluku $\boldsymbol{T}$ parametrille $\theta$ ja $\boldsymbol{T}(\boldsymbol{y}) = \boldsymbol{T}(\boldsymbol{y}')$, niin $\mathrm{Infr}_f[\boldsymbol{y}] = \mathrm{Infr}_f[\boldsymbol{y}']$.

Tässä on merkitty Mayon määrittelemä tyhjentävyysperiaate merkinnällä SP2, koska kyseessä ei ole aivan sama periaate kuin tutkielman määritelmän 3 periaate. Näin on tehty myös siksi, että Berger ja Penã merkitsivät periaatteen vastaavalla tavalla. Mayon määrittelemä tyhjentävyysperiaate voidaan vielä eksplisiittisesti ilmaista menetelmän $\mathbf{M}$ avulla:

- **SP2**: Jos on olemassa tyhjentävä tunnusluku $\boldsymbol{T}$ parametrille $\theta$ ja $\boldsymbol{T}(\boldsymbol{y}) = \boldsymbol{T}(\boldsymbol{y}')$, niin $\mathbf{M}(f, \boldsymbol{y}) = \mathbf{M}(f, \boldsymbol{y}')$.

Määritelmää 3 vastaava muotoilu olisi taas: Jos on olemassa tyhjentävä tunnusluku $\boldsymbol{T}$ parametrille $\theta$ ja $\boldsymbol{T}(\boldsymbol{y}) = \boldsymbol{T}(\boldsymbol{y}')$, niin $\mathbf{Ev}(f, \boldsymbol{y}) = \mathbf{Ev}(f, \boldsymbol{y}')$.

Näillä tyhjentävyys- ja ehdollistamisperiaatteen määritelmillä Bergerin ja Penãn mukaan Mayon johtopäätös on, ettei niistä seuraa vahvaa uskottavuusperiaatetta. Tarkemmin he tuovat tämän esiin esimerkissään [5, esim. 2], jossa he esittelevät yksityiskohtaisemmat perustelut tälle väitteelle. Yleisesti kuitenkin heidän mukaan tarkastelemalla näitä kahta periaatetta voidaan huomata, että CP on päätelmän **Ev** ominaisuus, kun taas SP2 on menetelmän **M** ominaisuus. Bergerin ja Penãn tulkinnan mukaan nimenomaan tästä erottelusta johtuen Mayon viitekehyksessä ei seuraa Birnbaumin osoittamaa tulosta. Jos kuitenkin määriteltäisiin ehdollistamisperiaate uudestaan, jotta se olisi myös menetelmän **M** ominaisuus, tällöin seuraisi vahva uskottavuusperiaate **M**:än ominaisuutena. Heistä Mayokaan ei siis osoita varsinaista Birnbaumin todistusta vääräksi, sillä Birnbaumin tyhjentävyysperiaatteen määritelmä oli erilainen kuin Mayolla. Birnbaum ei siis tehnyt vastaavaa erottelua menetelmien tuloksen ja tutkijan tekemän päätelmän välillä.

# 7 Yhteenveto ja pohdintaa

Työssä käsiteltiin uskottavuusfunktion heikkoa ja vahvaa versiota sekä tarvittavin osin tilastollisen päättelyn yleisiä periaatteita. Suuri osa työstä liittyi myös Birnbaumin lauseeseen, ja erityisesti sen todistukseen ja siihen kohdistuneeseen kritiikkiin.

Työtä olisi voinut jatkaa esimerkiksi tarkastelemalla, miten aputunnuslukuperiaatteen voidaan osoittaa olevan yhtäpitävä vahvan uskottavuusperiaatteen kanssa. Tästä lisää lähteissä [5, luku 4] ja [6]. Samoin aputunnuslukuperiaatetta eli vahvaa ehdollistamisperiaatetta olisi voinut käsitellä vielä tarkemmin, ja pohtia, millaisia seurauksia periaatteella yleisesti on tilastollisessa päättelyssä. Nämä teemat eivät kuitenkaan suoranaisesti liittyneet tutkielman aiheeseen, joten ne jäivät aiheen rajauksen ulkopuolelle.

Myös tyhjentävyyttä olisi voinut käsitellä nykyistä tarkemmin, ja määritellä esimerkiksi minimaalisen tyhjentävän tunnusluvun ja erottelevan tunnusluvun (complete statistic). Minimaalisella tunnusluvulla tarkoitetaan lyhyesti sellaista tunnuslukua, joka on maksimaalisesti tiivistänyt aineistoa samalla kuitenkin säilyttäen kaiken tiedon parametrista. Aiheista lisää lähteissä [4, luku 6.2.2] ja [2, luku 3]. Nämä konseptit eivät olleet oleellisia uskottavuusperiaatteen käsittelyssä, mistä syystä niitä ei tässä tutkielmassa tarkasteltu.

Mielenkiintoista olisi ollut myös tarkemmin käydä läpi Birnbaumin lausetta käsitteleviä artikkeleita. Alkuperäisissä artikkeleissa tarkastellaan Birnbaumin lausetta ja siihen liittyviä periaatteita tarkemmin kuin mitä tutkielman luvuissa 6.2.1 ja 6.2.2 oli mahdollista. Olisi ollut kiinnostavaa myös tarkastella muitakin Birnbaumin lauseeseen kohdistuvia artikkeleita, kuten esimerkiksi Kalbfleischin artikkeli, joka on mainittu ainakin [7, luku 5.1] ja [4, luku 6.4]. Kyseinen julkaisu on kuitenkin vuodelta 1975, joten se ei ole välttämättä enää niin ajankohtainen. Muun muassa tästä syyystä kyseinen artikkeli jäi käsittelemättä.

# Viitteet


[1] Henri Nyberg ja Pekka Nieminen (2024) *Tilastollinen päättely I ja II*, Turun yliopisto

[2] Pekka Nieminen (2024) *Tilastollisen päättelyn teoria*, Turun yliopisto

[3] Yudi Pawitan (2001) *In All Likelihood: Statistical Modelling and Inference Using Likelihood*, Oxford University Press

[4] George Casella ja Roger L. Berger (2002) *Statistical Inference*, Duxbury

[5] Víctor Penã ja James O. Berger (2017) *A note on recent criticisms to Birnbaum's theorem*, arXiv:1711.08093v1

[6] Michael Evans (2013) *What does the proof of Birnbaum's theorem prove*?, Electronic Journal of Statistics 7, 2645–2655

[7] Deborah G. Mayo (2014) *On the Birnbaum Argument for the Strong Likelihood Principle*, Statistical Science 29(2), 227–239

[8] Allan Birnbaum (1962) *On the Foundations of Statistical Inference*, Journal of the American Statistical Association, Vol. 57 Nr. 298 s. 269–306