Teollisuusyrityksen tuotteiden kysynnän ennustaminen asiantuntijahaastatteluiden avulla

Oskari Luomala

Tilastotieteen pro gradu -tutkielma

Matematiikan ja tilastotieteen laitos
Jyväskylän yliopisto

13. kesäkuuta 2018

JYVÄSKYLÄN YLIOPISTO
Matematiikan ja tilastotieteen laitos

Luomala, Oskari: Teollisuusyrityksen tuotteiden kysynnän ennustaminen asiantuntijahaastatteluiden avulla

Tilastotieteen pro gradu -tutkielma (45 sivua) + 3 liitettä (10 sivua)
13. kesäkuuta 2018

Tiivistelmä

Kysynnän ennustaminen on tyypillinen ongelma teollisuusyrityksillä. Ennusteiden laatimisessa usein hyödynnetään tilaushistoriaa tai jotain muuta aineistoa. Laadituilla ennusteilla on vaikutusta koko yrityksen toimintaan ja päätöksentekoon. Aina ei kuitenkaan ole riittävää aineistoa saatavilla, joten tarvitaan muita keinoja ennusteiden tekemiseen. Yrityksen henkilöstöä voidaan tämänkaltaisessa tilanteessa hyödyntää, sillä heillä on kokemusta ja tietoa tilauksiin liittyvistä tekijöistä.

Ennusteisiin tarvittava tieto voidaan hankkia yrityksen henkilöstöltä asiantuntijahaastatteluin. Tiedon hankkiminen edellyttää tilastotieteen, todennäköisyysjakaumien ja haastattelutekniikoiden yhdistämistä. Ennustejakauman mallintamiseksi tarvitaan joitain tunnuslukuja jakaumasta. Kvartiilien määrittäminen puolitusmenetelmällä on luotettavaksi todettu menetelmä tunnuslukujen määrittämiseen. Ennustejakauman sovittaminen haastattelun aikana edesauttaa kommunikointia ja helpottaa jakauman sopivuuden määrittämistä.

Tässä tutkielmassa esitellään interaktiivinen haastattelutyökalu, joka sovittaa jakauman eksaktisti asiantuntijan määrittämiin kvantiileihin. Jakauman sovittamisessa ja hienosäädössä hyödynnetään polynomisia kvantiilisekoituksia ja L-momentteja. Työkalu mahdollistaa jakauman tarkastelun numeerisesti ja visuaalisesti. Tulokset tallennetaan lokitietoina, joista haastatteluiden tulokset voidaan johtaa.

Työkalun avulla mallinnettiin jyväskyläläisen teollisuusyrityksen Black Bruin Oy:n tuotteille tilausten ennustejakaumat vuoden 2018 ensimmäisessä kvartaalissa. Yrityksestä valikoitiin asiantuntijoita, joille asiantuntijahaastattelu tehtiin. He saivat ennakkotietoina aineistoa aikaisemmista tilauksista ja täyden ohjeistuksen haastattelun kulusta ja huomioitavista seikoista. Kohteena oli muutama esimerkkituote, jotka kattavat suuren osan yrityksen liikevaihdosta. Asiantuntijoiden suoriutumista tutkittiin kollektiivisesti ja vertaillen tuotteita ja asiantuntijoita keskenään. Ennustejakaumia verrataan toteutuneisiin tilausmääriin.

Tulokset osoittivat, etteivät asiantuntijat suoriudu tehtävästä kovin hyvin yksinään. Yhdistetyt ennustejakaumat osoittivat, että kollektiivisesti ennusteet osuivat melko hyvin kohdalleen. Asiantuntijahaastattelu osoittautui toimivaksi tavaksi tuottaa ennusteita tilauksille aineiston hyödynnettävyyden ollessa vähäistä. Menetelmän kehittäminen yleiskäyttöisemmäksi edellyttää pieniä parannuksia haastattelumenetelmään ja esiteltyyn työkaluun. Tuloksien avulla Black Bruin pystyy kehittämään tilausten ja kysynnän ennustamista. Yritys sai lisää tietoa tilaushistorian hyödynnettävyydestä ja henkilöstön tietämyksestä tilausten ennustamisessa.

Avainsanat: kysynnän ennustaminen, asiantuntijahaastattelu, haastattelutyökalu, L-momentit, kvantiilisekoitukset, todennäköisyysjakaumien yhdistäminen

UNIVERSITY OF JYVÄSKYLÄ
Department of Mathematics and Statistics

Luomala, Oskari: Predicting the Incoming Orders of an Industrial Company Using Expert Elicitation

Master's Thesis in Statistics (45 pages) + 3 appendices (10 pages)
June 13, 2018

Summary

Forecasting incoming orders is a typical problem in the industry. The estimates of future orders affect the operations and the precision has an impact on finance. When relevant data for predictions cannot be found other approaches are needed.

Employees hold relevant information regarding orders. Summarizing this information as a distribution requires statistical methods combined with elicitation techniques. Common and reliable method is to elicitate the quantiles of a distribution using the bisection method. Fitting the distribution during the interview helps in giving feedback and deciding whether the distribution is acceptable.

In this thesis, I present an interactive visualization tool which fits a distribution exactly to the given quantiles. The fitting method uses polynomial quantile mixtures and L-skewness and L-kurtosis as fine tuning parameters. The tool also makes it possible to inspect the fitted distribution numerically and visually. The results are saved as a log file which consists of all the operations made by the user.

The objective of this study was to predict the number of ordered products from an industrial company during the first quarter of 2018. Multiple experts from the company were elicitated separately. The experts were given the data on the previous orders and full guidance on things to consider in the interview. The elicitation was performed for a few example products which cover a large portion of the total orders. The performance was compared between the experts and between the products. Predicted orders were compared with the actual orders.

The results show that none of the experts excelled in the order prediction. Collectively the experts were able to predict the orders fairly well. The methods and the visualization tool used were found suitable for elicitation and for forecasting orders. Improvements would be possible in the briefing of the experts and also in the elicitation tool. The company can improve prediction of the incoming orders by utilizing its data and the knowledge of its employees.

Keywords: order prediction, expert elicitation, elicitation software, L-moments, quantile mixtures, opinion pool, combining quantiles, lognormal distribution

Sisältö

1 Johdanto

Teollisuusyrityksille tyypillinen ongelma on tuotteiden kysynnän ennustaminen. Kysyntä määrittää yrityksen toimintaa ja kannattavuutta, ja siksi on tärkeää pystyä ennustamaan kysyntää tulevaisuuteen. Mikäli ennusteet ovat epävarmoja, joutuu yritys ennakoimaan kysyntää varastoimalla tuotteita tuotannon sujuvuuden varmistamiseksi. Tämä tarkoittaa ylimääräisiä kustannuksia yritykselle. (Christopher ja Lee, 2004)

Jyväskyläläisellä teollisuusyrityksellä Black Bruin Oy:llä on kuvatun kaltainen ongelma kysynnän ennustamisessa. Ennustamisessa ei ole aiemmin huomioitu sellaisia taustatekijöitä, jotka eivät ilmene tilaushistoriasta. Tilauksiin vaikuttavien taustatekijöiden huomioiminen on mahdollista hyödyntämällä yrityksen työntekijöillä olevaa informaatiota. Asiantuntijoiden mielipiteiden hyödyntäminen on yleistä päätöksenteossa ja monimutkaisten ilmiöiden ennustamisessa (O'Hagan *et al.*, 2006, s. 9, 179).

Ennusteiden laatimiseksi tilauksille on tarpeen muodostaa asiantuntijan mielipiteestä todennäköisyysjakauma. Todennäköisyysjakauman mallintamiseksi on tehtävä asiantuntijahaastattelu. Haastattelu on monimutkainen prosessi, jossa vaaditaan huolellista suunnittelua ja asiantuntijoiden valmistelua (Garthwaite *et al.*, 2005). Jakauman mallintamiseksi asiantuntijalta kysytään tunnuslukuja mallinnuksen kohteesta. Näitä tunnuslukuja haastattelijan on käytettävä jakauman muodostamiseen. Lopuksi on varmistettava vastaako mallinnus asiantuntijan näkemystä.

Graafisen työkalun käyttöä asiantuntijahaastattelussa ja jakauman mallintamisessa on tutkittu hieman, ja esimerkiksi verkossa on saatavilla työkalu: *the MATCH Uncertainty Elicitation Tool* (Morris *et al.*, 2014). Intuitiivisiksi ja asiantuntijan kannalta helposti määritettäviksi tunnusluvuiksi jakaumalle ovat osoittautuneet kvartiilit (O'Hagan *et al.*, 2006, s. 103). Perinteiset jakaumat eivät ole kuitenkaan mahdollistaneet jakauman sovittamista eksaktisti kolmeen kvartiiliin.

Yksittäisen asiantuntijan sijaan on mahdollista hyödyntää useampaa asiantuntijaa ennustejakauman laatimisessa. Useamman eri asiantuntijan ennustejakaumat yhdistetään yhdeksi jakaumaksi. Yhdistämistä voidaan tehdä tiheysfunktioiden tai kvantiilifunktioiden avulla (Lichtendahl Jr. *et al.*, 2013).

Tässä tutkielmassa laaditaan ennusteet Black Bruinin viiden eri tuoteperheen tilausmäärille. Mallinnus perustuu asiantuntijahaastatteluihin ja apuna käytetään graafista työkalua. Työkalu toteutetaan osana tutkielmaa. Graafisen työkalun on tarkoitus haastattelutilanteessa edistää jakauman mallinnusta ja tiedon välittymistä asiantuntijan ja haastattelijan välillä. Työkalussa hyödynnetään polynomisia kvantiilisekoituksia, jotka mahdollistavat jakauman mallintamisen eksaktisti kvantiilien avulla. Useamman eri asiantuntijan ennustejakaumat yhdistetään yhdeksi ennustejakaumaksi tuoteperheelle. Vertailun kohteena on kollektiivinen ja yksilöllinen kyky tuottaa ennusteita tilausmäärille. Ennustettuja tilausmääriä verrataan toteutuneisiin. Black Bruin saa tuloksista tietoa työntekijöidensä kyvyistä ennustaa tilausmääriä. Tulosten pohjalta asiantuntijat voivat kehittyä ennusteiden laatimisessa ja yritys voi hyödyntää työntekijöitään jatkossa paremmin päätöksenteossa.

2 Tilausten ennustaminen ilman historiatietoja

Tilausten ennustaminen on yleinen ongelma teollisuusalan yrityksissä. Yrityksillä on tuotteita, joiden valmistamiseen tarvitaan komponentteja, jotka hankitaan muilta yrityksiltä tai tuotetaan itse. Saatavuus ja toimitusajat on ennakoitava, jotta omia tuotteita pystytään valmistamaan aikataulussa riittävällä varmuudella. Komponenttien ja valmiiden tuotteiden varastointi sitoo yrityksen pääomaa. Kyse on siis yrityksen tuotantoketjun suunnittelusta ja hallinnasta. Kattava teos aiheesta on esimerkiksi *Logistics & Supply Chain Management* (Christopher, 2016). Tilausten ennustaminen vaikuttaa koko tuotantoketjuun, vaikka se on vain yksi osa sitä.

Tilausten ennustamisella tarkoitetaan yritykseltä tilattavien tuotteiden määrän arvioimista annetussa aikaikkunassa. Ennusteiden avulla pyritään ennakoimaan tuotantoprosessia ja saavuttamaan säästöjä. Toinen keskeinen tavoite on varmistaa tuotannon sujuvuus ja tuotteiden toimittaminen sovitussa ajassa.

Yrityksellä käytössä oleva aineisto ei aina pysty vastaamaan ennustamisen tarpeisiin. Esimerkiksi tilaushistoriaa voisi käyttää ennustamaan tulevia tilauksia. Tilaushistoriasta ei välttämättä käy ilmi osa taustatekijöistä, joilla on vaikutusta toteutuviin tilauksiin. Pelkkää tilaushistoriaa käytettäessä oletetaan myös, että kysyntää kuvaava prosessi pysyy tulevaisuudessa samanlaisena kuin menneisyydessä. Yrityksen henkilöstöllä voi olla informaatiota tilausmääriin vaikuttavista tekijöistä, jotka eivät näy olemassa olevasta aineistosta. Ongelman ratkaisemiseksi tässä työssä perehdytään ennustamiseen käytettävän informaation hankkimiseen yrityksen henkilöstöltä.

2.1 Tilausten ennustaminen Black Bruin Oy:n tuotteille

Kaikki tiedot Black Bruinista pohjaavat yrityksen talousjohtajan, Janne Mustosen kanssa käytyihin keskusteluihin ja yrityksen verkkosivuihin (Black Bruin Oy, 2017). Yritys valmistaa isokokoisia hydraulimoottoreita, jotka ovat hintavia ja ne koostuvat lukuisista pienemmistä osista ja komponenteista. Näiden suuri määrä tarkoittaa sitä, että varastoitavia komponentteja on myös paljon ja niihin sitoutuu suuri määrä yrityksen varallisuutta.

Optimoimalla varastointitarvettaan yritys saisi vähennettyä varastossa olevaa tavaran määrää. Varaston kutistaminen ei kuitenkaan saa tuottaa tilanteita, jossa tarvittavia komponentteja ei ole ajoissa saatavilla. Komponenttien tarvetta on siis pystyttävä ennustamaan huomioiden niiden toimitusaika muilta valmistajilta. Myös valmiita tuotteita varastoidaan. Lisäksi tuotannossa on huomioitava yrityksen tuotantokapasiteetti.

Ennusteita on järkevää laatia moottoreiden kokoluokan tasolla, sillä valmiit lopputuotteet poikkeavat hieman toisistaan. Lisäyksiä moottoriyksiköihin tehdään asiakkaiden toiveiden mukaan, mutta niitä voidaan ajatella lisävarusteina. Keskeisiltä osiltaan asiakkaille lähtevät tuotteet samasta moottorin kokoluokasta ovat siis samanlaisia. Täten valtaosa tarvittavista osista määräytyy tilattujen moottorien kokoluokan perusteella. Valmiit tuotteet voidaan kategorisoida tuoteperheeksi moottorin tyypin ja kokoluokan perusteella. Ennustamalla tilausmäärää tuoteperheille, jotka muodostavat suuren osan yrityksen liikevaihdosta, voidaan myös liiketoimintaa ohjata ennusteiden perusteella.

Yrityksen valmistamat hydraulimoottorit ovat komponentteina esimerkiksi maaja metsätaloudessa käytettävissä työkoneissa. Työkoneet ovat pitkän kehitysprosessin tuotoksia, jolloin ne ovat markkinoillakin pitkään. Tämä tarkoittaa Black Bruinin kannalta pitkiä asiakkuussuhteita koneita valmistaviin yrityksiin. Lopullinen kysyntä määräytyy kuitenkin valmiiden sovelluksien, eli työkoneiden kysynnän perusteella. Maa- ja metsätalouden kausiluonteisuus ja vaihteleva kannattavuus realisoituvat tilausmääriin.

Lopputuotteeseen käytetyistä komponenteista osa valmistetaan itse ja loput tulevat muilta valmistajilta. Toimitus- ja valmistusajat komponenteille vaihtelevat viikosta useampaan viikkoon. Komponenttien tilauksissa ja varastoinnissa on siis huomioitava niiden saatavuus. Tuotannon ja kokoonpanon suunnittelun kannalta tilausten ennakointia on tehtävä liukuvasti vähintään kolme kuukautta eteenpäin. Yrityksen kanssa käytyjen keskustelujen perusteella tilausmäärää päädytään ennustamaan kvartaalipohjaisesti. Tässä tutkielmassa ennustaminen toteutetaan vuoden 2018 ensimmäiselle kvartaalille, eli tammi-, helmi- ja maaliskuun aikana saapuville tilauksille. Ennusteet lasketaan tuoteperheestä tilattujen tuotteiden kappalemäärinä.

2.2 Tilausten ennustaminen asiantuntijahaastatteluilla

Black Bruinin tuotteiden tilausten ennustamiseen historia-aineisto ei ole riittävä, koska tuotteissa ja kysynnässä tapahtuu muutoksia. Tähän ongelmaan etsitään ratkaisua yrityksen henkilöstöstä, joilla on tietoa muuttuvista tekijöistä tuotteiden kysyntään liittyen. Yrityksen henkilöstön informaatio muodostetaan ennustejakaumaksi asiantuntijahaastatteluiden avulla.

Tulevia tilauksia ennustettaessa tilaushistorialla huomioidaan ainoastaan aikaisempaa kehitystä. Tällöin toiminnassa tapahtuvat muutokset tekevät mallista käyttökelvottoman. Seuraavaksi tunnistetaan Black Bruinin tilauksiin vaikuttavia tekijöitä, joita ei voida selittää tilaushistorialla. Samalla perustellaan, miksi yrityksen asiantuntijoilta voisi löytyä informaatiota näistä tekijöistä.

Yrityksen toiminnasta on löydettävissä kaksi laajempaa tekijäkokonaisuutta, joilla on suuresti vaikutusta tuleviin tilauksiin.

- Yrityksellä on pitkät asiakkuussuhteet, jolloin tietoa asiakkuuksista on mahdollista hyödyntää.
- Yrityksessä tehdään päätöksiä tuotteiden kehitykseen ja vastaanotettaviin tilauksiin liittyen.

Ensimmäinen tilauksiin vaikuttava tekijä on asiakkaiden toiminta. Pitkien asiakkuuksien ansiosta yrityksellä voi olla tietoa asiakkaan kehityksestä, investoinneista tai huonoista näkymistä. Asiakkaiden toiminta voi olla vakaata tilauksien suhteen, mutta muutokset voivat olla historia-aineiston valossa täysin ennakoimattomia. Työkoneiden kysynnän vaihtelu realisoituu työkoneita valmistavien asiakkaiden kautta, jolloin heidän toiminnastaan on tärkeää saada informaatiota. Informaatio asiakkaista välittyy yritykseen myyntiä ja tilauksia hoitavien henkilöiden kautta.

Yrityksen päätöksenteolla on vaikutusta saapuviin tilauksiin. Tuotekehityksen ansiosta uusia tuotteita syntyy ja vanhoja poistetaan myynnistä. Tuotteiden hinnoittelussa tapahtuu muutoksia ja myynnin käytänteissä on rajoituksia. Esimerkiksi tilaukselle on asetettu rajoite vähimmäistilausmäärästä. Näiden seurauksena asiakkaat mukautuvat uuteen tilanteeseen ja tilausmäärissä tapahtuu muutoksia. Päätöksentekoon lukeutuu lisäksi markkinointi ja asiakkaiden kanssa solmitut sopimukset.

Tehtyjen päätöksien huomioiminen tilausten ennusteissa on siis tärkeää ja perusteltua. Ylimmän johtoportaan henkilöt tekevät päätöksiä ja heidän on pystyttävä arvioimaan myös niiden vaikutuksia tilausmääriin.

Kysynnässä esiintyy kausiluonteisuutta, joka selittyy markkinoilla ja maa- ja metsätalouden toiminnan jaksottumisella. Kausiluonteisuus on estimoitavissa tilaushistoriasta, joskaan ei kaikissa tapauksissa, esimerkiksi uudempien ja muuttuneiden tuotteiden kohdalla. Vaikka tilaushistoria ei sellaisenaan riitä ennustamiseen, voidaan sitä hyödyntää pohjatietona asiantuntijoille.

Kaikki mainitut tilauksiin vaikuttavat tekijät voidaan ajatella olevan yrityksen sisällä tiedossa. Miten nämä tiedot hankitaan ja hyödynnetään tilausten ennustamisessa? Seuraavaksi perehdytään asiantuntijahaastatteluiden avulla tapahtuvaan informaation hankkimiseen ja mallintamiseen.

3 Informaation hankkiminen todennäköistämisellä

Yrityksen henkilöstöstä valitaan asiantuntijat, jotka ottavat huomioon ennusteissaan tilausmäärään vaikuttavat tekijät. Omaa tietämystään ja aineistoa hyödyntäen asiantuntijat rakentavat kokonaiskuvan tilausmäärän ennustejakaumasta. Tieto jakaumasta on saatava välittymään asiantuntijalta haastattelijalle. Tässä luvussa määritellään informaation hankkiminen todennäköistämisellä ja esitellään mallinnuksen toteutustapa. Teoria pohjaa teokseen *Uncertain Judgements: Eliciting Experts' Probabilities* (O'Hagan *et al.*, 2006).

3.1 Todennäköistämisen teoriaa

Yleisesti asiantuntijalta hankitun tiedon keräämisestä käytetään englannin kielessä nimitystä *expert elicitation*. Jakauman muodostamisen tilanteessa on käytetty nimitystä *probability distributions elicitation* (Garthwaite *et al.*, 2005), jonka suomennan todennäköistämiseksi. Todennäköistämisessä mallinnetaan haastattelulla asiantuntijan informaatio jakaumaksi. Haastattelun tueksi luodaan työkalu, joka auttaa hahmottelemaan sovitetun jakauman. Todennäköistämisen tuloksena syntynyttä jakaumaa voidaan käyttää priorijakaumana Bayes-mallinnuksessa, tai ennustejakaumana itsessään (O'Hagan *et al.*, 2006, s. 9). Tässä työssä perehdytään mallinnetun jakauman hyödyntämiseen ennusteena sellaisenaan.

Todennäköistäminen vaatii huolellista valmistelua. Todennäköistämisen tavoitteena on muodostaa asiantuntijan näkemyksen pohjalta todennäköisyysjakauma annetusta suureesta, joka tässä työssä on tuoteperheen tilausmäärä kolmen kuukauden ajanjaksolla. Todennäköistämisessä on neljä vaihetta, joista vaiheet 2-4 toistuvat:

1. valmistelu
2. tunnuslukujen selvittäminen asiantuntijalta
3. jakauman sovitus
4. tulosten riittävyyden tarkastelu

(Garthwaite *et al.*, 2005). Esitellään vaiheet ensin yleisemmällä tasolla. Myöhemmissä luvuissa perehdytään käytännön tasolla eri vaiheiden toteutukseen ja niihin liittyviin ongelmiin ja ratkaisuihin.

Valmistelu on aikaa vievin osa todennäköistämisessä. Se pitää siis sisällään asiantuntijoiden valinnan, heidän kouluttamisen ja tutkimusongelman mallinnettavan osan määrittämisen. Asiantuntijat on syytä perehdyttää käytettävään menetelmään ja heille on tarjottava riittävästi lisäinformaatiota haastateltavasta aiheesta. Lisäinformaatio voi esimerkiksi olla aikaisempia tilausmääriä. Vaarana on kuitenkin lukkiutuminen ennalta annettuun asetelmaan. Sopiva määrä ennakkotietoa auttaa kuitenkin asiantuntijaa muodostamaan lopullisen näkemyksen omiin tietoihin perustuen. Ennen haastattelua saatujen tietojen avulla asiantuntijat ehtivät perehtyä tutkimusongelmaan ja muodostaa paremman kuvan ilmiöstä ja pohtia ennakkoon asioita toivotulta suunnalta. Samasta aiheesta useampaa eri henkilöä haastateltaessa on syytä myös huomioida, etteivät he sovi keskenään vastaavansa tietyllä tavalla. Tällöin tuloksissa voi näkyä samansuuntaisuutta ja informaatiota menetetään ilmiöstä ja henkilöiden poikkeavista näkemyksistä.

Tunnuslukujen selvittäminen on merkittävin osa todennäköistämistä. Tunnuslukuihin perustuen määritetään todennäköisyysjakauma, joka kuvastaa asiantuntijan näkemystä. Jakauman mallintamiseksi on asiantuntijan pystyttävä kertomaan joitain suureita jakaumasta ja niiden arviointiin liittyy lukuisia psykologisia tekijöitä. Erilaiset psykologiset tekijät ja niiden vaikutusten arviointi sivuutetaan tässä työssä. Myöhemmin tässä luvussa esiteltävä puolitusmenetelmä on todettu melko luontevaksi lähestymistavaksi ihmismielen kannalta. Puolitusmenetelmässä jakaumasta selvitetään kvantiileja. (O'Hagan *et al.*, 2006, s. 97-106)

Jakauman sovittaminen tapahtuu annettuihin tunnuslukuihin perustuen. Valitut tunnusluvut, jakaumaperhe ja sovittamiseen käytetty menetelmä rajoittavat kaikki osittain toisiaan (Garthwaite *et al.*, 2005). Menetelmään kantaa ottamatta jakaumalle saadaan määritettyä tässä kohtaa muoto ja se voidaan visualisoida. Visualisoinnin tavoitteena on tuoda saataville lisäinformaatiota jakaumasta. Jakauman määrittämisen myötä siitä saadaan johdettua muitakin tunnuslukuja. Jakauman sovittamisessa käytettävään menetelmään perehdytään yksityiskohtaisesti luvussa 4.

Tulosten riittävyyden tarkastelussa on kyse sovitetun jakauman osuvuudesta asiantuntijan näkemykseen. Graafinen tarkastelu ja erilaiset tunnusluvut syntyneestä jakaumasta tuovat lisäinformaatiota ilmiöstä ja se mahdollistaa yksityiskohtaisemman keskustelun tuloksista. Haastattelijan tehtävänä on esittää tarkentavia kysymyksiä jakaumasta, joiden perusteella hän päättelee kuvaako mallinnettu jakauma asiantuntijan näkemystä täsmällisesti. Mikäli tuloksiin ei olla tyytyväisiä, korjataan selvitettyjä tunnuslukuja ja sovitetaan jakauma uudelleen. Tunnuslukuja korjataan tai lisätään kunnes haastattelija on varmistunut, että jakauma kuvaa riittävällä tarkkuudella asiantuntijan näkemystä ilmiöstä. Koko prosessi on siis iteratiivinen.

Seuraavaksi käsitellään yksityiskohtaisemmin jakauman tunnuslukujen hankkimiseen käytettävä menetelmä. Jakauman sovittaminen kvantiileihin ja todennäköistämisessä apuna käytettävä työkalu vaativat omat lukunsa.

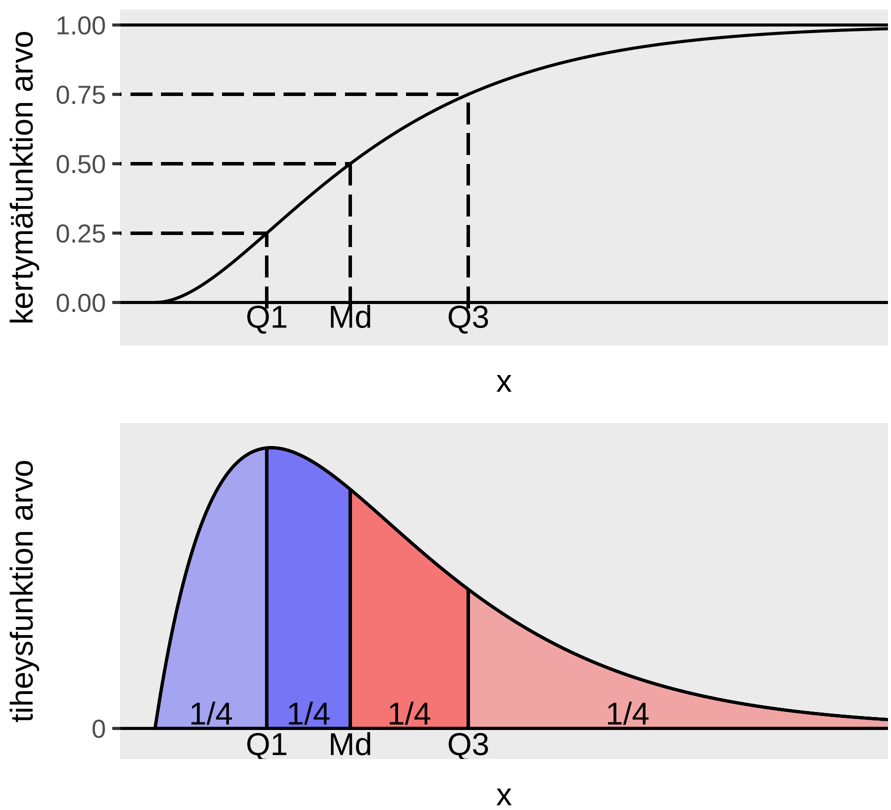
3.2 Puolitusmenetelmä

Jakauman tunnusluvut määritellään haastattelussa puolitusmenetelmällä (*bisection method*) (O'Hagan *et al.*, 2006). Puolitusmenetelmän avulla jakaumasta saadaan määrättyä kvantiileja. Puolitusmenetelmän ideana on kertymäfunktion jakaminen yhtä suuriin osiin. Osien jakopisteet ovat määritettäviä kvantiiliarvoja. Tyypillinen jako on kolmen pisteen avulla, jolloin tuloksena saadaan jakauman kvartiilit.

Puolitusmenetelmässä todennäköisyysjakauma jaetaan ensin kahtia - jakopiste on jakauman mediaani. Puolitetut osat jaetaan kahtia, jolloin jakaumalle saadaan ala- ja yläkvartiili. Puolitusmenetelmän idea on havainnollistettu kuvassa 1. Kuvassa nähdään tiheysfunktio ja sitä vastaava kertymäfunktio. Asiantuntijalle tähän tapaan hahmottaminen ei välttämättä ole tuttua, joten tarkka ohjeistus on tarpeen.

Asiantuntijaa pyydetään aluksi arvioimaan sellainen tilausmäärä, jonka ylittyminen ja alittuminen ovat hänen mielestään yhtä todennäköisiä. Tämä nimetty tilausmäärä on jakauman mediaani. Kuvasta katsottuna tarkoitus on jakaa todennäköisyysjakauma siten, että sinisen ja punaisen alueen pinta-alat ovat yhtä suuret.

Ylä- ja alakvartiilin määrittäminen vaativat hieman enemmän ohjeistusta. Ylä-



Kuva 1: Puolitusmenetelmä havainnollistettuna kertymä- (ylhäällä) ja tiheysfunktion (alhaalla) avulla, todennäköisyysjakauma on määräytynyt asiantuntijan kvartiileista eksaktisti; *pystyviivat:* asiantuntijan määräämät kvartiilit, *värit:* eri väriset alueet ovat todennäköisyysmassaltaan yhtä suuria

kvartiilin määrittämiseksi asiantuntijaa ohjeistetaan kertomalla: *Saat varman tiedon, että tilausmäärä ylittyy äsken nimeämästäsi arvioistasi. Nimeä uusi arvio ennusteelle tässä tilanteessa siten, että on yhtä todennäköistä ylittää uusi arvio, tai että se on uuden ja vanhan arviosi välissä.* Nyt mediaanin yläpuolella oleva todennäköisyysmassa on saatu jaettua kahteen yhtä suureen osaan. Kuvasta 1 katsottuna punainen alue tiheysfunktiossa on jaettu kahtia. Alakvartiilin määrittämisen kohdalla toimitaan vastaavasti. Tällöin puolestaan kuvan 1 sininen alue pyritään jakamaan kahtia.

Menetelmän etuna haastattelutilanteessa on kvartiilien intuitiivinen tulkinta. Yhtä suurten ositteiden perusteella asiantuntijan kanssa on helpompaa käydä keskustelua jakaumasta. Esimerkiksi haastattelija voi varmistaa, että jakopisteiden väliin jäävät alueet todellakin ovat asiantuntijan mielestä yhtä todennäköisiä. Myöhemmin esiteltävä graafinen työkalu on myös olennaisesti tässä apuna. Tarvittaessa kvartiileja korjataan jakauman tarkemman tarkastelun myötä vastaamaan paremmin asiantuntijan näkemystä.

Kvantiilien avulla jakauman parametrien sovittamiseksi on kirjallisuudessa esitetty pienimmän neliösumman menetelmää (Morris *et al.*, 2014). Pienimmän neliösumman menetelmässä sovitetaan jakauman malliparametrit minimoiden tappiofunktio. Tappiofunktio lasketaan teoreettisen jakauman kvantiilipisteiden neliöitynä etäisyytenä asiantuntijan antamista kvantiileista (Morris *et al.*, 2014). Menetelmän heikkoutena on, että jakauma ei useimmiten sovi täsmällisesti annettuihin kvantiileihin.

Tässä työssä tutkitaan vaihtoehtoista tapaa jakauman sovittamiseen, missä jakaumalla on vähintään yhtä monta malliparametriä, kuin on määritettyjä jakauman tunnuslukuja. Kvantiilisekoitusten avulla jakauman sovittaminen voidaan tehdä täsmällisesti useimmissa tilanteissa. Seuraavaksi esitellään kvantiilisekoitusten ja malliparametrien sovittamisen teoria.

4 Teoriaa mallinnuksesta - polynomiset kvantiilisekoitukset

Aikaisemmin esitelty puolitusmenetelmä tuottaa jakaumasta tunnuslukuina kvantiileja. Kvantiilien hyödyntämiseksi esittelen tässä luvussa yleisesti kvantiilisekoitukset ja tarkemmin polynomiset kvantiilisekoitukset. Kvantiilisekoitukset mahdollistavat jakauman täsmällisen määrittämisen kvantiileista. Polynomisia kvantiilisekoituksia käytetään myöhemmin graafisessa todennäköistämistyökalussa jakauman mallintamiseen ja esittämiseen. Erikoistapauksena esitellään log-normaali polynominen kvantiilisekoitus, jota käytetään asiantuntijalta saatuihin tunnuslukuihin sovitettavana jakaumana.

Kvantiilisekoitus määritellään kvantiilifunktioiden lineaarikombinaationa. Kvantiilifunktio on kertymäfunktion käänteisfunktio, jatkuvalle satunnaismuuttujalle määriteltynä $Q(u) = F^{-1}(u)$. Diskreetille tai epäjatkuvalle kertymäfunktiolle kvantiilifunktio voidaan määrittää muodossa:

$$Q(u) = \inf_{x \in \mathbb{R}} \{u \leq F(x)\},$$

missä $u \in]0,1[$. Mikäli $Q(\cdot)$ on derivoituva, niin tiheysfunktio voidaan esittää kvantiilifunktion avulla:

$$f(Q(u)) = \frac{1}{Q'(u)}. \tag{1}$$

(Parzen, 1979)

Kvantiilisekoitus on jakaumaperhe, joka määritellään kvantiilifunktion kautta. Kvantiilisekoitus on useamman eri jakauman kvantiilifunktioiden lineaarikombinaatio:

$$Q(u) = \sum_{i=1}^{k} c_i Q_i(u), \tag{2}$$

missä $Q_i(\cdot)$ on kvantiilifunktio ja c_i on mallipararametri. Indeksi k on sekoitettavien jakaumien määrä. Malliparametrien c_i rajoitteena on, että $Q(\cdot)$:n on oltava kvantiilifunktio. Täten funktion $Q(u)$ on oltava kasvava funktio välillä $u \in]0,1[$. (Karvanen, 2006)

Polynomiset kvantiilisekoitukset ovat erikoistapaus, jossa lineaarikombinaatio muodostuu polynomitermeistä ja yhdestä, pohjajakaumana toimivasta, kvantiilifunktiosta. Polynomitermit ovat itse asiassa hyvin yksinkertaisia kvantiilifunktioita. Kaavasta (2) johdettuna polynominen kvantiilisekoitus on muotoa:

$$Q_{P_m}(u) = bQ_0(u) + \sum_{i=0}^{m} a_i u^i, \tag{3}$$

missä $Q_0(\cdot)$ on pohjajakauman kvantiilifunktio ja indeksi m on polynomin aste. Nyt malliparametreja (b, a_i) on $k = m+2$ kappaletta. Funktion $Q_0(\cdot)$ ollessa derivoituva, saadaan tiheysfunktio esitettyä kaavan (1) ja käänteisfunktion derivaatan nojalla muodossa:

$$\begin{aligned} f(x) = f(Q_{P_m}(u)) &= \frac{1}{bQ_0'(u) + \sum_{i=1}^{m} i a_i u^{i-1}} \\ &= \frac{1}{\frac{b}{f_0(Q_0(u))} + \sum_{i=1}^{m} i a_i u^{i-1}}. \end{aligned}$$

Näin ollen polynomisen kvantiilisekoituksen tiheysfunktio voidaan määritellä helposti, mikäli pohjajakaumalle on olemassa funktiot tiheyden f_0 ja kvantiilien Q_0 laskemiseen. Kertymäfunktion määrittämiseksi polynomiselle kvantiilisekoitukselle on turvauduttava numeeriseen käänteisfunktion ratkaisemiseen kvantiilifunktiosta.

Kvantiilisekoituksille voidaan määrittää jakaumasta satunnaisesti generoiva funktio kvantiilifunktion avulla. Funktion avulla jakaumasta voidaan simuloida toisistaan riippumattomia havaintoja. Simulointi perustuu käänteisen jakauman menetelmään (Robert, 2004). Kvantiilisekoitukselle oletetaan kvantiilifunktion olevan määritettävissä vähintään numeerisessa muodossa. Lisäksi oletetaan, että tasajakautuneesta satunnaismuuttujasta $U \sim \mathcal{U}[0, 1]$ voidaan simuloida havaintoja. Tällöin muunnoksella $Q(U)$ voidaan generoida havaintoja kvantiilisekoituksesta, joka noudattaa jakaumaa kertymäfunktiolla $F = Q^{-1}$.

4.1 L-momentit

L-momentit ovat tunnuslukuja, joilla voidaan kuvata jakauman muotoa, eli muun muassa sijaintia, skaalaa, vinoutta ja huipukkuutta. Seuraavaksi esitellään L-momenttien teoriaa ja laskennallisia ominaisuuksia. Teoria L-momenteista perustuu artikkeliin *L-moments: analysis and estimation of distributions using linear combinations of order statistics* (Hosking, 1990). L-momentteja hyödynnetään myöhemmin polynomisen kvantiilisekoituksen parametrien sovittamisessa.

Olkoon satunnaismuuttujan X, jonka kvantiilifunktio on Q. Tällöin satunnaismuuttujan X asteen r L-momentti määritellään:

$$\lambda_r = \int_0^1 Q(u) P^*_{r-1}(u)\, du, \tag{4}$$

missä polynomi

$$P^*_r(u) = \sum_{j=0}^{r} (-1)^{r-j} \binom{r}{j} \binom{r+j}{j} u^j \tag{5}$$

tunnetaan nimellä siirretty Legendren polynomi.

Kaavoista (4) ja (5) laskettuna neljä ensimmäistä L-momenttia ovat kvantiilifunktion avulla esitettyinä:

$$\begin{aligned}
\lambda_1 &= \int_0^1 Q(u)\, du \\
\lambda_2 &= \int_0^1 Q(u)(2u-1)\, du \\
\lambda_3 &= \int_0^1 Q(u)(6u^2-6u+1)\, du \\
\lambda_4 &= \int_0^1 Q(u)(20u^3-30u^2+12u-1)\, du.
\end{aligned} \tag{6}$$

L-momenttien tulkinnat ovat analogisia verrattuna tavallisiin momentteihin. λ_1 kuvaa jakauman sijaintia, λ_2 skaalaa, λ_3 vinoutta ja λ_4 huipukkuutta. Lisäksi määritellään skaalariippumattomat tunnusluvut L-vinous τ_3 ja L-huipukkuus τ_4 L-momenttien suhteina:

$$\tau_3 = \lambda_3/\lambda_2, \qquad \tau_4 = \lambda_4/\lambda_2. \tag{7}$$

Kvantiilisekoituksille L-momentit voidaan laskea yksittäisten kvantiilifunktioiden L-momenttien avulla:

$$\lambda_r(Q) = \lambda_r\left(\sum_{i=1}^{k} c_i Q_i\right) = \sum_{i=1}^{k} c_i \lambda_r(Q_i). \tag{8}$$

Ominaisuus seuraa kaavoista (2) ja (4).

4.2 Polynomisen kvantiilisekoituksen parametrien sovittaminen

Edellä esiteltiin L-momenttien ja kvantiilien yhteydet kvantiilifunktioon, joka määrittää jakauman täsmällisesti. Näitä yhteyksiä hyödyntäen on mahdollista sovittaa polynomisen kvantiilisekoituksen parametrit olettaen, että pohjajakauma on tiedossa ja lisäksi jakaumasta tiedetään tai on laskettavissa kvantiileja ja muotoa koskevia tunnuslukuja. Tämä menetelmä mahdollistaa siis parametrisen jakauman sovittamisen eksaktisti asiantuntijalle intuitiivisten tunnuslukujen avulla.

Todennäköistämisessä tarvitaan monenlaisiin tilanteisiin sopiva jakaumaperhe. Syynä on, ettei haastattelija tiedä ennakkoon, millainen näkemys asiantuntijalla on mitattavasta suureesta. Tämä tarkoittaa jakaumaperheen joustavuutta muodon suhteen. Polynomisen kvantiilisekoituksen ja L-momenttien avulla tätä ominaisuutta tavoitellaan estimoinnissa. Polynomisen kvantiilisekoituksen aste määrää samalla sekä vaadittavien tunnuslukujen lukumäärän että jakauman joustavuuden. Seuraavaksi esillään, miten kolme asiantuntijan antamaa kvantiilia ja tieto jakauman vinoudesta ja huipukkuudesta kytketään polynomiseen kvantiilisekoitukseen.

Olkoon sovitettava jakauma kolmannen asteen polynominen kvantiilisekoitus, Q_{P_3}, annetulla pohjajakaumalla Q_0. Lisäksi tunnetaan kvantiili x_i ja tätä vastaava kertymäfunktion arvo u_i Sijoittamalla kaavaan (3) tunnetut arvot saadaan yhtälö:

$$x_i = bQ_0(u_i) + a_3u_i^3 + a_2u_i^2 + a_1u_i + a_0, \tag{9}$$

missä tuntemattomia malliparametrejä ovat b, a_3, a_2, a_1, a_0. Parametrit voidaan määrätä täsmällisesti, mikäli jakaumasta tunnetaan viisi kvantiilia. Tällöin kvantiileja vastaavat yhtälöt voidaan esittää matriisimuodossa:

$$\begin{bmatrix} 1 & u_1 & u_1^2 & u_1^3 & Q_0(u_1) \\ 1 & u_2 & u_2^2 & u_2^3 & Q_0(u_2) \\ 1 & u_3 & u_3^2 & u_3^3 & Q_0(u_3) \\ 1 & u_4 & u_4^2 & u_4^3 & Q_0(u_4) \\ 1 & u_5 & u_5^2 & u_5^3 & Q_0(u_5) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ b \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}. \tag{10}$$

Yhtälöryhmän (10) ratkaisu edellyttää kuitenkin matriisin kääntyvyyttä. Ratkaisun löytyminen ei itsessään takaa jakauman olemassaoloa, sillä lisäksi malliparametrien määräämän funktion on oltava kvantiilifunktio. Ehto voidaan tarkistaa numeerisesti tarkastelemalla kvantilifunktion kasvavuutta. Jakauman olemassaolon tarkistamiseen palataan Todennäköistämistyökalun esittelyn yhteydessä.

Aiemmin luvussa 3.2 esitellyllä puolitusmenetelmällä saadaan asiantuntijalta haastateltua jakauman kolme kvantiilia. Kolmannen asteen polynomiseen kvantiilisekoitukseen jää siis kaksi parametriä kiinnittämättä. Jakauman vinouden ja huipukkuuden hyödyntämiseksi tutkitaan L-momenttien laskemista polynomiselle kvantiilisekoitukselle. L-momenttien laskeminen polynomiselle kvantiilisekoitukselle yksinkertaistuu lineaaristen ominaisuuksien ansiosta (kuten kaavassa (8)):

$$\lambda_r(Q_{P_m}) \overset{(3)}{=} \lambda_r\Big(bQ_0(u) + \sum_{i=0}^{m} a_i u^i\Big) \overset{(4)}{=} b\lambda_r(Q_0) + \sum_{i=0}^{m} a_i \lambda_r(u^i). \tag{11}$$

Näin ollen L-momentin laskeminen voidaan hajottaa yksittäisten kvantiilifunktioiden L-momenttien laskemiseksi. Pohjajakaumalle L-momentit lasketaan tarvittaessa numeerisesti, mutta polynomitermeille L-momenttien laskeminen on mahdollista suljetussa muodossa:

$$\begin{aligned}\lambda_r(u^m) &= \int_0^1 (u^m) P^*_{r-1}(u)\, du \\ &= \int_0^1 \sum_{j=0}^{r-1} u^{j+m} (-1)^{r-1-j} \binom{r-1}{j}\binom{r-1+j}{j}\, du \\ &= \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j}\binom{r-1+j}{j} \int_0^1 u^{j+m}\, du \\ &= \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j}\binom{r-1+j}{j} \frac{1}{j+m+1}.\end{aligned} \tag{12}$$

Skaalariippumattomat L-vinous τ_3 ja L-huipukkuus τ_4 ovat käyttökelpoisempia jakauman muodon säätämisessä, kuin λ_3 ja λ_4. Käyttökelpoisuus perustuu siihen, että L-vinouden ja L-huipukkuuden mahdolliset arvot ovat rajattuja (Hosking, 1990). Määritelmästä (7) johdetaan lineaarinen yhtälö, jossa malliparametrit ovat tuntemattomia:

$$\begin{aligned} & \tau_r = \frac{\lambda_r(Q_{P_m})}{\lambda_2(Q_{P_m})} \\ \Leftrightarrow \quad & \tau_r\, \lambda_2(Q_{P_m}) - \lambda_r(Q_{P_m}) = 0 \\ \Leftrightarrow \quad & b\left(\tau_r\, \lambda_2(Q_0) - \lambda_r(Q_0)\right) + \sum_{i=0}^{m} a_i \left(\tau_r\, \lambda_2(u^i) - \lambda_r(u^i)\right) = 0, \end{aligned} \tag{13}$$

missä $r \in \{3, 4\}$. Yhtälöä voidaan sieventää sijoittamalla polynomitermien L-momentit laskettuna kaavalla (12):

$$\begin{array}{llll} \lambda_2(1) = 0 & \lambda_2(u) = \frac{1}{6} & \lambda_2(u^2) = \frac{1}{6} & \lambda_2(u^3) = \frac{3}{20} \\ \lambda_3(1) = 0 & \lambda_3(u) = 0 & \lambda_3(u^2) = \frac{1}{30} & \lambda_3(u^3) = \frac{1}{20} \\ \lambda_4(1) = 0 & \lambda_4(u) = 0 & \lambda_4(u^2) = 0 & \lambda_4(u^3) = \frac{1}{140}. \end{array}$$

Malliparametrit voidaan nyt ratkaista hyödyntäen yhtälöitä (9) ja (13). Yhtälöryhmälle saadaan matriisiesitys:

$$\begin{bmatrix} 1 & u_1 & u_1^2 & u_1^3 & Q_0(u_1) \\ 1 & u_2 & u_2^2 & u_2^3 & Q_0(u_2) \\ 1 & u_3 & u_3^2 & u_3^3 & Q_0(u_3) \\ 0 & \tau_3 \frac{1}{6} & \tau_3 \frac{1}{6} - \frac{1}{30} & \tau_3 \frac{3}{20} - \frac{1}{20} & \tau_3\, \lambda_2(Q_0) - \lambda_3(Q_0) \\ 0 & \tau_4 \frac{1}{6} & \tau_4 \frac{1}{6} & \tau_4 \frac{3}{20} - \frac{1}{140} & \tau_4\, \lambda_2(Q_0) - \lambda_4(Q_0) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ b \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \\ 0 \end{bmatrix}. \tag{14}$$

Matriisiesityksen kolme ylintä riviä määräytyvät kvantiileista. Kaksi alinta riviä määräytyvät L-vinoudesta ja L-huipukkuudesta. Yhtälöryhmä on nopea ratkaista käänteismatriisin avulla, sillä ainoastaan malliparametrien vektori (a_0, a_1, a_2, a_3, b) on tuntematon. Kaikki tarvittavat funktiot malliparametrien määrittämiseen kvantiileista ja L-momenteista on toteutettu osana tätä työtä. Funktiot on totetutettu R-kielellä (R Core Team, 2017) ja ne löytyvät liitteestä C.

Esitetty ratkaisutapa yleistyy siten, että malliparametrien sovittamisessa voidaan käyttää mielivaltaista yhdistelmää kvantiileja ja L-momentteja. Ainoastaan polynomisen kvantiilisekoituksen aste muuttuu sovitettavien malliparametrien määrän mukaan. Asiantuntijoiden haastatteluissa selvitetään puolitusmenetelmällä ennustejakauman kvartiilit. Tällöin edellisissä merkinnöissä $u_1 = 0.25, u_2 = 0.5$ ja $u_3 = 0.75$. L-vinous ja L-huipukkuus puolestaan määritellään haastattelijan toimesta. Niiden tarkoituksena on auttaa hienosäätämään jakaumaa asiantuntijan antamien kommenttien perusteella.

4.3 Pohjajakauman valinta kvantiilisekoituksessa

Edellä esiteltiin kvantiileiden avulla määriteltävä jakaumaperhe. Jakauman sovittamisessa ei otettu vielä kantaa polynomisen kvantiilisekoituksen pohjajakauman valintaan. Pohjajakauma vaikuttaa olennaisesti jakauman ominaisuuksiin, joten valinnassa on otettava huomioon mallinnettava ilmiö. Valintakriteerien perusteella tehdään valinta käytettävästä pohjajakaumasta.

Tilausten ennustamisessa todennäköisyysjakauma halutaan rajoittaa positiiviselle reaaliakselille. Lisäksi voidaan ajatella, että todella suurten tilausmäärien kohdalla todennäköisyydet ovat lähes olemattomia. Eli todennäköisyysjakauma on lyhythäntäinen. Arviointi perustuu tietoon siitä, että yrityksen asiakkuudet ovat vakaita kysynnältään ja asiakkailla on rajat omalle tuotantokapasiteetilleen ja varastointimahdollisuuksilleen.

Pohjajakauma antaa jakaumalle määrittelyjoukon, jota polynomitermit siirtävät ja venyttävät rajallisesti. Koska tilausten määrä voi saada vain positiivisia reaaliarvoja, niin esimerkiksi gamma- ja log-normaalijakauma ovat sopivia pohjajakaumaksi. Molempien jakaumien tilanteessa on mahdollista edelleen sovittaa jakauma, joka saa positiivisia todennäköisyyksiä negatiivisillakin arvoilla. Tähän ominaisuuteen vaikuttaa kuitenkin ainoastaan polynomisen kvantiilisekoituksen vakiotermi a_0. Vakiotermin vaikutuksesta em. jakaumien tapauksessa tiheysfunktio saa arvon 0 välillä $]-\infty, a_0]$. Nollasta poikkeavan vakiotermin arvot tuottavat aina hieman ongelmia. Negatiivinen vakiotermi on ei-toivottu mallinnettavan ilmiön kannalta. Positiivinen vakiotermi puolestaan tarkoittaa, että pienet tilausmäärät ovat mahdottomia. Ongelman voisi korjata siten, että vakiotermille kiinnitetään arvo $a_0 = 0$. Se kuitenkin voisi vaikeuttaa parametrien sovittamista ja tuoda muita ei-toivottuja ominaisuuksia jakaumaan.

Pohjajakauman valitseminen lueteltujen vaatimusten pohjalta toteutetaan visuaalisten tarkastelujen avulla ja kokeilemalla erilaisia syötearvoja käytännössä. Kokeiluilla ei voida kuitenkaan saavuttaa täydellistä varmuutta jakauman toimivuudesta, sillä toivottu jakauman muoto paljastuu vasta aidossa todennäköistämisprosessissa. Tästä huolimatta jakauma kiinnitetään ennakkoon, jotta vaihtoehtoja jakauman sovittamisessa ei ole liikaa.

Lukuisten kokeilujen ja graafisten tarkastelujen myötä pohjajakaumaksi valitaan log-normaalijakauma. Valitun jakauman lisäksi testattiin gamma- ja normaalijakaumaa. Myös polynomisen kvantiilisekoituksen eksponenttimuunnosta kokeiltiin, mutta jakauman sovittamisessa ilmeni numeerisia ongelmia. Tulokset -osiossa tarkastellaan pohjajakuman valinnan onnistuneisuutta asiantuntijan näkemyksen mallintamisessa.

4.4 Log-normaalijakauma

Kokeilujen myötä pohjajakaumaksi valittiin log-normaalijakauma. Pohjajakauman hyödyntämiseksi polynomisen kvantiilisekoituksen sovittamisessa log-normaalijakaumasta käsitellään vain tarpeelliset ominaisuudet.

Olkoon normaalijakautunut satunnaismuuttuja $Y = \log X \sim \mathcal{N}(\mu, \sigma^2)$. Tällöin satunnaismuuttuja X on log-normaalijakautunut, $X \sim \text{Lognormal}(\mu, \sigma^2)$ (Crow ja Shimizu, 1987). Log-normaalijakauman tiheysfunktio on johdettavissa satunnaismuuttujan muunnoksella. Tiheysfunktioksi saadaan:

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma x} \exp\Big(- \frac{(\log(x) - \mu)^2}{2\sigma^2} \Big), \quad x > 0\,.$$

Kvantiilifunktio on puolestaan:

$$Q(u; \mu, \sigma^2) = \exp\big(\mu + \sigma\,\Phi^{-1}(u)\big), \quad 0 < u < 1\,.$$

Tiheys- ja kvantiilifunktioiden laskemiseen log-normaalijakaumalle on R-kielessä toteutettuna valmiit funktiot (R Core Team, 2017).

Polynomisen kvantiilisekoituksen pohjajakaumaksi valitaan Lognormal(0, 1). Jatkossa pohjajakaumaan viittaavat tiheysfunktio f_0 ja kvantiilifunktio Q_0 määritellään:

$$f_0(x) = \frac{1}{\sqrt{2\pi}\,x} \exp\Big(- \frac{\log^2(x)}{2} \Big), \quad x > 0\,,$$

$$Q_0(u) = \exp\big(\Phi^{-1}(u)\big), \quad 0 < u < 1\,.$$

Yhtälöryhmässä (14) esiintyy pohjajakauman L-momentteja. Log-normaalijakaumalle L-momentit lasketaan numeerisesti integroimalla, hyödyntäen kaavoja (6). Yhtälöryhmön (14) ratkaisemisessa tarvittavat L-momenttien arvot ovat: $\lambda_2(Q_0) = 0.858$, $\lambda_3(Q_0) = 0.397$ ja $\lambda_4(Q_0) = 0.252$.

5 Työkalu asiantuntijahaastatteluihin

Aiemmin esiteltiin teoreettiset perusteet parametrisen jakauman sovittamiseen kvantiileiden ja L-momenttien avulla. Tässä luvussa sovitetaan teoria käytäntöön ja esitellään graafinen työkalu jakaumien sovittamiseen. Työkalusta esitellään yksityiskohtaisesti toteutustapa ja ominaisuudet, joista on hyötyä todennäköistämisessä.

Graafisen työkalun käyttö on perusteltua asiantuntijahaastatteluissa, sillä todennäköisyysjakauman määrittäminen on mutkikas prosessi. Haastattelussa määritettyjen tunnuslukujen tueksi on pystyttävä esittämään tarkentavia kysymyksiä jakaumasta. Tämä edellyttää jo määritettyjen tunnuslukujen pohjalta jakauman muodostamista ja uusien tunnuslukujen johtamista. Graafista työkalua voidaan siis hyödyntää jakauman sovittamiseen, kuvaamiseen ja tunnuslukujen laskemiseen. Näin ollen graafisella työkalulla voidaan edistää kommunikointia ja helpottaa jakauman sopivuuden määrittämistä (James *et al.*, 2010).

Työkalun toteutus jakaantuu kolmeen osaan: käyttöliittymä, palvelin ja laskentafunktiot. Käyttäjälle näkyvä osa työkalua on käyttöliittymä, missä on tekstiä, tekstin syöttökenttiä, graafeja ja erilaisia painikkeita. Laskenta tapahtuu syötetietoja käyttäen ja tulokset tulevat käyttöliittymään näkyviin. Erillistä komentoa ei vaadita, että jakauma sovittuu, vaan se tapahtuu aina siihen vaikuttavien syötetietojen muutoksen jälkeen. Palvelin puolestaan toimii kahden edellämainitun välissä ja kutsuu annetuilla tiedoilla laskentafunktioita ja välittää tulokset takaisin graafisten elementtien näytettäväksi. Työkalu on toteutettu R-kielellä (R Core Team, 2017).

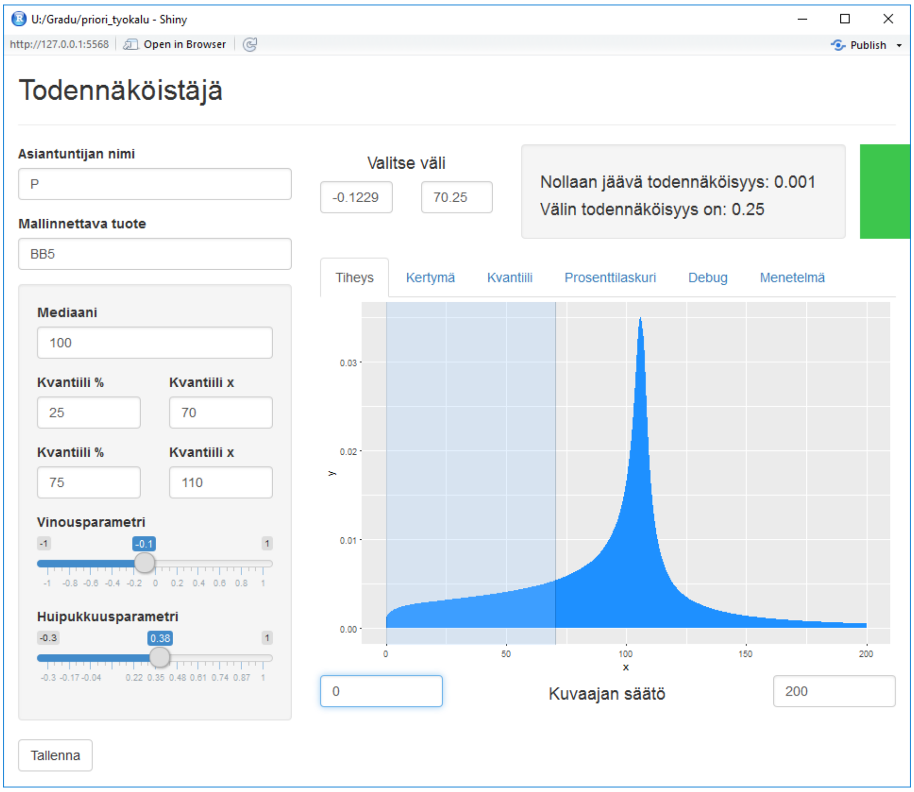
Työkaluun kerättyjen tunnuslukujen selitteet ilmenevät taulukosta 1. Taulukon merkinnät vastaavat yhtälöryhmän (14) merkintöjä. Mediaania x_2 vastaava todennäköisyyskertymä u_2 on 0.5.

Taulukko 1: Asiantuntijahaastatteluissa kerätyt tunnusluvut, niiden nimitykset työkalussa ja selitteet nimityksille

Tunnusluku	Nimitys työkalussa	Selite
x_1	Kvantiili x	Jakauman kvantiili
x_2, Md	Mediaani	Jakauman mediaani
x_3	Kvantiili x	Jakauman kvantiili
u_1	Kvantiili %	Kvantiilipisteen x_1 kertymä
u_3	Kvantiili %	Kvantiilipisteen x_3 kertymä
τ_3	Vinousparametri	Jakauman L-vinous
τ_4	Huipukkuusparametri	Jakauman L-huipukkuus

5.1 Todennäköistäjän käyttöliittymä

Työkalun käyttöliittymä on web-pohjainen ja se on toteutettu käyttäen shiny-pakettia (Chang *et al.*, 2017). Paketti toimii rajapintana HTML-elementteihin, joilla on myös toiminnallisuuksia. Ulkoasuun voidaan vaikuttaa CSS-tyylimäärittelyillä. Toiminnallisuudet puolestaan perustuvat JavaScriptiin. Työkalun ulkoasu on esiteltu kuvassa 2.

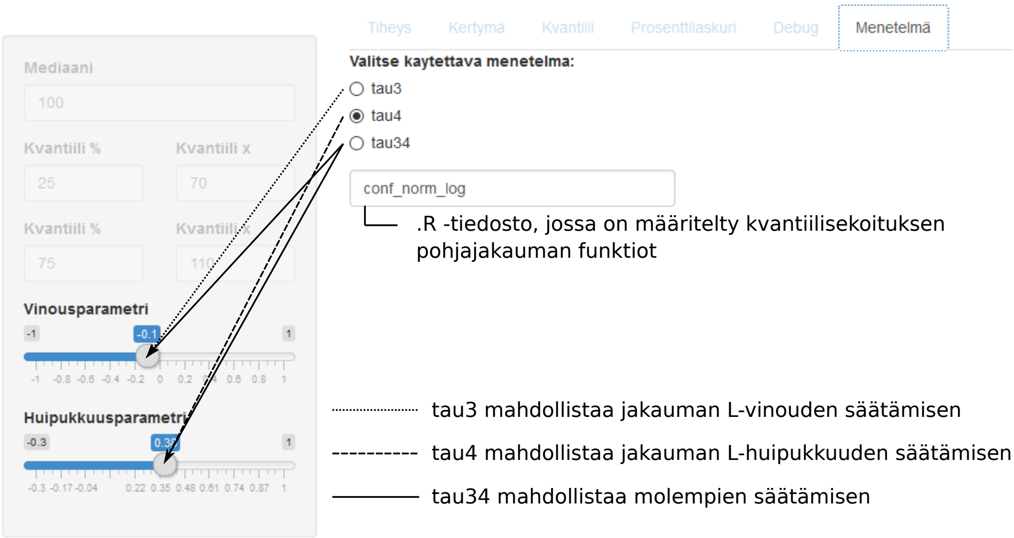


Kuva 2: Työkalun käyttöliittymä

Ulkoasu jakaantuu useampaan eri näkymään, jotka löytyvät välilehtien takaa. Seuraavaksi esitetään yksityiskohtaisesti kuvia hyödyntäen työkalun toiminallisuudet eri näkymissä. Kuvissa on huomattava, että epäoleellisia komponentteja on häivytetty kuvista, jotta esiteltävät ominaisuudet korostuisivat. Oikeassa käyttötilanteessa näin ei ole. Työkalun käytön ohjeistuksessa kuvataan työkalun käyttöä asiantuntijahaastatteluissa.

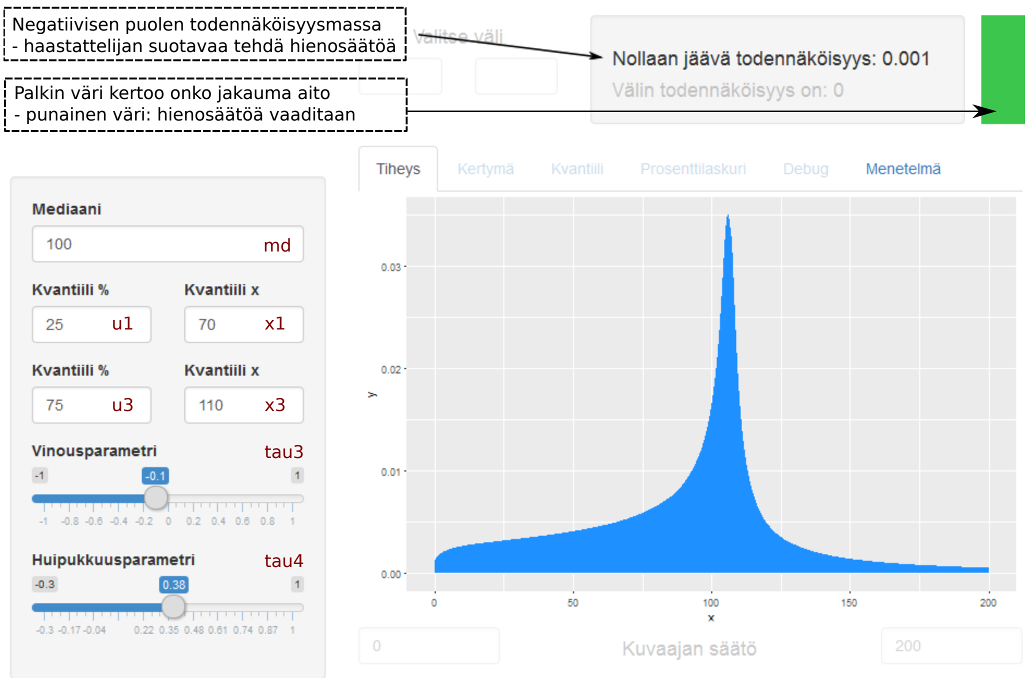
Työkalun käyttäminen alkaa mallinnettavan asian tietojen kirjaamisella. Tähän on varattu kaksi kenttää: *Asiantuntijan nimi* ja *Mallinnettava tuote*. Ne määrittävät myös tiedoston nimen, johon työkalussa olevat tiedot tallentuvat. Seuraavaksi käyttäjä eli haastattelija määrittää jakauman sovittamiseen käytettävän menetelmän ja pohjajakauman. Nämä toiminnallisuudet löytyvät *Menetelmä* -välilehdeltä ja ne näkyvät kuvassa 3.

Jakauman sovittamisessa on mahdollista hyödyntää jakauman L-vinoutta, L-huipukkuutta tai molempia. Lisäksi polynomisen kvantiilisekoituksen pohjajakaumaa on mahdollista vaihtaa. Vaihtaminen tapahtuu määritystiedoston avulla, jossa on R-koodina nimetty pohjajakauman funktiot. Pohjajakauman määrittämiseksi on nimettävä kertymä- ja kvantiilifunktio. Oletusasetuksena pohjajakaumana on Lognormaalijakauma.



Kuva 3: Menetelmä -välilehden toiminnallisuudet

Alkusäätöjen jälkeen todennäköistäminen aloitetaan selvittämällä asiantuntijalta kvantiilit. Kvantiileille on omat syötekentät, kuten kuvasta 4 ilmenee. Näillä tiedoilla Todennäköistäjä sovittaa lähes reaaliaikaisesti jakauman. Jakauman hienosäätö tapahtuu liukusäätimien avulla, riippuen aiemmin valitusta menetelmästä. Tarvittaessa menetelmää voidaan vaihtaa haastattelun missä tahansa vaiheessa. Jakaumasta nähdään sovittamisen myötä tiheys- kertymä- ja kvantiilifunktio, joita tarkastellen jakaumaa voidaan hienosäätää.

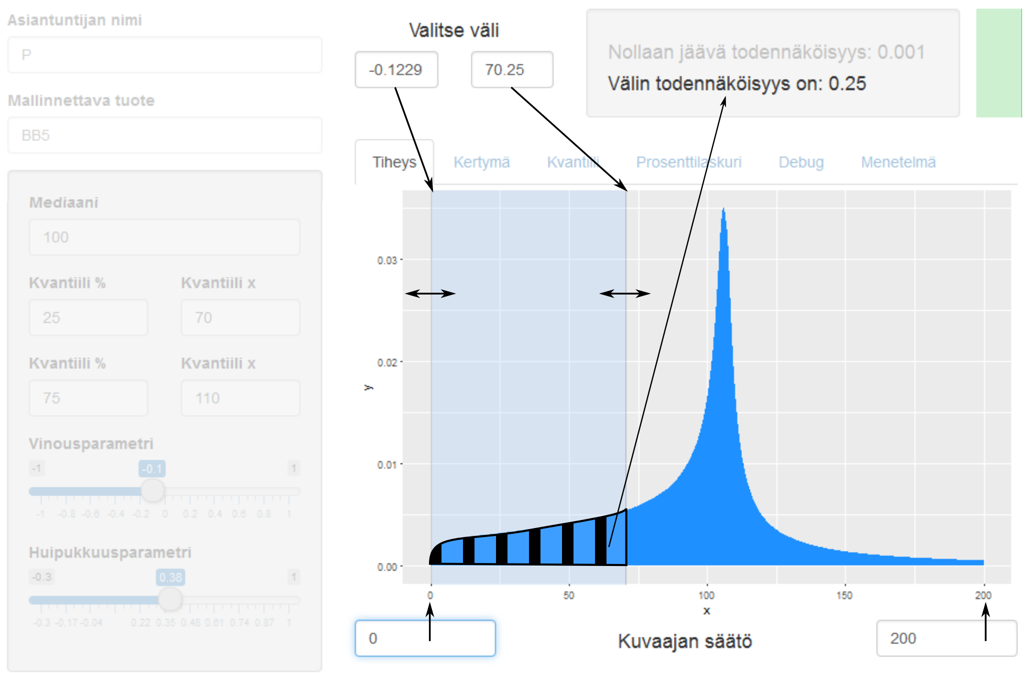


Kuva 4: *Vasemmalla:* jakauman sovittamiseen vaikuttavat säädöt, *oikealla:* sovitettu jakauma, *ylhäällä:* ohjeistus haastattelijalle jakauman hienosäätämiseksi

Hienosäätöön ja jakauman tarkasteluun on lukuisia eri ominaisuuksia. Kuvassa 4 nähdään indikaattoripalkki, joka kertoo käyttäjälle, muodostuuko sovitettujen malliparametrien myötä aito kvantiilifunktio. Lopullisen jakauman on oltava ehdottomasti aito todennäköisyysjakauma, jotta se on käyttökelpoinen. Hienosäätöä on siis tehtävä, kunnes palkki saadaan vihreäksi, jolloin sovitettu kvantiilifunktio on kasvava. Muussa tapauksessa voidaan joutua toteamaan, että asiantuntijan antamilla tiedoilla ei ollut mahdollista sovittaa mallia. Tällöin joko pohjajakauma ei sovi annettuun ilmiöön, tai asiantuntijan antamissa tiedoissa esiintyi ristiriitaisuuksia.

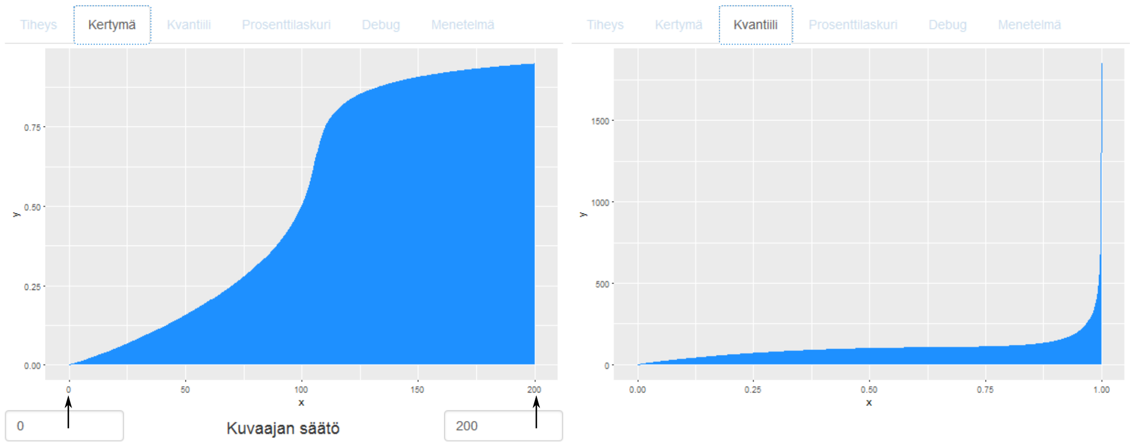
Indikaattoripalkin vieressä kerrotaan myös selkokielisesti, mikä on negatiiviselle puolelle jäävä todennäköisyysmassa. Tämä auttaa jakauman hienosäätämisessä, kun oletetaan tuotteiden tilausten saavan vain positiivisia arvoja. Tavoitteena olisi sovittaa jakauma siten, ettei nollan alapuolella todennäköisyysmassaa jäisi ollenkaan. Käytännössä tämä ei kaikissa tilanteissa onnistu, joten tälle todennäköisyydelle on tehtävä jokin tulkinta. Esimerkiksi jakaumaa voidaan jälkikäsitellä katkaisemalla jakauma ja skaalaamalla jäljelle jäänyt todennäköisyys.

Sovitetun jakauman todennäkäisyyksiä on mahdollista tarkastella maalaustyökalulla, jolla voi valita tiheysfunktiosta välin. Työkalu ilmoittaa valitun välin todennäköisyyden. Toiminnallisuus on havainnollistettu kuvassa 5. Esimerkiksi voidaan miettiä, mikä olisi todennäköisyys, että ensimmäisen kvartaalin ennustettu tilausmäärä olisi jotain väliltä $0 - 70$. Tällöin se voidaan laskea syöttämällä arvot tekstikenttiin, tai vaihtoehtoisesti piirtämällä alue kursorilla kuvaajaan. Saatuun todennäköisyyteen asiantuntijalta kysytään kommentit. Kommenttien perusteella jakaumaa hienosäädetään ja tarvittaessa myös kvantiileja muutetaan.



Kuva 5: Tiheysfunktion pinta-alan määrittäminen annetulla välillä, alareunassa kuvaajan x-akselin säädöt

Lisäksi sovitetun jakauman tarkasteluun voidaan hyödyntää kertymä- ja kvantiilifunktioita. Näille on omat näkymänsä erillisillä välilehdillä. Välilehdet on esitetty kuvassa 6. Tiheys- ja kertymäfunktion piirtoaluetta on mahdollista säätää asettamalla piirtovälin päätepisteet kuvaajan alapuolella oleviin syötekenttiin.

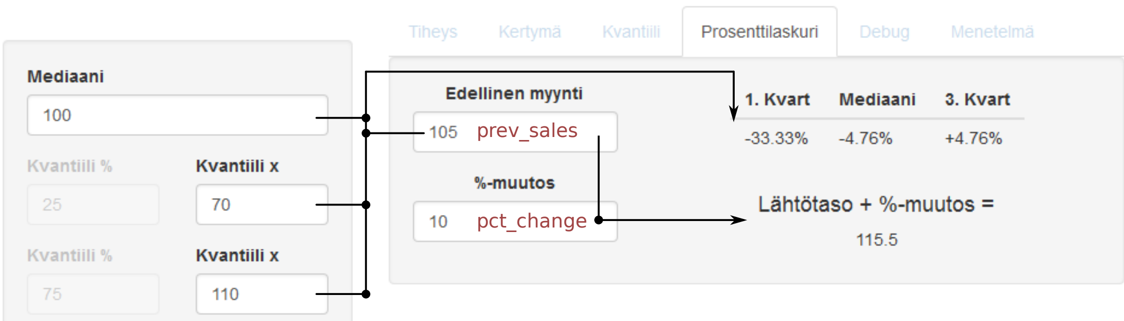


(a) Kertymäfunktion kuvaaja

(b) Kvantiilifunktion kuvaaja

Kuva 6: Sovitetun jakauman kertymä- ja kvantiilifunktioiden tarkastelu välilehdillä

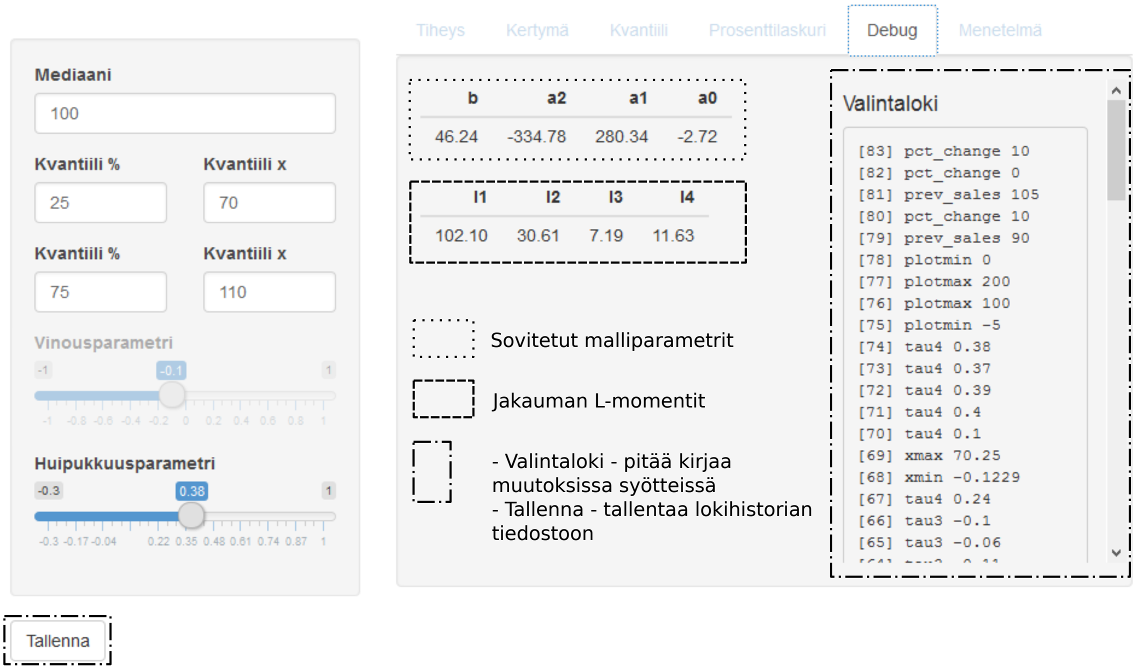
Mikäli kvantiilien määrityksessä on haasteita, voidaan hyödyntää aiemman vastaavan ajankohdan tilausmääriä apuna. Asiantuntija voi ajatella kysynnässä tapahtuvia muutoksia prosentuaalisesti. Tätä varten työkalussa on mukana yksinkertainen prosenttilaskuri, joka on esitetty kuvassa 7. Asiantuntija voi ilmaista ennustetut tilausmäärän kvantiilit prosentuaalisina muutoksina, jolloin haastattelijan on johdettava kvantiilit tästä tiedosta. Prosenttilaskuri mahdollistaa asetettujen kvantiilien vertaamisen prosentuaalisesti aikaisempaan tilausmäärään. Lisäksi voidaan laskea paljonko ennustettu prosentuaalinen muutos vastaisi kappalemääränä.



Kuva 7: Työkalun prosenttilaskuri, nuolet osoittavat laskemistavan

Jakauman sovittamisen jälkeen haastattelija voi tarkastella polynomisen kvantiilisekoituksen malliparametreja *Debug* -välilehdeltä. Samalta välilehdeltä löytyvät myös jakauman L-momentit. Oikealla kuvassa 8 nähdään valintaloki, jossa on tiivistetysti esitettynä kaikki haastattelijan tekemät muutokset työkalussa. Valintalokissa on tiedot kuhunkin syötekenttään tehdyistä muutoksista tietoineen ja aikaleimoineen. Tämä mahdollistaa tulosten toistettavuuden kaikissa haastattelun vaiheissa. Työkalusta löytyy lisäksi *Tallenna* -painike, joka tallentaa valintalokin tiedostoon.

Tällöin viimeisimmistä syötteiden arvoista on johdettavissa asiantuntijahaastattelun tulos. Virheiden varalta työkalu tallettaa myös suljettaessa oman tiedoston.



Kuva 8: Debug -välilehden ominaisuudet

5.2 Esimerkki työkalun käyttötilanteesta

Työkalun todellinen käyttötilanne on mutkikkaampi kuin pelkistetty kuvaus työkalun ominaisuuksista. Kattavampi kuva haastattelutilanteesta ja työkalun käytöstä saadaan tutkimalla työkalun lokitiedostoa. Seuraavaksi esitellään esimerkkinä yhden tuoteperheen asiantuntijahaastattelu lokitiedoston avulla. Lokitietojen ohella kerrotaan haastattelun kulusta.

Haastattelutilanteessa haastattelija käy ensimmäisenä läpi haastattelun kulun ja tavoitteen. Asiantuntijan kanssa käydään läpi hänen saamansa ennakkomateriaalit. Asiantuntijalta kysytään lähestymistapoja ja tekijöitä, joiden hän arvelee vaikuttavan tilausmääriin. Haastattelija ohjeistaa pohtimaan huolellisesti näiden tekijöiden kautta kokonaisarviota sapuuville tilauksille. Lisäksi asiantuntijan kanssa keskustellaan hänen työtehtävästään ja tilausten ennustamisesta yleisesti. Taustojen kartoituksen jälkeen aloitetaan varsinainen asiantuntijahaastattelu, joka tehdään kullekin tuoteperheelle erikseen.

Taulukko 2 on esimerkki eräästä haastattelutilanteesta. Taulukko on muokattu versio lokitiedoista ja jakauman tarkasteluita on karsittu niiden suuren määrän vuoksi. Jakauman hienosäätöä on siistitty syötekentän *tau4* osalta. Aika on kuvattu työkalun käynnistyshetken suhteen. Työkalun syötekenttien nimet on esitelty taulukossa 4 ja kuvissa 3, 4 ja 7.

Työkalu käynnistetään uuden tuotteen haastattelun alkaessa. Kohdat 1–3 ovat työkalun oletusasetukset. Haastattelija täyttää asiantuntijan nimen, mallinnettavan tuoteperheen ja aiemman tilausmäärän kohdissa 4–6. Puolitusmenetelmän mukaisesti selvitetään jakauman kvartiilit kohdissa 7–9. Asiantuntija tarkastelee aiempia

Taulukko 2: Asiantuntijahaastattelun kulku haastattelutyökalun lokitietojen mukaan

	aika (mm:ss)	työkalun syötekenttä	syötetty arvo		aika (mm:ss)	työkalun syötekenttä	syötetty arvo
1	00:00	u1	0.25	22	07:25	tau4	0.12
2	00:00	u3	0.75	23	07:31	plotmin	100
3	00:00	tau4	0.1	24	08:06	xmin	220.9
4	00:01	asiantuntija	j	25	08:09	xmax	241
5	00:05	tuotenimi	BB6	26	08:20	x3	215
6	01:10	prev_sales	150	27	08:26	x3	219
7	01:31	md	210	28	08:39	tau4	0.08
8	02:25	x1	160	29	08:44	tau4	0.04
9	03:36	x3	220	30	08:46	tau4	0.1
10	03:59	menetelma	tau4	31	09:12	xmin	210.6
11	03:52	plotmin	50	32	09:14	xmax	248.5
12	04:06	plotmax	250	33	09:47	x3	210
13	04:23	xmax	218.4	34	10:22	md	190
14	04:24	xmin	67.03	35	11:18	xmin	189.1
15	04:56	tau4	0.3	36	11:31	xmax	244.3
16	06:30	tau4	0.39	37	14:10	xmax	230.3
17	06:37	xmin	209.1	38	14:18	xmin	145.2
18	06:37	xmax	214.7	39	15:58	pct_change	-20
19	07:00	md	205	40	16:04	x1	160
20	07:09	x1	170	41	16:43	xmin	116
21	07:17	tau4	0.33	42	16:59	xmax	125.1

tilausmääriä ja miettii mihin suuntaan tuoteperheen kysyntä on muuttumassa. Lisäksi otetaan mm. huomioon asiakkaat, tuotteen kehitys ja käyttökohteet. Epävarmuus tuodaan ennusteeseen mukaan taustatekijöihin liittyvästä epävarmuudesta. Asiantuntija on selvästi sitä mieltä, että yläkvartiilin ja mediaanin väli on kapea. Hienosäädön menetelmäksi valitaan huipukkuuden säätäminen ja kuvaaja asetetaan kohdilleen (10–12). Tässä kohtaa jakauma on sovittunut ensimmäisen kerran ja siitä nähdään työkalussa tiheysfunktio. Haastattelussa on kulunut neljä minuuttia ja on aika siirtyä jakauman hienosäätämiseen.

Jakauman hienosäätämiseksi jakaumasta valitaan väli (13, 14). Väli on jakauman vasen häntä, jonka todennäköisyysmassasta asiantuntijalta kysytään kommentti. Jakauman huipukkuutta säätämällä (15, 16) ja toisen välin tarkastelun (17, 18) perusteella jakauman alakvartiilia ja mediaania korjataan (19, 20). Asiantuntija pienentää ennustettaan alkuperäisestä ja samalla tarkentaa arviotaan (kaventaa jakaumaa). Vasempaan häntään ollaan toistaiseksi tyytyväisiä.

Jakauman yläpäätä tutkitaan seuraavaksi ja yläkvartiilia korjataan hieman (26, 27). Huipukkuutta säädetään edelleen jakaumasta (28–30), jonka jälkeen yläkvartiili saadaan asetettua lopullisesti paikalleen (33). Samalla tehdään myös viimeinen korjaus mediaaniin (34). Alakvartiilin korjaamiseksi tarkatellaan vielä todennäköisyyskertymiä kahdella eri välillä (35–38). Prosenttilaskinta (39) hyödnnetään pahimman skenaarion laskemiseksi ja alakvartiili (40) asetetaan takaisin ensimmäiseen arvioon. Vielä tehdään yksi välin tarkastelu (41–42) jakaumalle. Tämän jälkeen haastattelija on varmistunut siitä, että mallinnettu jakauma kuvastaa asiantuntijan arviota. Haastatteluun kului yli 17 minuuttia tämän tuoteperheen osalta. Työkalusta painetaan *Tallenna*-painiketta, jolloin lokitiedosto tallentuu. Yhden tuoteperheen osalta asiantuntijahaastattelu on ohi ja työkalu suljetaan.

Työkalun syötearvoissa tapahtuneet muutokset ilmenevät tarkemmin taulukosta 3. Taulukon rivien indeksit vastaavat täydellisten lokitietojen arvoja. Taulukon arvot pysyvät samana ylhäältä alaspäin luettaessa, kunnes muutos tapahtuu. Lokitietojen avulla on myös mahdollista toisintaa jakauman muuttuminen ajan suhteen. Vaakaviiva rivien 9 ja 15 välissä kertoo ajanhetkestä, jolloin jakauma on saatu sovitettua ensimmäisen kerran asiantuntijan määrittämiin tunnuslukuihin.

Taulukko 3: Asiantuntijahaastattelun kulku syötearvojen muutoksina kuvattuna

	aika (mm:ss)	x1	md	x3	tau4
3	00:00				0.1
7	01:31		210		
8	02:25	160			
9	03:36			220	
15	04:56				0.3
16	06:30				0.39
19	07:00		205		
20	07:09	170			
21	07:17				0.33
22	07:25				0.12
26	08:20			215	
27	08:26			219	
28	08:39				0.08
29	08:44				0.04
30	08:46				0.1
33	09:47			210	
34	10:22		190		
40	16:04	160			

Taulukon 3 avulla haastattelun kulku hahmottuu selkeämmin. Ajallisesti tarkasteltuna pitempien taukojen kohdalla jakauman muotoa on mietitty tarkemmin ja haastattelija on esittänyt tarkentavia kysymyksiä. On huomattava, että taulukosta

2 on karsittu peräkkäisiä muutoksia samassa syötearvossa esimerkin selkiyttämiseksi. Syötearvoa saatetaan korjata pienin askelin kunnes esimerkiksi valitun välin todennäköisyys saadaan vastaamaan asiantuntijan määräämää arvoa. Tällöin tarpeettomia askelia kertyy lokitietoihin runsaasti.

Esimerkkitapauksessa asiantuntija pohti pitkään ennustettaan ja kokonaiskuvan muodostaminen vei aikaa. Haastattelun kesto oli pisimmästä päästä. Joidenkin tuoteperheiden ja asiantuntijoiden kohdalla kvartiilit määritettiin suoraan paikoilleen, jolloin niiden määrittämiseen käytettiin alussa enemmän aikaa. Tämän jälkeen hienosäätöön kului vähemmän askelia.

6 Asiantuntija-arvioiden yhdistäminen

Tuoteperheen tilausmäärän ennusteen määrittäminen edellyttää useamman eri asiantuntijan arvion yhdistämistä. Tavoitteena on saada tiivistettyä useammalta eri asiantuntijalta saatava informaatio yhdeksi käyttökelpoiseksi ennusteeksi. Yhdistämisessä on tavoitteena ottaa myös huomioon asiantuntijoiden epävarmuus tulevasta tilausmäärästä. Asiantuntijahaastattelun tuloksena saadaan tuoteperheen tilausmäärälle ennustejakauma. Luonnollinen tapa lähestyä ongelmaa on siis tarkastella todennäköisyysjakaumien yhdistämistä.

Asiantuntijoiden mielipiteiden yhdistämisestä käytetään englannin kielessä termiä *opinion pool*, joka esiintyy mahdollisesti ensimmäisen kerran artikkelissa (Stone *et al.*, 1961). Yleisimpiä menetelmiä ovat lineaarinen ja logaritminen jakaumien yhdiste (Clemen ja Winkler, 1999). Tämän kaltainen jaottelu perustuu jakaumien painotusmenetelmiin. Tässä työssä perehdytään lineaarisiin painotusmenetelmiin.

Menetelmät voidaan myös jaotella sen mukaan, miten jakaumaa hyödynnetään yhdisteessä. Tyypillisesti jakaumien yhdistäminen toteutetaan tiheysfunktioiden avulla. Tiheysfunktiota yhdistettäessä tulos ei yleensä ole samasta jakaumaperheestä kuin yksittäiset jakaumat. Tähän esitetään ratkaisuksi polynomisen kvantiilisekoituksen sovittamista tiheysfunktioiden yhdisteeseen. Vaihtoehtoinen menetelmä on kvantiilifunktioiden yhdistäminen (Lichtendahl Jr. *et al.*, 2013). Kvantiilifunktioiden yhdistäminen vastaa luvussa 4 esiteltyjä kvantiilisekoituksia.

Seuraavaksi käsitellään asiantuntijoiden välistä painotusta, joka otetaan huomioon jakaumien yhdistämisessä. Tämän jälkeen esitellään varsinaiset yhdistämismenetelmät. Kolme tutkittavaa menetelmää jakaumien yhdistämiseksi ovat:

- tiheysfunktioiden keskiarvoistus
- kvantiilisekoituksen sovittaminen
- kvantiilifunktioiden keskiarvoistus.

6.1 Asiantuntijoiden painotus jakaumien yhdistämisessä

Todennäköisyysjakaumien yhdistämisessä on mahdollista huomioida asiantuntijoiden väliset erot. Eroilla tarkoitetaan eroavaisuuksia siinä, miten ja mitä asioita asiantuntija ottaa huomioon mallinnettaessa ennustetta tilausmäärälle. Tilausmääriä ennustettaessa yrityksen eri asiantuntijat työskentelevät eri osa-alueilla ja siten heillä on myös erilaisia lähestymistapoja ongelmaan. Ennusteiden välillä voi olla myös eroja siinä, miten kattavasti asiantuntija huomioi erilaisia selittäviä tekijöitä. Painotus otetaan kussakin varsinaisessa jakaumien yhdistämismenetelmässä huomioon.

Jakaumat voidaan yhdistää monella tavalla, mutta menetelmästä riippumatta tarkoituksena on huomioida asiantuntijoiden erot painotuksien avulla. Painoille käytetään merkintää w_j, missä indeksi j viittaa asiantuntijaan. Painojen summaksi asetetaan $\sum_j w_j = 1$. Tällöin kunkin asiantuntijan paino kuvastaa osuutta, jolla se vaikuttaa yhdistettyyn ennustejakaumaan. Painotuksen tarkempi tulkinta riippuu käytettävästä menetelmästä.

Painoarvojen valintaan käytetään haastatteluissa kertynyttä tietoa. Perusteena ovat esimerkiksi asiantuntijan työn vastuualue ja kuinka laajasti he kertovat huomioivansa eri asioita ennusteissaan. Myös tieto aikaisemmasta tilausmäärien seurannasta huomioidaan. Suuria eroja asiantuntijoiden välille ei saada johdettua.

Asiantuntijahaastattelu toteutettiin neljälle eri henkilölle *J, P, R* ja *K*. Painotusta ei aseteta erikseen eri tuoteperheille, koska eroja ei voida muutenkaan arvioida suuriksi. Asiantuntijat laitetaan paremmuusjärjestykseen perustuen heidän kokonaisvaltaiseen tietämykseen tilauksista. Annetaan painoarvoiksi $w_J = 0.29$, $w_P = 0.26$, $w_R = 0.24$ ja $w_K = 0.21$. Painotusta voidaan kuvata melko tasaiseksi. Seuraavaksi käsitellään jakaumien yhdistämistä. Näemme myös miten nämä painot itse asiassa vaikuttavat yhdistettyyn jakaumaan.

6.2 Tiheysfunktioiden keskiarvoistus

Tiheysfunktioiden lineaarinen keskiarvoistus on tyypillisin jakaumien yhdistämiseen käytetty menetelmä (Clemen ja Winkler, 1999). Lineaarinen jakaumien yhdiste saadaan ottamalla painotettu keskiarvo tiheysfunktioista pisteittäin. Yhdistetyn jakauman tiheysfunktio esitetään muodossa:

$$f(y) = \sum_j w_j f_j(y), \tag{15}$$

missä f_j on asiantuntijan j määrittämän jakauman tiheysfunktio.

Lineaarisen tiheysfunktioiden yhdistämisen tuloksena syntyy aina leveämpi jakauma, kuin mitä yksittäiset asiantuntijat ovat arvioineet (Lichtendahl Jr. *et al.*, 2013). Lineaarinen yhdiste voi mahdollisesti olla myös monihuippuinen, mikäli asiantuntijoiden arviot poikkeavat selkeästi toisistaan. Monihuippuisuudelle on annettava kuitenkin tulkinta, jotta yhdistetty jakauma on mielekäs ennusteena. Yhdistetty tiheysfunktio ei pääsääntöisesti tuota helposti tulkittavaa parametrista jakaumaa. Tämän ratkaisemiseksi toisena tutkittavana menetelmänä on kvantiilisekoituksen sovittaminen tiheysfunktioiden yhdisteeseen.

6.3 Kvantiilisekoituksen sovittaminen yhdistettyyn tiheysfunktioon

Tiheysfunktioiden keskiarvoistuksen avulla yhdistetty jakauma ei ole samaa jakaumaperhettä kuin alkuperäiset jakaumat. Seuraavaksi tutkitaan, miten yhdistettyä tiheysfunktiota voidaan approksimoida polynomisella kvantiilisekoituksella. Polynomiselle kvantiilisekoitukselle jakauman parametrit ovat laskettavissa suoraan jakauman L-momenteista. L-momentit puolestaan ovat helposti laskettavissa empiirisestä jakaumasta. Empiirinen yhdistetty jakauma tuotetaan simulointien avulla.

Polynomisen kvantiilisekoituksen tapauksessa voidaan helposti simuloida yksittäisistä jakaumista havaintoja, kuten luvussa 4 osoitettiin. Oletetaan, että yhdistetty jakauma on muodostuu tiheysfunktioiden painotettuna keskiarvona. Tätä yhdistettyä jakaumaa voidaan simuloida yksittäisten jakaumien avulla seuraavasti:

1. olkoon n valittu otoskoko
2. kullekin jakaumalle annetaan painot w_j, $\sum w_j = 1$
3. simuloidaan kustakin jakaumasta f_j otos $Y^{(j)}$, kokoa $n \times w_j$
4. yhdistetään simuloidut otokset, $Y = \{Y^{(J)}, Y^{(P)}, Y^{(R)}, Y^{(K)}\}$.

Yhdistetty otos on peräisin jakaumasta, johon halutaan sovittaa L-momenttien avulla polynomisen kvantiilisekoituksen parametrit.

Otokselle Y asteen r L-momentti voidaan laskea seuraavasti:

$$l_r = \frac{1}{r}\binom{n}{r}^{-1} \sum_{i=1}^{n} \left[\left(\sum_{j=0}^{r-1} (-1)^j \binom{r-1}{j} \binom{i-1}{r-1-j} \binom{n-i}{j} \right) Y_{(i)} \right], \tag{16}$$

missä $Y_{(i)}$ on järjestetyn otoksen alkio paikassa i (Elamir ja Seheult, 2003). L-momenttien laskeminen otokselle on toteutettu R-paketissa *Lmoments* (Karvanen, 2016).

Otoksesta estimoidaan neljä ensimmäistä L-momenttia, sillä yksittäisten jakaumien sovittamisessa on myös käytössä neljä parametria. Yhdistetyn jakauman malliparametrit saadaan lineaarisen yhtälöryhmän ratkaisuna, joka on johdettavissa luvun 4 kaavasta (11). Yhtälöryhmä perustuu tietoon yhdistetyn jakauman ja polynomisen kvantiilisekoituksen komponenttien L-momenteista. Tuntemattomina muuttujina ovat yhdistetyn jakauman malliparametrit. Tällöin tuloksena on yhdistetty ennustejakauma, joka on samaa jakaumaperhettä kuin yksittäiset asiantuntijoiden määrittämät jakaumat. Menetelmän keskeinen ajatus on jakauman approksimointi polynomisella kvantiilisekoituksella.

6.4 Kvantiilifunktioiden keskiarvoistus

Useampi ennustejakauma voidaan vaihtoehtoisesti yhdistää keskiarvoistamalla yksittäisten jakaumien kvantiilifunktiot (Lichtendahl Jr. *et al.*, 2013). Painotetun keskiarvon ottaminen kvantiilifunktioista vastaa kvantiilisekoitusta. Osoitetaan, että saman jakaumaperheen jakaumille menetelmä vastaa jakaumien L-momenttien painottamista. L-momenttien avulla menetelmälle saadaan intuitiivinen tulkinta.

Todetaan ennen tuloksien esittelyä, että kaikissa haastatteluissa päädyttiin hyödyntämään jakauman sovittamisessa kvantiileja ja jakauman huipukkuutta. Täten malliparametreja on sama määrä kunkin asiantuntijan määräämässä ennustejakaumassa, jolloin ne ovat samaa jakaumaperhettä. Lasketaan yhdistetty ennustejakauma painotettuna keskiarvona kvantiilifunktioista. Kullekin tuoteperheelle ennustejakauma määräytyy seuraavasti:

$$\begin{aligned} Q_{Yhd}(u) &= \sum_j w_j \, Q_j(u) = \sum_j w_j \left(b_j Q_0(u) + a_j u^2 + a_{1j} u + a_{0j} \right) \\ &= \beta Q_0(u) + \alpha_2 u^2 + \alpha_1 u + \alpha_0, \end{aligned} \tag{17}$$

missä

$$\beta = \sum_j w_j b_j \qquad \alpha_2 = \sum_j w_j a_{2j} \qquad \alpha_1 = \sum_j w_j a_{1j} \qquad \alpha_0 = \sum_j w_j a_{0j}.$$

Indeksointi j viittaa asiantuntijaan. Menetelmän etuna on, että yhdistetty jakauma on samaa muotoa kuin asiantuntijahaastatteluissa käytetty jakaumaperhe. Yhdistetyn jakauman malliparametrit ovat painotettuja keskiarvoja yksittäisten jakaumien malliparametreista. Tiheysfunktio yhdistetylle jakaumalle on johdettavissa kuten luvussa 4.

Huomataan myös, että yhdistetyn jakauman (17) L-momentit määräytyvät yksittäisten jakaumien vastaavien tunnuslukujen painotettuna keskiarvona. Tämä ominaisuus seuraa kaavasta (8):

$$\lambda_r(Q_{Yhd}) = \sum_j w_j \lambda_r(Q_j).$$

Asiantuntijat määrittävät siis kollektiivisesti jakauman muodon. Tällöin esimerkiksi jakauman sijainti asettuu keskimäärin heidän arvioihin. Yhdisetyn jakauman skaala on myös keskimäärin yksittäisten jakaumien skaala. Tämä tarkoittaa sitä, että jakauman leveys ei kasva, vaikka yksittäisten jakaumien sijainnit poikkeaisivat suurestikin toisistaan. Vinouden ja huipukkuuden kohdalla tulkinnat ovat vastaavia.

6.5 Asiantuntijahaastatteluiden tulosten arviointimenetelmät

Tähän mennessä huomion keskipisteessä on ollut koko asiantuntijahaastatteluiden toteutus ja tulosten tiivistäminen ennustejakaumiksi. Asiantuntijahaastattelun onnistumisen arvioimisessa tarkastellaan kaikkia käytettyjä menetelmiä ja toteutusvaiheita. Ennen tuloksia käydään lyhyesti läpi mitkä asiat vaikuttavat tuloksiin ja miten niiden vaikutuksia voidaan arvioida.

Haastattelumenetelmiä arvioidaan sen perusteella, miten hyvin asiantuntijat pystyivät omaksumaan ne. Työkalun sopivuutta haastatteluihin tarkastellaan käytettävyyden, jakauman sovittamisen ja hienosäädön kannalta. Asiantuntijoille tarjottujen ennakkotietojen vaikutusta ennustejakaumaan arvioidaan heidän antamien kommenttien perusteella.

Asiantuntijoiden sovittamia jakaumia vertaillaan toteutuneeseen tilausmäärään. Black Bruinin toteutuneet tilausmäärien lukemat eri tuoteperheille saatiin huhtikuun alussa. Ensimmäisenä kohtana on tutkia ovatko asiantuntijoiden mallintamat ennustejakaumat osuneet lähelle toteutuneita arvoja. Poikkeavien tulosten kohdalla on selvitettävä, miksi asiantuntijan ennuste ei ole osunut kohdalleen.

Mallinnettujen ennustejakaumien tuloksia tarkastellaan kolmella eri tasolla:

- yksittäiset ennustejakaumat
- vertailu asiantuntijoittain ja tuoteperheittäin
- kokonaisvaltainen ennustamiskyky.

Käytettävinä menetelminä ovat tiheysfunktion graafinen tarkastelu, kertymäfunktion arvo toteutuneella tilausmäärällä ja mediaaniestimaatin vertaaminen toteutuneeseen tilausmäärään. Näillä menetelmillä saadaan kokonaisvaltainen käsitys asiantuntijahaastattelun onnistumisesta ja käytettävyydestä tilausmäärän ennustamisessa.

7 Asiantuntijahaastatteluiden tulokset

Työkalun toimivuus ja polynomisten kvantiilisekoitusten sopivuus tilausmäärän ennustejakauman mallintamisessa voidaan todeta vain käytännössä. Tässä luvussa esitellään tulokset haastatteluista jakaumineen.

7.1 Asiantuntijat Black Bruinilla

Yrityksen kanssa käytyjen keskustelujen perusteella asiantuntijoiksi valittiin henkilöitä, jotka ovat aktiivisesti tilausten kanssa tekemisissä. Näihin lukeutuivat saapuvia tilauksia hoitavat henkilöt ja osa myynti- ja markkinointihenkilöstöä. Valitut asiantuntijat saivat itselleen ennakkoon tietoon aikaisemman vuoden osalta kvartaaleittain tilausmäärien lukemat. He saivat myös ohjeistuksen haastattelutilanteesta ja kysymyksistä noin viikkoa ennen sovittuja haastattelupäivämääriä.

Ensimmäisenä vuorossa olleet kaksi tilaustenkäsittelijää totesivat haastattelutilanteessa, etteivät he pysty antamaan mitään arvioita tulevien tilausten lukumääristä. Syyksi he arvioivat, että heidän työnsä on melko mekaanista, eikä asiakkailta välity heidän suuntaan informaatiota tulevista tilauksista. He kertoivat myös, etteivät ole seuranneet tilausten lukumääriä aikaisemmin. Esitietoina saadut tilausmäärien lukemat olivat siis jokseenkin uusia heille. Tämän seurauksena todennäköistämistä ei heidän osaltaan toteutettu. Mahdollisina tuloksina heidän haastattelustaan olisi luultavasti saatu leveitä jakaumia, jotka sijoittuvat täsmälleen aiemman tilausmäärän ympärille.

Tilaustenkäsittelijöiden tilalle valikoitiin kaksi muuta henkilöä haastateltavaksi. Neljän valitun asiantuntijan osalta todennäköistäminen saatiin suoritettua ilman ongelmia. Henkilöt J, P, K ja R muodostivat kukin näkemyksensä viiden eri tuoteperheen hydraulimoottoreiden tulevista tilausmääristä. Haastattelut toteutettiin erikseen kullekin asiantuntijalle. Kaikissa haastatteluissa oli mukana Black Bruin Oy:n talousjohtaja Janne Mustonen. Tarkoituksena oli varmistaa kommunikoinnin sujuvuus minun, eli haastattelijan, ja asiantuntijan välillä. Mustonen ei juuri kuitenkaan puuttunut haastattelutilanteisiin, sillä keskustelu eteni sujuvasti jokaisen asiantuntijoiden kanssa. Tuoteperheet käytiin kaikkien asiantuntijoiden kanssa läpi samassa järjestyksessä: BB3, BB4, BB5, BB6 ja BB7.

Henkilön K osalta todennäköistäminen toteutettiin etänä siten, että puheyhteyden lisäksi jaoimme tietokoneen työpöytänäkymän hänelle. Asiantuntijalla K oli näkyvillä todennäköistämisen käyttöliittymä, jossa esitettyihin kuvaajiin ja tunnuslukuihin oli samat keskustelun ja analysoinnin mahdollisuudet kuin muillakin asiantuntijoilla. Muihin haastatteluihin verrattuna etänä kommunikointi ei poikennut informaation välittymisessä.

Valittujen asiantuntijoiden ymmärrys jakaumiin ja tilastollisiin malleihin oli vaihtelevaa. Kaikkien kanssa saavutettiin kuitenkin käsitys siitä, että mitä tilausmäärien ennustejakauma kuvastaa. Eroja havaittiin myös tilausmäärien seurantatavoissa. Henkilöt J ja P ovat aktiivisesti olleet yrityksessä toteuttamassa myynnille ennusteita. Heidän käsityksensä tilausmääristä vaikutti kokonaisvaltaisemmalta ja he ottivat useampia eri vaikuttavia seikkoja arvioissaan huomioon.

Henkilö R puolestaan seurasi tarkkaan aikaisempia lukuja ja arvio kehitystä muutosten kautta. Hänellä oli käsitys kausivaihtelusta ja tilausmäärien kehittymisestä, joihin hän myös nojasi arvionsa.

Henkilö K asetti arvionsa puhtaasti lukumääriin, jotka koosti yrityksen asiakkuuksien kautta yhteenlaskemalla. Hän arvioi siis asiakastasolla tapahtuvia muutoksia tilausmäärissä ja muodosti niistä kokonaiskuvan. Henkilön K rooli yrityksessä kattaa yhden asiakassektorin, mistä hänellä voisi olettaa olevan tarkasti tietoa. Toimintasektorinsa ulkopuolelta ei voida olettaa häneltä löytyvän tarkkaa tietoa.

Asiantuntijahaastatteluiden perusteella yrityksen asiantuntijat ottivat eri tavoin todennäköistämisessä huomioon tilauksiin vaikuttavia tekijöitä. Kokonaisuudessaan he huomioivat todennäköistämisessä kaikkia niitä tekijöitä, joihin heidän arvionsa toivottiinkin perustuvan. Tämän perusteella asiantuntijoiden määrittämät ennustejakaumat pitävät sisällään sellaista informaatiota, joka ei tilaushistoriassa näy. Tämä ei kuitenkaan vielä kerro mallinnuksen osuvuudesta, eikä vastaa kysymykseen, miten asiantuntijan informaatio näkyy sovitetussa ennustejakaumassa.

7.2 Työkalun toimivuus haastattelutilanteessa

Asiantuntijahaastatteluissa hyödynnettiin luvussa 5 esiteltyä todennäköistämistyökalua. Työkalu mahdollisti visuaalisen tarkastelun sovitettavasta ennustejakaumasta. Seuraavaksi tarkastellaan työkalun toimivuutta ja arvioidaan sen käyttökelpoisuutta osana todennäköistämistä.

Työkalu toimi käytössä luotettavasti. Riittävällä virheiden käsittelyllä huolehdittiin, ettei ajonaikaisia virheitä pääse syntymään. Näin ollen sovellus ei kaatunut haastatteluiden aikana. Kaatuminen tarkoittaa sovelluksen odottamatonta sulkeutumista virheestä johtuen. Kaatumisessa sovelluksen sisältämät tiedot eivät välttämättä tallentuisi, jolloin todennäköistäminen olisi aloitettava alusta kyseisen tuoteperheen kohdalla. Sovellus on myös mahdollista sulkea epähuomiossa. Tämä otettiin huomioon siten, että työkalu tallentaa tiedot myös sulkeutuessaan.

Asiantuntijan kannalta työkalu edisti kokonaisuuden hahmottamista. Kerätyt jakauman tunnusluvut olivat yksinkertaisia ja intuitiivisia puolitusmenetelmän ansiosta. Visuaalinen tarkastelu ja prosenttilaskuri olivat asiantuntijalle tärkeimmät ominaisuudet todennäköistämisessä. Niiden avulla asiantuntijan informaatio saatiin muutettua jakaumaksi. Asiantuntijan arvioita kvantiileista pystyttiin tarkentamaan työkalun ominaisuuksia hyödyntäen. Kvartiiliväliä ja jakaumien häntien tarkastelua käytettiin tarkentamisessa apuna.

Haastattelijalle työkalu oli erityisen tärkeä, sillä ilman sitä tarkentavia kysymyksiä olisi liki mahdotonta esittää. Visuaalinen ja jakauman häntäpäiden tarkastelu on tärkeää, koska haastattelijan on varmistuttava, että sovitettu jakauma sopii ilmiöön ja asiantuntijan puheisiin. Hienosäätö jakaumalle oli helppoa, sillä jakauma sovittui reaaliajassa. Kaikki toivotut ominaisuudet sovitetusta jakaumasta eivät eksaktisti täyttyneet, joten hienosäädöllä on rajansa.

Polynomiset kvantiilisekoitukset osoittautuivat erittäin käyttökelpoiseksi jakaumaperheen valinnaksi. Kvantiileihin pohjautuva jakauman määrittäminen mahdollisti jakauman eksaktin määräämisen asiantuntijalle intuitiivisiin tunnuslukuihin perustuen. Pohjajakauman valinta polynomiseen kvantiilisekoitukseen oli haasteellinen Myöhemmin sovitettujen jakaumien graafisessa tarkastelussa nähdään, miten jakauma on katkaistu. Pohjajakauman parametreja μ ja σ muuttamalla pohjajakaumasta saisi paremmin toimivan. Jakauman vasemman hännän katkeaminen voitaisiin estää kiinnittämällä malliparametri $a_0 = 0$. Vastaavasti polynomin astetta pitäisi kasvattaa yhdellä, jotta vapaita malliparametrejä olisi edelleen sama määrä. Pohdinnassa palataan tutkimaan kehitysehdotuksia todennäköistämistyökalulle.

Todennäköistämisessä jakaumien tunnuslukuja korjattiin useaan otteeseen kaikkien asiantuntijoiden ja tuoteperheiden kohdalla. Useimmiten jakaumaa kavennettiin todennäköistämisen edetessä. Ajallisesti prosessi oli hitain ensimmäisen tuoteperheen kohdalla ja nopeutui kohti viimeistä tuoteperhettä. Sekä asiantuntija, että haastattelija kehittyivät haastattelutilanteessa. Asiantuntija osasi ottaa nopeammin keskeisiä asioita esille ja huomioon arvioissaan. Vastaavasti myös haastattelija ohjasi keskustelua tehokkaammin loppua kohti. Kunkin asiantuntijan kanssa käytettiin tunnista puoleentoista tuntiin viiden jakauman muodostamiseen. Haastatteluihin varattiin kaksi tuntia per henkilö, joten todennäköistäminen sujui tavoiteajassa.

7.3 Työkalulla kerätty aineisto asiantuntijahaastatteluista

Todennäköistämistyökalu mahdollisti haastatteluiden tulosten helpon tallentamisen. Työkalun avulla tuoteperheelle mallinnettu ennustejakauma tallennettiin kerättyjen tunnuslukujen muodossa.

Työkalulla kerätyt jakauman tunnusluvut on esitetty taulukossa 4. Taulukon rivi vastaa yhden tuoteperheen todennäköistämisen tulosta yhden asiantuntijan kanssa. Taulukon luvut on esitetty sillä tarkuudella kuin ne on haastattelutilanteessakin kerätty. Näihin tunnuslukuihin on sovitettu ja esitetty haastattelutilanteessa lopullinen jakauma. Kyseinen jakauma on todennäköistämisen lopputulos. Tällöin kunkin tuoteperheen kohdalla asiantuntijan kanssa varmistuttiin mallinnetun jakauman kuvastavan asiantuntijan näkemystä.

Taulukosta 4 havaitaan, että ennusteita tuoteperheiden tilausmäärille on kyetty arvioimaan kymmenien kappaleiden tarkkuudella. Tätä tarkemmat arviot ovat seurausta pääosin jakauman hienosäädöstä. Liitteessä A on tunnuslukuja vastaavien sovitettujen ennustejakaumien malliparametrit ja L-momentit. Seuraavaksi perehdytään graafisesti todennäköistämisen tuloksiin.

Taulukko 4: Asiantuntijahaastatteluissa työkaluun kerätyt, ennustejakaumien määrittämiseen käytetyt tunnusluvut, joiden selitteet löytyvät taulukosta 1

asiantuntija	tuote	x_1	x_2	x_3	u_1	u_3	τ_4
J	BB3	6	10	20	0.25	0.75	0.35
K	BB3	8	12	19	0.25	0.75	0.16
P	BB3	5	10	20	0.25	0.75	0.29
R	BB3	4	6	8	0.25	0.75	0.36
J	BB4	216	237	247	0.25	0.75	0.10
K	BB4	160	170	180	0.25	0.75	0.10
P	BB4	200	250	300	0.25	0.75	0.33
R	BB4	216	230	242	0.25	0.75	0.46
J	BB5	279	300	310	0.25	0.75	0.13
K	BB5	270	300	320	0.25	0.75	0.10
P	BB5	275	350	400	0.25	0.75	0.25
R	BB5	300	330	340	0.25	0.75	0.18
J	BB6	160	190	210	0.25	0.75	0.10
K	BB6	120	130	140	0.25	0.75	0.26
P	BB6	150	200	250	0.25	0.75	0.04
R	BB6	150	160	170	0.25	0.75	0.12
J	BB7	62	75	80	0.10	0.90	0.12
K	BB7	57	76	87	0.25	0.75	0.21
P	BB7	80	100	120	0.25	0.75	0.27
R	BB7	70	80	85	0.25	0.75	0.06

7.4 Tilausmääräennusteiden graafinen tarkastelu

Todennäköistämisessä asiantuntijan kanssa määritettiin ennustejakauma tuoteperheen tilausmäärälle. Jakaumista esitetään tiheysfunktiot, jotka olivat myös todennäköistämisessä tärkein apuväline asiantuntijan tuottaman informaation tarkastelussa. Vertailun helpottamiseksi yhdistämisen tulokset esitetään yksittäisten asiantuntijoiden jakaumien yhteydessä. Näin pystytään muodostamaan kokonaiskuva kunkin tuoteperheen tilausmäärän ennusteesta.

Kuvissa 9–13 yläosassa on esitetty eri asiantuntijoiden kanssa sovitetut yksittäiset jakaumat. Kuvien alaosassa nähdään puolestaan yksittäisistä jakaumista eri menetelmin yhdistetyt jakaumat. Lisäksi kuviin on merkittynä pystysuoralla katkoviivalla kyseisen tuoteperheen toteunut tilausmäärä.

Kuvissa 10, 11 ja 13 nähdään yksittäisten asiantuntijoiden jakaumissa todella teräviä huippuja. Tilausmääriä ajatellen nämä eivät luultavasti kovin hyvin kuvasta asiantuntijan todellista näkemystä. Sen sijaan määritettyjen kvantiilien välissä jakauman odottaisi näyttävän sileämmältä. Tähän jakaumaperheen ominaisuuteen ei keksitty ratkaisua todennäköistämisen suunnitteluvaiheessa. Pohdinnassa käsitellään parannusehdotuksia jakaumaperheelle.

7.4.1 Asiantuntijahaastatteluissa sovitetut jakaumat

Yksittäisistä jakaumista nähdään, että pääsääntöisesti ennusteet ovat samansuuntaisia. Asiantuntijoille löydetään kuitenkin ominaisia piirteitä. Tuoteperheiden tasolla vertailtaessa tulkitaan eroja perusteluissa ja ennustejakaumien poikkeavuuksista.

Asiantuntija K on selkeästi antanut muita pienempiä ennusteita. Taustalla vaikuttavaksi tekijäksi arvioidaan hänen toimintasektorinsa, joka on muita rajatumpi. Tällöin hänen ennusteensa voivat olla aliestimoituneita, mutta jakauman muoto on sopiva. Asiantuntija P puolestaan on järjestelmällisesti arvioinut epävarmuutta tulevasta tilausmäärästä muita asiantuntijoita suuremmaksi. Ennusteissaan hän ei yhtä vahvasti huomioinut aiempia tilauksia. Lähestymistapana hänellä oli kaikkien mahdollisten skenaarioiden huomioiminen ennusteissa. Asiantuntijat J ja R antoivat kauttaaltaan samansuuntaisia ennustejakaumia. Heidän lähestymistapansa oli aiempiin tilausmääriin pohjaava, mutta kuitenkin he huomioivat vahvasti omaa informaatiotaan.

Tuoteperheiden välisten erojen tarkastelulla voidaan selvittää, missä tuoteperheissä asiantuntijoiden oma informaatio näkyy selvimmin. Tuoteperheen BB3 kohdalla ei voida juurikaan puhua poikkeavista arvioista. Määrät ovat lähes identtisiä, eikä haastattelutilanteessa ilmennyt eroja perusteluissa. Arviot perustuivat lähes täydellisesti oletukseen, että tuotteita menee saman verran kuin vuotta aikaisemminkin vastaavana ajankohtana. Huomioitavana seikkana, on että edellisvuotena ensimmäisessä kvartaalissa kysyntä on ollut heikointa muuhun vuoteen nähden.

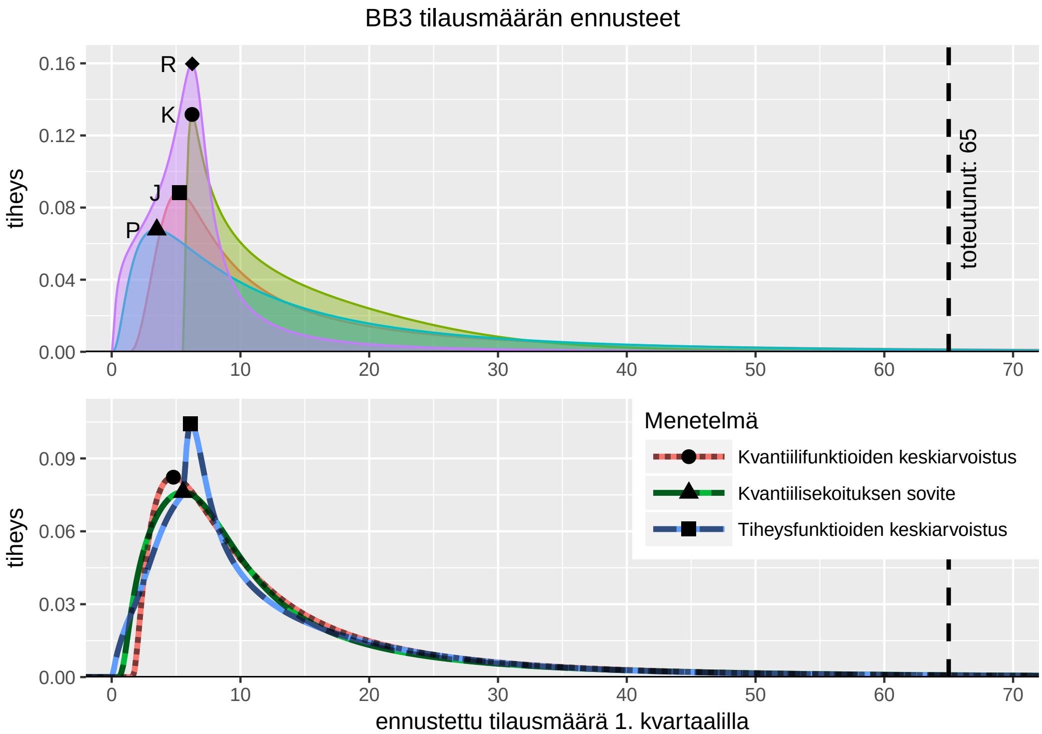
Tuoteperheen BB4 kysynnän arvioitiin reilusti kasvavan vuoden takaisesta. Henkilön K arvio poikkesi muista selkeästi pienemmällä ennusteella. Tämän tuoteperheen kohdalla ensimmäinen kvartaali on aikasempana vuotena ollut selkeästi paras kysynnältään. Asiantuntijat arvioivat tilausmäärään vaikuttavan ensisijaisesti vuodenajan ja sen lisäksi ennustivat merkittävää kasvua kysynnässä.

Tuoteperheen BB5 kysynnän arvioitiin lievästi kasvavan vuoden takaisesta. Tässä asiantuntijoiden välillä on havaittavissa pieniä eroja keskenään Asiantijat perustivat näkemyksensä selkeästi edellisen vuoden vastaavaan ajankohtaan.

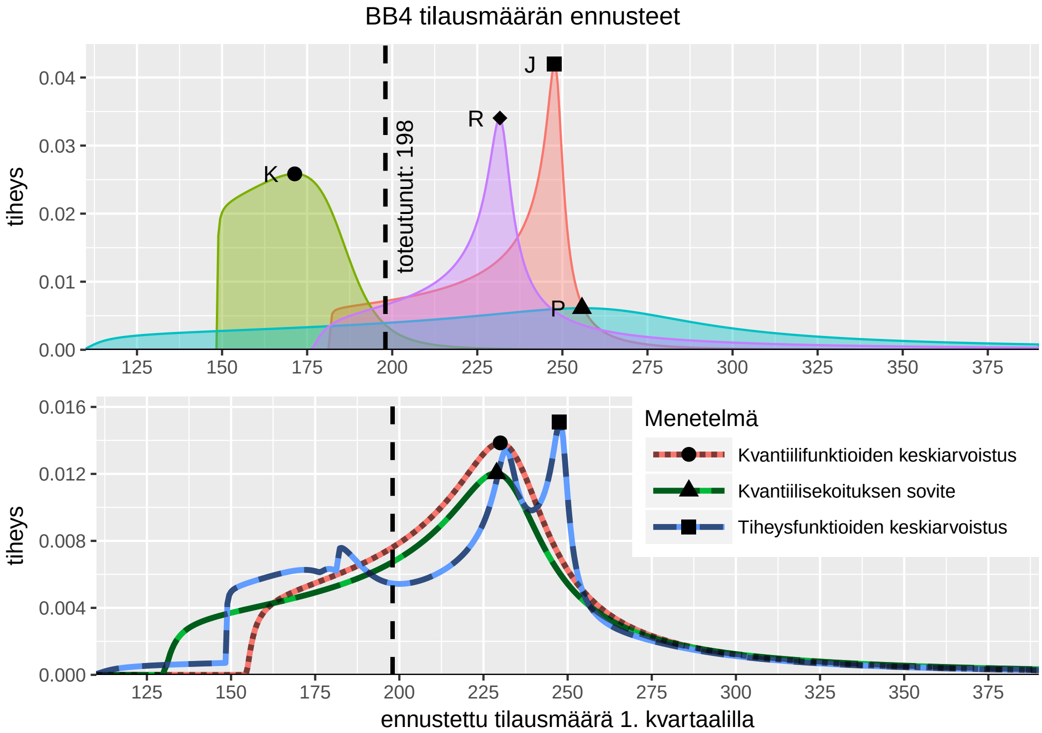
Tuoteperheen BB6 kysynnän arvioitiin lievästi kasvavan vuoden takaisesta. Tämän tuoteperheen kohdalla asiantuntijoilla oli selkeästi eniten eroja ennustetuissa jakaumissa. Poikkeavat arviot tulkitaan siten, että asiantuntijat ovat tuoneet selkeästi omaa informaatiotaan ennustejakaumaan. Arviot tulevasta tilausmäärästä perusteltiin aiemman vuoden vastaavan ajankohdan tuloksella ja tuoteperheen vakaalla kysynnällä.

Tuoteperheen BB7 kysynnän ei arvioitu muuttuvan vuoden takaisesta. Asiantuntijoiden väliset erot ovat lieviä. Haastattelutilanteissa kysynnän arvioitiin olevan kausiluontoista ja lisäksi tuoteperhettä pidettiin erityisen vakaana kysynnän suhteen.

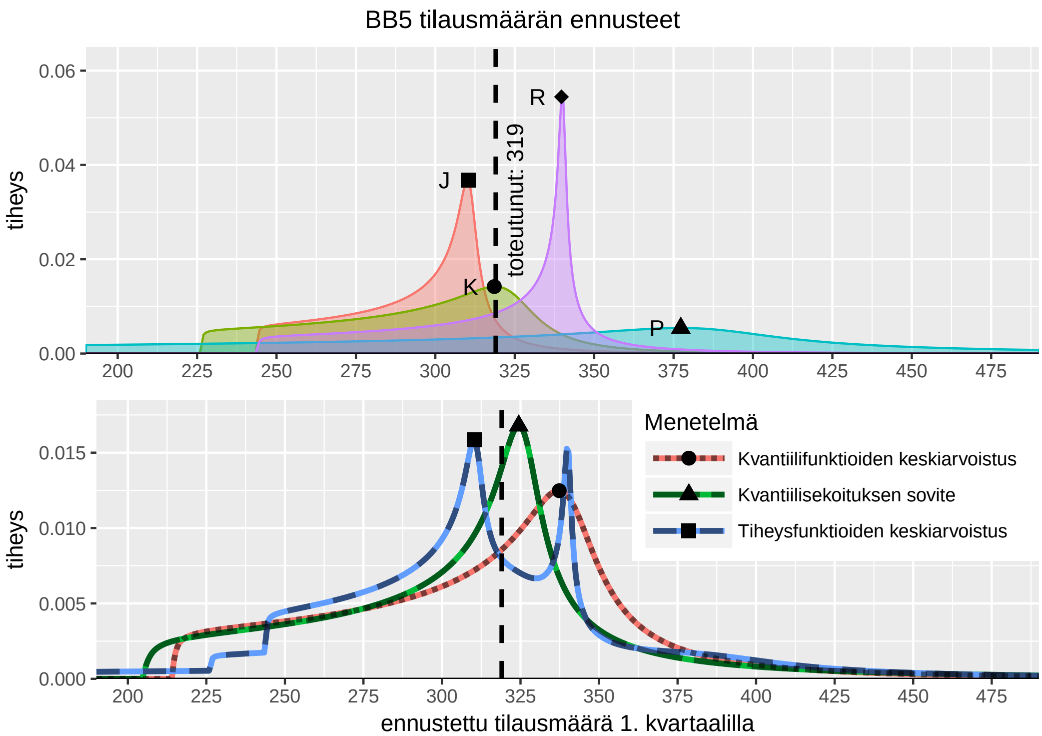
Kokonaisuutena näyttäisi siltä, että tilausmäärälle on annettu ennuste pohjautuen vuotta aikaisempiin tilauksiin. Asiantuntijat ovat vaihtelevasti huomioineet taustalla vaikuttavia tekijöitä, joka näkyy eroina etenkin tuoteperheen BB6 kohdalla. Tuoteperheen BB3 kohdalla eroja ei löydetä asiantuntijoiden välillä. Tuoteperheiden BB4 ja BB5 osalta asiantuntijoiden perusteluista ilmeni, että todennäköistämisessä on huomioitu kohtalaisen kattavasti taustatekijöitä. Kasvun ennustaminen näkyi leveämpinä jakaumina. Tuoteperheiden BB6 ja BB7 osalta taustatekijöitä huomioitiin eniten. Johtopäätöksissä palataan tulkitsemaan asiantuntijoiden mahdollista harjaantumista todennäköistämisprosessin aikana.



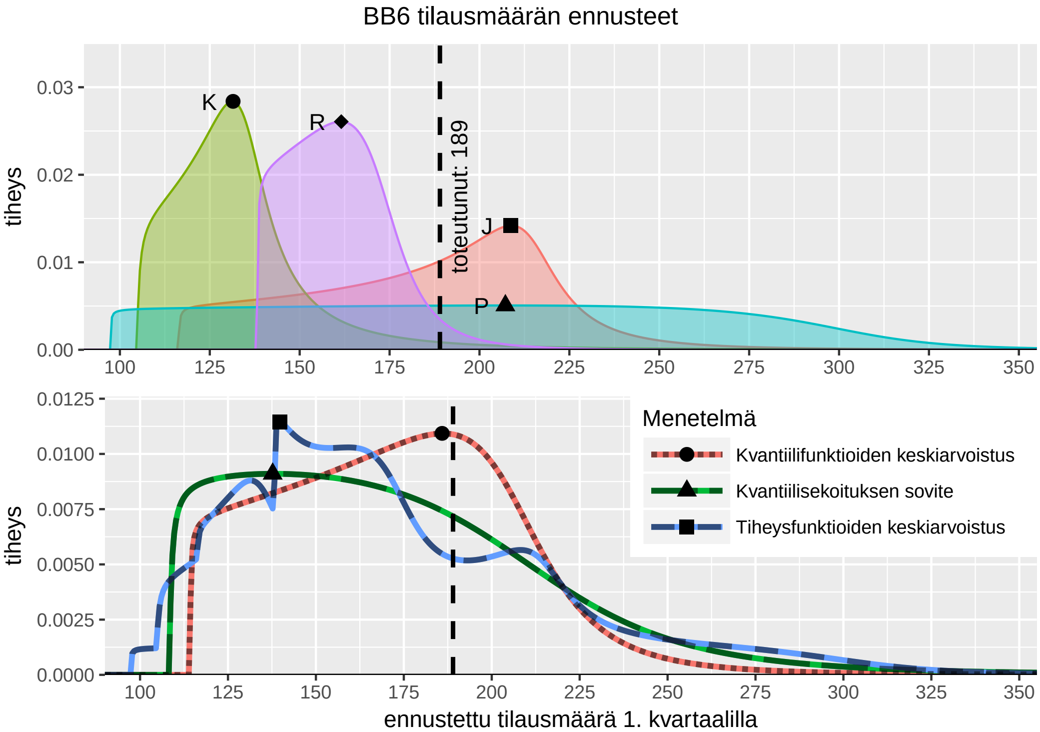
Kuva 9: Ennustejakaumien tiheysfunktiot tuoteperheelle BB3; *ylhäällä:* asiantuntijoiden ennustejakaumat tilausmäärälle, *alhaalla:* ennustejakaumien yhdisteet eri menetelmin, *katkoviivalla:* toteutunut tilausmäärä



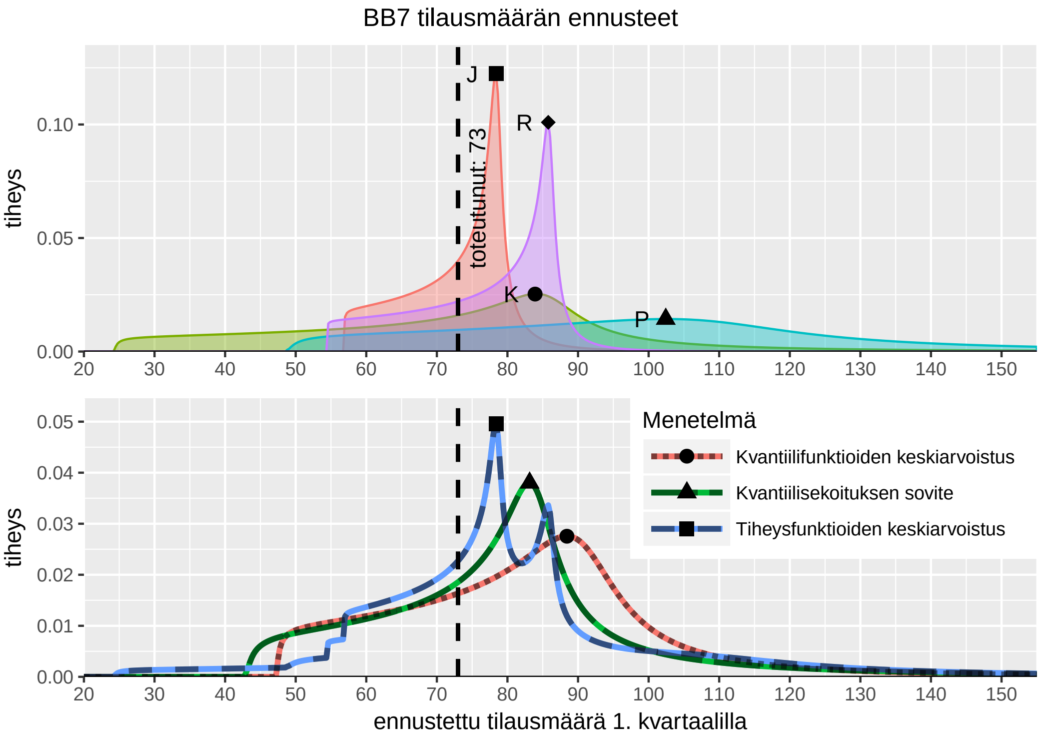
Kuva 10: Ennustejakaumien tiheysfunktiot tuoteperheelle BB4; *ylhäällä:* asiantuntijoiden ennustejakaumat tilausmäärälle, *alhaalla:* ennustejakaumien yhdisteet eri menetelmin, *katkoviivalla:* toteutunut tilausmäärä



Kuva 11: Ennustejakaumien tiheysfunktiot tuoteperheelle BB5; *ylhäällä:* asiantuntijoiden ennustejakaumat tilausmäärälle, *alhaalla:* ennustejakaumien yhdisteet eri menetelmin, *katkoviivalla:* toteutunut tilausmäärä



Kuva 12: Ennustejakaumien tiheysfunktiot tuoteperheelle BB6; *ylhäällä:* asiantuntijoiden ennustejakaumat tilausmäärälle, *alhaalla:* ennustejakaumien yhdisteet eri menetelmin, *katkoviivalla:* toteutunut tilausmäärä



Kuva 13: Ennustejakaumien tiheysfunktiot tuoteperheelle BB7; *ylhäällä:* asiantuntijoiden ennustejakaumat tilausmäärälle, *alhaalla:* ennustejakaumien yhdisteet eri menetelmin, *katkoviivalla:* toteutunut tilausmäärä

7.4.2 Asiantuntijoiden ennustejakaumien yhdistämisen tulokset

Kuvien 9–13 alaosassa on esitetty asiantuntijoiden ennustejakaumien yhdistämisen tulokset. Menetelmien tulokset poikkeavat toisistaan lähinnä muodon suhteen. Tärkeä huomio on, että kvantiilifunktioiden painotus ja parametrisen mallin sovittaminen yhteistiheydelle tuottavat erilaisia tuloksia. Molemmat ovat samasta jakaumaperheestä, kuin yksittäiset asiantuntijoiden määrittämät jakaumat. Näiden menetelmien avulla sovitettujen jakaumien malliparametrit ja L-momentit löytyvät liitteestä B.

Kuvista nähdään, että tiheysfunktioiden painottaminen tuottaa monihuippuisia ja kulmikkaita jakaumia. Monihuippuisuus ilmentää asiantuntijoiden arvioiden poikkeavuutta toisistaan. Kulmikkuus johtuu puolestaan polynomisen kvantiilisekoituksen ominaisuuksista ja siihen liittyvän pohjajakuman valinnasta. Etenkin tuoteperheen BB4 kohdalla monihuippuisuus ja kulmikkuus korostuvat.

Kvantiilifunktioiden painotuksella yhdistetyt ennustejakaumat näyttävät sopivan hyvin yksittäisiin jakaumiin. Yhdistetty jakauma on muita menetelmiä kapeampi, jolloin useampaa asiantuntijaa hyödynnettäessä epävarmuus tilausmäärästä ei kasva. Kvantiilifunktioiden sovite on kahden muun menetelmän välimuoto ja tiheysfunktiot näyttävät eri tuoteperheissä sileiltä ja tilauksien ennustamiseen sopivilta.

Tuoteperheen BB6 kohdalla edelliset menetelmät tuottavat mielenkiintoisen tuloksen. Kuvassa 12 nähdään, miten jakauman vasen häntä katkeaa nopeasti. Käytännössä se tarkoittaisi, että asiantuntijat ovat varmoja siitä, että tuotteita tilataan vähintään noin 110 kappaletta ja sitä suuremmille määrille annetaan melko tasainen ennuste. Yhdistämismenetelmien väliset erot havaitaan tarkemmin verrattaessa ennustejakaumia toteutuneisiin tilausmääriin.

Yksittäisiä ja yhdistettyjä jakaumia ei voida pitää käyttökelpoisena pienien tilausmäärien todennäköisyyksien arvioinnissa. Yleisellä tasolla mallinnusta voidaan pitää hyvänä kvartiilivälillä, mutta sen ulkopuolella yksittäiset, eivätkä yhdistetyt jakaumat luultavasti vastaa asiantuntijan näkemystä tai todellisuutta.

7.5 Ennusteiden vertailu toteutuneisiin tilausmääriin

Black Bruinin tuotteiden toteutuneet tilausmäärät auttavat hahmottamaan sovitettujen jakaumien mielekkyyttä. Ennusteiden osuvuutta mitataan tutkimalla kertymäfunktioiden arvoja toteutuneen tilausmäärän arvossa. Täydellinen osuma kertymäfunktiolla tarkasteltuna tarkoittaa siis arvoa 0.5. Piste-estimaattina tilausmäärälle käytetään jakauman mediaania. Piste-estimaatit löytyvät asiantuntijoittain taulukosta 4. Asiantuntijoista esitetään lisää johtopäätelmiä, mutta huomio on kuitenkin kollektiivisessa ennustamisessa. Tuoteperheen BB3 kohdalla tarkastelut tehdään erikseen, sillä ennusteet poikkesivat täysin toteutuneesta.

Asiantuntijoiden ennustejakaumien osuvuus ilmenee taulukosta 5. Riittävimmällä epävarmuudella ennusteet on asettanut asiantuntija P. Tämä näkyy kertymäfunktioiden arvoissa, jotka kaikki ovat asiantuntijalla P lähellä kvartiiliväliä. Hän on kuitenkin kaikissa ennusteissaan yliarvioinut tulevia tilausmääriä. Asiantuntijoiden J, K ja R jakaumat ovat pääsääntöisesti lähes yhtä leveitä. Heidän ennustejakaumansa ovat siis muutamissa tapauksissa osuneet joko hieman sivuun tai epävarmuutta on aliestimoitu. Asiantuntija K on kuitenkin selkeästi aliestimoinut tilausmääriä, joka nähtiin jo graafisissa tarkasteluissa. Piste-estimaattien tasolla asiantuntijat J ja R ovat suoriutuneet parhaiten ennusteissaan.

Taulukko 5: Asiantuntijoiden ennustejakaumien kertymäfunktioiden arvot toteutuneissa tilausmäärissä

nimi	BB3	BB4	BB5	BB6	BB7
J	0.96	0.10	0.92	0.49	0.41
K	1.00	0.97	0.74	0.97	0.45
P	0.97	0.24	0.38	0.44	0.18
R	1.00	0.10	0.38	0.96	0.31

Asiantuntijoiden ennusteista yhdistetyt jakaumat kertovat koko pitkän todennäköistämisprosessin tuloksen. Eri menetelmillä yhdistettyjen jakaumien mediaanit käyvät ilmi taulukosta 6. Tiheysfunktioiden keskiarvoistuksella ennusteet ovat pienimpiä ja kvantiilien keskiarvoistuksella ennusteet ovat suurimpia läpi tuoteperheiden. Kappalemääräisesti toteutuneiden tilauksien ja ennustettujen erot ovat korkeintaan 10% koko tuoteperheen kysynnästä kvartaalin aikana. Poikkeavuuksia voidaan pitää kohtuullisina tuoteperheille BB4 – BB7.

Taulukko 6: Ennustejakaumien yhdisteiden mediaanien vertailu toteutuneeseen tilausmäärään

	BB3	BB4	BB5	BB6	BB7
Tiheysfunktioiden keskiarvoistus	9	226	309	164	78
Kvantiilisekoituksen sovite	9	222	315	165	80
Kvantiilifunktioiden keskiarvoistus	9	225	320	173	83
Toteutuneet tilaukset	65	198	319	189	73

Jakaumamielessä ero toteutuneeseen tilausmäärään käy ilmi taulukosta 7. Parhaiten asiantuntijat ovat kyenneet ennustamaan tuoteperheen BB5 kysynnän, joka on osunut ihan täsmälleen ennustettuun. Tämän jälkeen tulevat järjestyksessä BB6, BB7 ja BB4. Näiden neljän tuoteperheen ennusteet osuvat suunnilleen ennustejakaumien kvartiilivälien sisään. Tulosta voidaan pitää hyvänä. Tämän tutkielman johtopäätöksissä palataan tarkastelemaan jakaumien epävarmuuden käyttökelpoisuutta.

Taulukko 7: Ennustejakaumien yhdisteiden kertymäfunktioiden arvot toteutuneissa tilausmäärissä

	BB3	BB4	BB5	BB6	BB7
Tiheysfunktioiden keskiarvoistus	0.98	0.32	0.61	0.69	0.33
Kvantiilisekoituksen sovite	0.98	0.29	0.56	0.69	0.33
Kvantiilifunktioiden keskiarvoistus	0.98	0.23	0.49	0.67	0.31

Tuoteperheen BB3 kohdalla asiantuntijat ennustivat täsmälleen vuodentakaisten tilausmäärän mukaan. Heidän saamissaan ennakkotiedoissa BB3 tilausmäärät olivat muissa kvartaaleissa muutamissa kymmenissä. Toteutunut tilausmäärä 65 kpl oli moninkertainen ennustettuun. Odottamaton tilausmäärä johtui jälkikäteen tehdyn selvityksen mukaan lukuisista syistä. Loppuvuotta kohti tuoteperheen BB3 tilausmäärä oli kasvanut edellisenä vuotena. Tästä syystä arvioitiin, että asiakkaiden varastoissa tuotetta olisi tarpeeksi, eikä tilauksia saapuisi. Tuoteperhe BB3 on poistumassa yrityksen valikoimasta ja asiakkaat haalivat tästä tiedosta johtuen itselleen tuotetta varastoon tavallistakin enemmän. Asiakkailta vaadittiin myös isompia kertatilauksia. Nämä tiedot eivät tulleet huomioiduiksi asiantuntijoiden arvioissa, joten tässä kohtaa todennäköistäminen epäonnistui.

Kvartiiliväliä voidaan pitää hyvänä mittarina osuvuudesta, sillä jokainen asiantuntija on määrittänyt sen. Yksittäisten asiantuntijoiden määrittämissä ennustejakaumissa osuvuus toteutuneeseen tilausmäärään oli vaihtelevaa. Asiantuntijoiden työtehtävistä johtuen tilausmäärää ennustettiin myös eri tavalla.

Kollektiivisesti asiantuntijat sen sijaan pystyivät melko luotettavasti tuottamaan ennusteita. Tärkein huomio on, että toteutuneet tilausmäärät osuivat pääsääntöisesti yhdistettyjen ennustajakaumien kvartiilivälien sisäpuolelle. Mediaaniestimaateissa erot olivat menetelmien välillä melko pieniä, mutta jakauman muodossa huomattavia. Kvantiilifunktioiden keskiarvoistuksella saadaan käyttökelpoisimpia ennustejakaumia, sillä asiantuntijoilla oli vaihtelevasti aiempaa kokemusta myynnin ennustamisesta. Tiheysfunktioiden avulla yhdistetty ennustejakauma on useimmissa tapauksissa monihuippuinen. Kvantiilisekoituksen sovitteella yhdistetty ennustejakaumas on kahden muun menetelmän välimuoto.

8 Johtopäätökset

Tilausten ennustaminen asiantuntijahaastatteluilla on monivaiheinen ja laaja tehtävä. Tässä osiossa pohditaan kokonaisuutta ja esitetään parannusehdotuksia asiantuntijahaastattelun toteuttamiselle. Lisäksi arvioidaan yrityksen saavuttamaa hyötyä ja esitetään ajatuksia menetelmän käytettävyydestä kysynnän ennustamisessa ja muissa tilanteissa.

8.1 Polynomiset kvantiilisekoitukset tilausmäärän ennustamisessa

Ennustejakauman mallintaminen perustui kvartiilien hyödyntämiseen. Aiemmin ei ole esitelty työkalua ja menetelmää, jotka mahdollistaisivat jakauman sovittamisen eksaktisti kvantiileihin. Polynomiset kvantiilisekoitukset sopivat tarkoitukseen hyvin. Jakauman muodon hienosäädössä log-normaali pohjajakauma ei täyttänyt kaikkia odotuksia, mutta toimi tehtävän kannalta riittävän hyvin.

Ongelmina jakauman muodossa olivat huipukkuus ja jakaumien häntien käyttäytyminen. Polynomisessa kvantiilisekoituksessa pohjajakuma asettaa rajoitteita jakauman muodolle. Pohjajakauman valintaan käytettiin runsaasti aikaa, mutta siltikään ei voinut ennakkoon varmistua jakaumaperheen sopivuudesta tilausmäärien ennustamiseen ja asiantuntijoiden näkemyksiin. Valinnassa oltaisiin voitu tutkia vielä paremmin pohjajakauman muotoa ja parametrien valintaa. Asiantuntijahaastatteluiden toteuttamisen jälkeen parannusehdotuksia on helppo nimetä.

Useissa mallinnetuissa jakaumissa jakauman vasenta häntää voisi kuvailla pystysuoraksi. Tämän lisäksi pienet tilausmäärät saavat 0-todennäköisyyksiä. Todellisuudessa se ei voi pitää paikkaansa, joten jakaumien vasenta häntää alakvartiilista laskettuna ei voida pitää realistisesti mallintuneena. Log-normaalijakauman μ ja σ -parametreja muuttamalla hännän käyttäytymiseen voidaan vaikuttaa. Näiden parametrien säätömahdollisuutta voisi myös harkita työkaluun lisättäväksi. Log-normaalijakauman mediaani on $\exp(\mu)$. Täten pohjajakauman parametrit voisivat myös muuttua asiantuntijan määrittämien kvartiilien mukaan.

Polynomisen kvantiilisekoituksen malliparametrien ratkaisemisessa on myös mahdollista vaikuttaa jakaumaperheen ominaisuuksiin. Vakiotermin asettamisella $a_0 = 0$ polynomisen kvantiilisekoituksen todennäköisyysjakauma kattaisi täsmälleen positiivisen reaaliakselin. Kvantiilifunktioon on tämän muutoksen myötä lisättävä yksi korkeamman asteen polynomitermi, jotta malliparametreja on edelleen sama määrä kuin haastattelussa määrättäviä tunnuslukuja.

Ennustejakaumien yhdistäminen kvantiilifunktioiden avulla osoittautui toimivaksi menetelmäksi. Kvantiilifunktioiden avulla jakaumia yhdistettäessä tuloksena on kvantiilisekoitus. Kvantiilisekoitus helpottaa tulkintaa, sillä yhdistämisen tulos on samasta jakaumaperheestä kuin yksittäiset ennustejakaumat. Tämän lisäksi menetelmän eduksi voidaan katsoa L-momenttien tulkinta yhdistämisessä. Useampaa asiantuntijaa käytettäessä ennustejakauman voisi ajatella tarkentuvan. Tiheysfunktioiden lineaarisella keskiarvoistuksella tätä ominaisuutta ei saavuteta, vaan jakauma on sitä leveämpi, mitä enemmän asiantuntijoita. Kvantiilifunktioiden avulla yhdistäminen sopii siis hyvin tilanteeseen, jossa yksittäiset ennustejakaumat ovat kvantiilisekoituksia. Tämän lisäksi suuresti poikkeavilla yksittäisillä ennustejakaumilla on vähemmän vaikutusta yhdisteeseen.

Tulosten perusteella pohjajakauma Lognormal$(0,1)$ toimi riittävän hyvin polynomisessa kvantiilisekoituksessa. Osa mallinnetuista jakaumista on hyvinkin huipukkaita, mutta osa puolestaan hyvinkin tasaisia. Liiallisen huipukkuuden korjaaminen jää haastattelijan vastuulle hienosäätöparametrin avulla. Polynomisista kvantiilisekoituksista on ennustejakauman mallintamisessa neljä selvää etua. Ne mahdollistavat asiantuntijalle intuitiivisten tunnuslukujen käytön, jakauman eksaktin sovittamisen, hienosäädön L-momenttien avulla ja lisäksi parametrien ratkaiseminen on laskennallisesti helppoa. Muutamat jakaumaperheen huonot ominaisuudet ovat luultavasti väistettävissä luetelluilla parannusehdotuksilla.

8.2 Asiantuntijahaastatteluiden valmistelun parantaminen

Asiantuntijoita valmisteltiin todennäköistämiseen tarjoamalla heille ennakkoon tietoa haastatteluista. Asiantuntijoiden valmistelusta heille uuteen tehtävään löydetään parannettavaa. Harjoittelukierros todennäköistämisessä ja asiantuntijoiden huolellisempi valikoiminen veisivät entistä parempiin tuloksiin.

Asiantuntijoiden saama ennakkotiedon määrä oli riittävä haastatteluun valmistautumisessa. Asiantuntijahaastattelu oli kaikille asiantuntijoille uusi haaste, mutta tästä huolimatta todennäköistämisessä onnistuttiin lähes kokonaan. Ajallisesti tarkasteltuna asiantuntijat kehittyivät ja todennäköistäminen sujui nopeammin viimeisten tuoteperheiden kohdalla. Syynä on kommunikoinnin sujuvoituminen todennäköistämisen tultua tutuksi. On myös mahdollista että asiantuntijat oppivat huomioimaan olennaiset tekijät ennusteensa kannalta nopeammin. Olisi kuitenkin ollut järkevää tehdä yksi harjoittelukierros, sillä ensimmäisenä mallinnetun tuoteperheen BB3 ennuste ei osunut lähellekään toteutunutta.

Kahden ensimmäisenä haastateltavan asiantuntijan kanssa ei päästy todennäköistämisessä edes alkuun. Heidän kohdallaan todennäköistäminen olisi selvästi vaatinut lisää valmistelua. Heidän tehtävissä ei ole aikaisemmin ollut mukana tilausmäärien tarkkaa seurantaa, mutta he ovat kuitenkin asiakkuuksien kanssa tekemisissä. Työtehtäviä kehittämällä he voisivat tuoda merkittävää informaatiota asiakkailta osaksi tilausten ennustamista. Esimerkiksi asiakkailta voitaisiin tiedustella ennakkotietoa tulevista tilauksista.

Muiden haastateltavien kanssa todennäköistämisessä onnistuttiin. Asiantuntijat voivat kuitenkin kehittyä edelleen tehtävässä. Esimerkiksi tietojen ylöskirjaaminen merkittävistä päätöksistä ja asiakkaista edistäisivät kokonaiskuvan muodostamista. Black Bruinin toiminnan ohjaamisesta vastaavien on syytä tuoda tilausmäärien seuraaminen useamman avainasemassa olevan henkilön työnkuvaan. Tämä auttaa jatkossa tekemään tarkempia tietoon perustuvia päätöksiä.

Asiakkuudet ovat niin merkittävässä asemassa, että ennusteita voidaan laatia asiakkaittain ja moottorin kokoluokan perusteella. Toteuttaminen vaatii kuitenkin työn jakamista useammalle eri työntekijälle.

8.3 Työkalun jatkokehitys

Työkalu asiantuntijahaastatteluun ei ole luvussa 5 esitellyssä muodossaan vielä valmis yleiseen käyttöön. Työkalu on rakennettu tiettyyn tilanteeseen sopivaksi, eikä siinä ole ohjeistusta mukana. Ominaisuuksia ja ohjeistusta lisäämällä työkalu voisi toimia yleispätevästi jakauman mallintamisessa asiantuntijahaastattelulla. Tämän lisäksi tulosten johtaminen työkalulla tallennettuista tiedoista on työlästä.

Jatkokehitettynä työkalua voisi ajatella käytettävän samankaltaisiin ongelmiin muissakin yrityksissä. Työkalun ominaisuudet on toteutettu tilausten ennustamista Black Bruinilla ajatellen. Työkalun ominaisuuksia olisi kehitettävä toimimaan yleispätevämmissä tilanteissa ja etenkin tulosten tulkitsemisen helpottamiseksi.

Esimerkiksi muistiinpanot ovat tärkeässä roolissa tulosten tulkinnassa. Haastattelija voisi kirjoittaa muistiinpanonsa työkaluun ja jokaisen lauseen kohdalla näkisi aikaleiman. Tämä edesauttaisi haastattelutilanteen palautumista mieleen ja mahdollistaisi tulosten mielekkyyden tarkastamisen jälkikäteen. Aikaleimojen perusteella olisi myös mahdollista luoda raportti, jossa kävisi ilmi pääpiirteissä todennäköistämisen vaiheet. Raportissa olisi ajan suhteen kuvattuna työkalun syötearvot, sovitettu jakauma ja asiantuntijan kommentit. Muistiinpanojen lisäysmahdollisuus on helppo toteuttaa, mutta valmis raportointiominaisuus on vaativa kokonaisuus.

Työkalun käyttäminen vaatii haastattelijalta tietämystä puolitusmenetelmästä ja kvantiilisekoitusten teoriasta. Tämän lisäksi haastattelijan on kyettävä esittämään sovitetusta jakaumasta tarkentavia kysymyksiä. Haastattelijan on kyettävä selittämään todennäköisyysjakaumista jopa sellaiselle asiantuntijalle, jolla ei ole aiempaa tietoa jakaumista. Työkalun käyttö edellyttää siis haastattelijaltakin asiantuntemusta mallinnettavasta suureesta ja koko työkalusta. Tätä kynnystä on mahdollista kuitenkin madaltaa tiivistämällä teoriaa ja haastattelun ohjeistusta työkaluun liitettäväksi. Tulosten tuottamisen helpottamiseksi haastattelutyökalu voisi tallettaa haastattelun päätteeksi mallinnettuun jakaumaan liittyvät tunnusluvut ja funktiot käyttökelpoisina objekteina lokitietojen ohella.

Tässä työssä ennusteet tilauksille laadittiin useamman eri asiantuntijan toimesta. Mahdollista olisi myös käyttää vain yhtä asiantuntijaa, mutta kuten tuloksissa huomattiin, yksittäisten asiantuntijoiden ennusteet osuivat huonommin kuin kollektiiviset ennusteet. Ennusteiden yhdistäminen oli ohjelmoinnillisesti haastavaa ja tulosten tuottamisen helpottamiseksi valmis apuohjelma jakaumien yhdistämiseen on tarpeen. Useamman eri asiantuntijan haastatteluiden tulokset tallennetaan lokitietoina. Apuohjelman avulla haastatteluiden tulokset saisi automaattisesti purettua lokitiedostoista. Tämän lisäksi ohjelma tuottaisi jakaumien kuvaajat. Yhdistäminen tapahtuisi kvantiilifunktioiden painotuksena, joka vastaa kvantiilisekoituksia. Asiantuntijoiden ennusteiden yhdisteen tulkitsemiseksi apuohjelman avulla olisi kyettävä muodostamaan tuloksista R-objekteiksi tiheys-, kertymä, kvantiili- ja jakaumasta satunnaisesti generoivat funktiot. Polynomisten kvantiilisekoitusten teorian perusteella ne on mahdollista toteuttaa. Pidemmälle vietynä apuohjelma olisi haastattelutyökalua vastaava työkalu, jolla voisi tarkastella visuaalisesti ja numeerisesti jakaumaa.

Vertailukohta tämän työn työkalulle löydettiin työkalun kehityksen loppumetreillä. Asiantuntijahaastatteluihin sopiva verrokkityökalu on kuvattu artikkelissa *A web-based tool for eliciting probability distributions from experts* (Morris *et al.*, 2014). Vertailukohteen työkalussa on mahdollista sovittaa jakauma useammalla eri menetelmällä, mukaan lukien kvartiilien avulla. Kvartiileja hyödynnettäessä työkalu sovittaa ennaltamääritetyn jakauman parametrit pienimmän neliösumman menetelmällä. Valittavina jakaumien joukossa ovat muun muassa normaali-, t-, ja gammajakauma. Menetelmän etuna on sovitetun jakauman helppo tulkinta, mutta haittapuolena voidaan pitää sitä, ettei jakauma sovitu täsmällisesti annettuihin kvartiileihin. Vertailukohteessa on kuitenkin toteutettu monia käytännöllisiä ominaisuuksia, esimerkiksi työkalussa on ohjeistus mukana ja tuloksiin voi palata tallennettavan linkin avulla.

8.4 Black Bruinin saavuttama hyöty asiantuntijahaastatteluista

Tämä työ toteutettiin yrityksen tarpeesta ennustaa tuotteidensa kysyntää. Mitä jyväskyläläinen teollisuusyritys Black Bruin hyötyy tästä tutkielmasta? Entä miten yritys voi konkreettisesti hyödyntää asiantuntijahaastatteluiden tuloksista? Black Bruin saa yritykselle suunnatun raportin tämän tutkielman tuloksista ja johtopäätöksistä, joiden avulla se voi kehittää eteenpäin kysynnän ennustamista.

Todennäköistämisessä mallinnettuja ennustejakaumat sopivat kohtalaisen hyvin toteutuneisiin. Asiantuntijoilla oli kollektiivista informaatiota tilausten ennustamiseksi. Asiantuntijahaastattelut ovat aikaa vieviä ja eivät tässä työssä esitellyssä muodossa sovellu yrityksen hyödynnettäväksi tilausten ennustamisessa. Sen sijaan merkittäviä löydöksiä ovat kysyntään vaikuttavien tekijät ja henkilöstön opastaminen näiden huomioimiseksi. Tärkeää yrityksen kannalta on myös huomata, miten luotettavasti eri avainasemissa olevat asiantuntijat osasivat arvioida tilausmäärälle ennusteita. Asiantuntijat voivat kehittyä ennustamisessa pohtimalla systemaattisemmin tilauksiin mahdollisesti vaikuttavia tekijöitä. Tämä ajattelutapa voidaan tuoda myös yrityksen päätöksentekoon.

Asiantuntijoiden suoriutuminen oli vaihtelevaa, mutta ainakin piste-estimaatteja osattiin antaa oman toimialueen perusteella. Näyttäisi siis siltä, että tilausten ennusteiden laatiminen olisi mahdollista asiakkaittain ja tuoteperheittäin henkilöstön toimesta. Tämän avulla katetaan suurin osa myytävistä tuotteista. Ulkopuolelle jää kuitenkin varaosakauppa, joten joidenkin yksittäisten komponenttien kohdalle ennustamismenetelmä ei sovi lainkaan. Kaikkien komponenttien tilausten ennustaminen vaatisi automatisoidun järjestelmän käyttämistä. Valmiiden tuotteiden tilausmäärien ennusteita voitaisiin hyödyntää osana tätä, mutta loppuosa ennustamisesta on hoidettava tilaushistoriaan perustuvilla ennustemalleilla. Syynä on erilaisten komponenttien runsas lukumäärä ja asiantuntijahaastatteluiden vaatimat resurssit.

Tässä työssä käsiteltiin laajasti kysynnän ennustamisen ongelmaa kohdeyrityksessä. Tutkielman edetessä kävi selväksi, miten tärkeää Black Bruinin kannalta kysynnän ennustamisen tutkiminen on. Tämän työn myötä Black Bruinilla on hahmotettu kysynnän ennustamisen tarpeellisuus koko tuotantoprosessin kannalta. Kysyntää voidaan arvioida tilauksia ennustamalla, sillä se on kysynnästä Black Bruinille realisoituva osa. Ennustamista kvartaaleittain voidaan edelleen kehittää asiakkailta konkreettisen informaation hankkimisen ja tallettamisen avulla.

8.5 Asiantuntijahaastattelun sovelluskohteet

Asiantuntijahaastattelun toteutus on laajaa osaamista vaativa kokonaisuus. Mallintaminen asiantuntijahaastatteluilla vaatii kaikilta osin runsaasti aikaa ja joskus myös kiinnostavan suureen selvittäminen on haastava ongelma, kuten tämän työn tapauksessa. Black Bruinin kysynnän ennustamiseen kaltaisia ongelmia löytyy luultavasti muiltakin keskisuurilta teollisuusyrityksiltä. Kysynnän ja tilausten ennustamiselle on tarvetta muillakin toimialoilla, mutta haasteet ennustamisessa ovat toisenlaisia.

Valmiiden työkalujen puuttuessa kaikki vaiheet asiantuntijahaastattelusta on toteutettava itse. Tämä tutkielma tarjoaa kattavan menetelmäkokonaisuuden vastaavanlaisten asiantuntijahaastatteluiden toteuttamiseksi. Menetelmiä on mahdollista kehittää pidemmälle valmiin työkalun muodossa.

Mallintaminen asiantuntijahaastattelulla sopii erityisesti tilanteisiin, joissa ei ole käytössä sopivaa aineistoa. Suureen ennusteen tai nykytilan mallintamiseksi on oltava kuitenkin asiantuntijoita, joilla on kokonaisvaltainen käsitys suureesta. Menetelmän käyttökelpoisuus rajautuu tilanteisiin, joissa asiantuntijoiden informaation hankkimiseen on järkevää käyttää runsaasti aikaa. Valmiille työkalulle raportointiominaisuuksineen voisi ajatella löytyvän käyttöä tilanteissa, joissa mallinnusta joudutaan toistamaan. Asiantuntijahaastatteluille puolestaan voisi olla tarvetta päätöksenteossa, jonka vaikutuksia on arvioitava rahallisten kustannusten kautta.

Viitteet

Black Bruin Oy. Black Bruin Oy:n nettisivut. `https://www.blackbruin.com/fi/`, 2017. [Online; haettu 4.10.2017].

Winston Chang, Joe Cheng, JJ Allaire, Yihui Xie ja Jonathan McPherson. *shiny: Web Application Framework for R*, 2017. R package version 1.0.3.

Martin Christopher. *Logistics & supply chain management.* Pearson UK, 2016.

Martin Christopher ja Hau Lee. Mitigating supply chain risk through improved confidence. *International journal of physical distribution & logistics management*, **34**, no. 5, 388–396, 2004.

Robert T. Clemen ja Robert L. Winkler. Combining probability distributions from experts in risk analysis. *Risk analysis*, **19**, no. 2, 187–203, 1999.

Edwin L. Crow ja Kunio Shimizu. *Lognormal distributions.* Marcel Dekker New York, 1987.

Elsayed A. H. Elamir ja Allan H. Seheult. Trimmed L-moments. *Computational Statistics & Data Analysis*, **43**, no. 3, 299–314, 2003.

Paul H. Garthwaite, Joseph B. Kadane ja Anthony O'Hagan. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, **100**, no. 470, 680–701, 2005.

Jonathan R. M. Hosking. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, **52**, 105–124, 1990.

Allan James, Samantha Low Choy ja Kerrie Mengersen. Elicitator: An expert elicitation tool for regression in ecology. *Environmental Modelling & Software*, **25**, no. 1, 129–145, 2010.

Juha Karvanen. Estimation of quantile mixtures via L-moments and trimmed L-moments. *Computational Statistics & Data Analysis*, **51**, no. 2, 947–959, 2006.

Juha Karvanen. *Lmoments: L-moments and quantile mixtures*, 2016. R package version 1.2-3.

Kenneth C. Lichtendahl Jr., Yael Grushka-Cockayne ja Robert L. Winkler. Is it better to average probabilities or quantiles? *Management Science*, **59**, no. 7, 1594–1611, 2013.

David E. Morris, Jeremy E. Oakley ja John A. Crowe. A web-based tool for eliciting probability distributions from experts. *Environmental Modelling & Software*, **52**, 1–4, 2014.

Anthony O'Hagan, Caitlin E. Buck, Alireza Daneshkhah, J. Richard Eiser, Paul H. Garthwaite, David J. Jenkinson, Jeremy E. Oakley ja Tim Rakow. *Uncertain judgements: eliciting experts' probabilities.* John Wiley & Sons, 2006.

Emanuel Parzen. Nonparametric statistical data modeling. *Journal of the American statistical association*, **74**, no. 365, 105–121, 1979.

R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2017.

Christian P. Robert. *Monte Carlo methods.* Wiley Online Library, 2004.

Mervyn Stone *et al.* The opinion pool. *The Annals of Mathematical Statistics*, **32**, no. 4, 1339–1342, 1961.

Liitteet

Liite A: Malliparametrit ja L-momentit asiantuntijoittain

Taulukko A: Asiantuntijoiden määrittämien ennustejakaumien malliparametrit ja L-momentit

asian-tuntija	tuote	b	a_2	a_1	a_0	l_1	l_2	l_3	l_4
J	BB3	13.52	-3.12	-8.20	1.36	18.51	9.72	5.26	3.40
K	BB3	3.06	12.44	0.67	5.50	15.02	4.81	1.63	0.77
P	BB3	10.13	1.73	-1.17	0.03	16.71	8.78	4.08	2.55
R	BB3	4.10	-15.50	11.58	-0.01	7.37	2.87	1.11	1.03
J	BB4	4.82	-106.24	154.21	181.63	231.28	12.14	-1.63	1.21
K	BB4	3.11	-11.76	42.72	148.47	171.04	7.83	0.84	0.78
P	BB4	85.77	-324.19	274.83	107.86	278.63	65.38	23.23	21.58
R	BB4	50.04	-205.12	111.65	175.42	245.37	27.36	13.02	12.59
J	BB5	6.62	-113.01	155.77	243.75	294.88	12.81	-1.14	1.66
K	BB5	7.78	-109.41	186.79	226.18	295.93	19.57	-0.56	1.96
P	BB5	65.87	-448.97	507.46	142.64	355.32	66.28	11.18	16.57
R	BB5	13.02	-209.22	251.36	243.60	321.01	18.20	-1.81	3.28
J	BB6	7.78	-109.41	186.79	116.18	185.93	19.57	-0.56	1.96
K	BB6	11.23	-42.43	49.79	104.48	133.75	10.86	3.04	2.82
P	BB6	5.64	-21.30	204.92	97.23	201.88	35.44	1.53	1.42
R	BB6	3.87	-14.63	43.38	138.10	161.29	8.11	1.05	0.97
J	BB7	1.94	-36.41	50.84	56.74	73.23	4.07	-0.44	0.49
K	BB7	12.13	-109.86	134.58	24.04	74.72	14.53	1.15	3.05
P	BB7	23.89	-90.31	100.84	48.26	107.98	22.26	6.47	6.01
R	BB7	1.31	-44.95	71.14	54.36	77.10	5.49	-0.98	0.33

Liite B: Malliparametrit ja L-momentit jakaumien yhdistämisessä

Taulukko B: Yhdistettyjen ennustejakaumien malliparametrit ja L-momentit

jakaumien yhdistämismenetelmä	tuote	b	a_2	a_1	a_0	l_1	l_2	l_3	l_4
Kvantiilifunktioiden keskiarvoistus	BB3	8.18	-1.56	0.24	1.55	14.64	6.80	3.19	2.06
Kvantiilisekoituksen sovite	BB3	9.85	-9.05	2.08	0.46	14.73	7.29	3.61	2.48
Kvantiilifunktioiden keskiarvoistus	BB4	36.36	-166.80	151.94	154.00	234.32	28.73	8.87	9.15
Kvantiilisekoituksen sovite	BB4	49.01	-235.94	203.97	130.09	234.23	36.73	11.58	12.33
Kvantiilifunktioiden keskiarvoistus	BB5	23.81	-222.70	276.67	213.73	317.09	29.42	2.02	5.99
Kvantiilisekoituksen sovite	BB5	44.99	-312.64	287.33	203.90	317.53	34.39	7.43	11.32
Kvantiilifunktioiden keskiarvoistus	BB6	7.01	-49.69	128.31	114.06	173.21	19.12	1.13	1.76
Kvantiilisekoituksen sovite	BB6	13.28	-11.90	94.99	107.84	173.27	25.25	4.88	3.34
Kvantiilifunktioiden keskiarvoistus	BB7	9.64	-67.90	86.30	47.10	83.50	11.34	1.56	2.42
Kvantiilisekoituksen sovite	BB7	18.82	-106.04	90.75	42.39	83.45	13.61	3.94	4.74

Liite C: Polynomiset kvantiilisekoitukset - R-funktiot

```
# Functions for solving polynomial quantile mixture -distribution
# parameters, polynomial quantile mixture referred as 'pqm'
# Oskari Luomala
# 5.6.2018

# Calculates inverse of f at given points y.
# Interval is the search area for the inverse value.
# Transforms inverse problem into finding the root of the function
inversef <- function(y, f, interval=c(0,1)) {
  # Function which root is searched
  inv_fkt <- function(y0) {
    f0 <- function(x0){f(x0)-y0}
    uniroot(f0,interval=interval, extendInt="yes")$root
  }
  inv_fkt_na <- function(y0){
    tryCatch(inv_fkt(y0), error = function(e) {0})
  }
  x <- sapply(y, inv_fkt_na)
  as.numeric(x)
}

# Simple polynom function
# x: vector of points,
# param: vector of polynom coefficients
polynom <- function(x, param) {
  n <- length(param)
  k <- 0:(n-1)
  sapply(x, function(x0){sum(param * x0^k)})
}

# Quantile function of pqm
# u: vector of probabilities
# param: vector of coefficients: b, a_k, ..., a0
# q0: base distribution, quantile function
qquantpol <- function(u, param, q0) {
  n <- length(param)
```

```
  b <- param[1]
  Q0_u <- q0(u)
  if(n < 2) return(b * Q0_u) # if there is no polynom pieces
  a_coef  <- param[n:2]
  b * Q0_u + polynom(u, a_coef)
}

# Cumulative density function of pqm, solved numerically
# u: vector of probabilities
# param: vector of coefficients: b, a_k, ..., a0
# q0: base distribution, quantile function
pquantpol <- function(y, param, Q0, ...) {
  Q <- function(u) {qquantpol(u, param, Q0)}
  inversef(y, Q, ...)
}

# Density function of pqm
# u: vector of probabilities
# param: vector of coefficients: b, a_k, ..., a0
# q0: base distribution, quantile function
# f0: base distribution, density function
dquantpol <- function(y, param, q0, f0, ...) {
  n <- length(param)
  b <- param[1]
  u <- pquantpol(y, param, q0, ...)
  d_value <- f0(q0(u))
  # if there is no polynom components
  if(n < 2) return( 1 / (b / d_value))
  da_coef  <- param[(n-1):2] * (1:(n-2))
  pol <- polynom(u, da_coef)
  abs(1 / (b / d_value + pol)  )
}

# Random generation function of pqm
# u: vector of probabilities
# param: vector of coefficients: b, a_k, ..., a0
# q0: base distribution, quantile function
```

```
rquantpol<-function(n, param, Q0) {
  u <- runif(n)
  qquantpol(u, param, Q0)
}

# Gives names for coefficient vector correctly
rename_coefs <- function(param) {
  n <- length(param)
  kmax <- (n-2)
  k <- kmax:0
  a_names <- if(kmax<0) numeric(0) else {paste0("a", k)}
  names(param) <- c("b", a_names)
  param
}

# Gives names for L-moment vector correctly
rename_lmoms <- function(lmoms) {
  n <- length(lmoms)
  k <- 1:n
  names(lmoms) <- paste0("l", k)
  lmoms
}

# Shifted Legendre's polynomial coefficients
# r: polynom order
legendre_P <- function(r) {
  if(r < 0) stop("Negative r specified")
  k <- 0:r
  leg_coefs <- (-1)^(r-k) * choose(r,k) * choose(r+k,k)
  names(leg_coefs) <- paste0("a", k)
  leg_coefs
}

# Calculates values of Legendre's shifted polynomial
# r: polynom order
# u: vector of points, where value is calculated
legendre_Pu <- function(r,u) {
```

```
  k <- 0:r
  leg_coefs <- legendre_P(r)
  sapply(u, function(u_i){sum(leg_coefs * u_i^k)})
}

# Calculates L-moment of given quantile function
# by numerical integration
# Q0: quantile function
# lmom_r: degree of moment
# lower, upper: integration limits
lmom_Q <- function(Q0, lmom_r, lower=0, upper=1, ...) {
  if(lmom_r<1) {
    warning("Order of L-moment should be positive.")
    return(NA)
  }
  Qpol <- function(u) {Q0(u) * legendre_Pu(lmom_r-1, u)}
  lmom <- integrate(Qpol, lower=lower, upper=upper, ...)$value
  names(lmom) <- paste0("l", lmom_r)
  lmom
}

# Simple function to test whether given quantile function
# defines a real distribution
is_quantile <- function(Q, ntest=100, eps=1e-5, ...) {
  u <- seq(eps, 1-eps, length.out=ntest)
  all(diff(Q(u, ...))>0)
}

# Calculates multiple L-moments of given quantile function
# by numerical integration
# Q0: quantile function
# rmax: number of first L-moments to be calculated
lmomQ_n <- function(Q0, rmax, ...) {
  if(rmax<1) stop("n should be >= 1")
  r <- 1:rmax
  lmoms <- sapply(r, function(r_i) {lmom_Q(Q0, lmom_r=r_i)})
  rename_lmoms(lmoms)
```

```
}

# Calculates theoretical L-moment of monomial f(x) = x^k
# lmom_degree: degree of L-moment to be calculated
# mon_degree: degree of monomial
lmom_polynom <- function(lmom_degree, mon_degree) {
  coefs <- legendre_P(lmom_degree - 1)
  d_pol <- (1:lmom_degree) + mon_degree
  sum( coefs/d_pol )
}

# Calculates theoretical L-moments of polynom
# f(x) = x^k + x^(k-1) + ... + x + 1
# lmom_ranks: degrees of L-moment to be calculated
# pol_degree: degree of polynom, k
lmom_u_matrix <- function(l_ranks, pol_degree) {
  if(pol_degree < 0) return(numeric(0))
  l_u <- expand.grid(l_ranks, pol_degree:0)
  lambda_pol <- mapply(lmom_polynom, l_u[,1], l_u[,2])
  matrix(lambda_pol, ncol = pol_degree + 1, byrow = FALSE)
}

# Calculates L-moments of pqm by components
# lmom_r: degrees of L-moments to be calculated
# Q0: quantile function
# pol_degree: degree of polynomial part
lmom_Q_polynom <- function(lmom_r, Q0, pol_degree) {
  if(length(lmom_r)==0) {return(numeric(0))}
  lq0 <- lmomQ_n(Q0, lmom_r)
  lu <-  lmom_u_matrix(lmom_r, pol_degree)
  cbind(lq0,lu)
}


# Solving functions -----------------------------------------------

# Solves pqm coefficients by given quantile function and L-moments
```

```
# lmom: vector of L-moments
# Q0: quantile function
lmom2polycoef <- function(lmom, Q0) {
  n <- length(lmom)
  A <- lmom_Q_polynom(1:n, Q0, n-2)
  param <- solve(A, lmom)
  rename_coefs(param)
}

# Calculates L-moments by given coefficients and quantile function
# of polynomial quantile mixture
# param: vector of coefficients b, a_k, ..., a0
# Q0: quantile function
polycoef2lmom <- function(param, Q0) {
  n <- length(param)
  A <- lmom_Q_polynom(1:n, Q0, n-2)
  lmoms <- as.vector(A %*% param)
  rename_lmoms(lmoms)
}

# Calculates quantile function values of each component
# of polynomial quantile mixture, no coefficients included
# u: vector of quantiles
# pol_degree: polynom degree
# Q0: base distribution, quantile function
calc_qmatrix <- function(u, pol_degree, Q0) {
  Q0_u <- matrix(Q0(u))
  if(pol_degree < 0) return(Q0_u)
  k <- pol_degree:0
  u_pol_matrix <- outer(u,k, FUN="^")
  cbind(Q0_u, u_pol_matrix)
}

# Solves pqm parameters by given quantiles and L-moments
# ux: matrix of quantiles u and equivalent distribution values x
#     as columns
# lmoms: known L-moments of given distribution
```

```
# Q0: base distribution, quantile function
solve_polqm <- function(ux, lmoms, Q0) {
  l_vec <- which(!is.na(lmoms))
  n <- nrow(ux) + length(l_vec)
  pol_degree <- n - 2
  u_vec <- ux[,"u"]
  x_vec <- ux[,"x"]
  b <- c(x_vec, lmoms[l_vec])

  qmatrix <- calc_qmatrix(u_vec, pol_degree, Q0)
  lmatrix <- lmom_Q_polynom(l_vec, Q0, pol_degree)
  A <- rbind(qmatrix, lmatrix)
  param <- solve(A, b)
  rename_coefs(param)
}

# Solves pqm parameters by given quantiles and
# possibly with L-skewness and L-kurtosis values
# ux: matrix of quantiles u and equivalent distribution values x
#     as columns
# Q0: base distribution, quantile function
# tau3: L-skewness - not mandatory
# tau4: L-kurtosis - not mandatory
solve_polqm_tau34 <- function(ux, Q0, tau3=NA, tau4=NA) {
  n <- nrow(ux) + (!is.na(tau3)) + (!is.na(tau4))
  pol_degree <- n-2
  u_vec <- ux[,"u"]
  x_vec <- ux[,"x"]
  # linear equation coefficients from given quantiles
  qmatrix <- calc_qmatrix(u_vec, pol_degree, Q0)
  # linear equation coefs from given L-skewness and L-kurtosis
  l2_coef<- lmom_Q_polynom(2, Q0, pol_degree)
  l3_coef<- lmom_Q_polynom(3, Q0, pol_degree)
  l4_coef<- lmom_Q_polynom(4, Q0, pol_degree)
  l3 <- if(is.na(tau3)) {numeric(0)} else {tau3 * l2_coef - l3_coef}
  l4 <- if(is.na(tau4)) {numeric(0)} else {tau4 * l2_coef - l4_coef}
  # Combining equation coefficients into matrix form
```

```
  A <- rbind(qmatrix, l3, l4)
  b <- c(x_vec,
         if(is.na(tau3)) {numeric(0)} else {0},
         if(is.na(tau4)) {numeric(0)} else {0})
  # Solving
  param <- tryCatch(solve(A, b),
                    error = function(e) {
                      print(e)
                      return(c(1,0,0,0))
                    }
  )
  rename_coefs(param)
}
```