

Testauksella apua tilaryhmien erojen havaitsemiseen

Anne-Mari Sepponen¹⁾, Alina Sinisalo²⁾, Arto Latukka²⁾

¹⁾*Tilastopalvelut, Luonnonvarakeskus (Luke), Kampusranta 9 C, 60320 Seinäjoki, etunimi.sukunimi@luke.fi*

²⁾*Tilastopalvelut, Luonnonvarakeskus (Luke), Latokartanonkaari 9, 00790 Helsinki, etunimi.sukunimi@luke.fi*

Taloustohtorin Maa- ja puutarhatalous -verkkopalvelu (www.luke.fi/taloustohtori/maatalous) on raportointi- ja analyysijärjestelmä, jossa on saatavilla erilaisia maatalouden tunnuslukuja eri luokittelijoittain. Tiedot perustuvat kannattavuuskirjanpidon tuloksiin noin 850 tilalta. Luvut painotetaan koskemaan kaikkia Suomen maa- ja puutarhatiloja, joiden taloudellinen tilakoko on yli 8000 euroa.

Taloustohtorissa näytettävät tunnusluvut ovat keskiarvolukuja. Niiden välinen vertailu on hankalaa, koska pelkkien keskiarvojen perusteella ei voida päätellä, ovatko erot merkitseviä. Ryhmien välillä voi näyttää olevan paljonkin eroa keskiarvoissa, mutta jos lukujen taustalla oleva hajonta on suurta, erot saattavat olla vain normaalia vaihtelevuutta. Ja päinvastoin, pienikin ero saattaa olla merkitsevää, jos luvut normaalisti poikkeavat toisistaan hyvin vähän eli ryhmä on hyvin homogeeninen. Jos tutkittavia ryhmiä on useampia, ei myöskään varmuudella voida päätellä, mitkä ryhmistä eroavat muista ja kuinka paljon.

Keskilukujen eroja voidaan tutkia tilastollisilla testeillä. Taloustohtori-verkkopalveluun on tuotettu analyysijärjestelmä, joka tutkii ja testaa automaattisesti eroavatko eri ryhmien tulokset toisistaan tilastollisesti merkitsevästi. Näin voidaan esimerkiksi tuotantokustannusraportista tarkastella, missä kustannuserissä on todellisia eroja eri tukialueiden välillä. Tilastollinen testaus ja siihen liittyvien oletusten ja ehtojen tutkiminen suoritetaan automaattisesti Taloustohtorissa SAS-ohjelmistolla ja tuotetaan käyttäjälle helposti tulkittavat merkitsevyystasot. Käyttäjä saa halutessaan tuloksista myös lisätietoja, kuten testien tarkat p-arvot sekä käytettyjen testien nimet, viemällä osoittimen rivin päälle. Jos verrattavia ryhmiä on enemmän kuin kaksi, voidaan ryhmiä tutkia tarkemmin parittaisella testillä, jonka käyttäjä saa esiin valitsemalla rivin hiirellä.

Taloustohtori-verkkopalvelua on haluttu kehittää yhä monipuolisemmaksi ja tuottaa konkreettista apua tutkimukselle, päätöksenteolle ja neuvonnalle. Ideana on käyttää vuosittain kerättyä kirjanpitoaineistoa paremmin tuottamalla lisätietoja ja erilaisia analyysimahdollisuuksia Taloustohtorin käyttäjille. Testausjärjestelmän käytettävyyttä ja raportteja on pyritty kehittämään siten, että ne vastaisivat visuaalisesti ja teknisesti parhaiten käyttäjien tarpeisiin.

Asiasanat: tilastollinen testaus, Taloustohtori

Johdanto

Kannattavuuskirjanpidossa seurataan maatalous- ja puutarhayritysten talouden kehitystä noin 850 kirjanpitotalta kerättävän aineiston perusteella. Luvut painotetaan koskemaan kaikkia Suomen maa- ja puutarhatiloja, joiden taloudellinen koko on yli 8000 euroa. Kannattavuuskirjanpito toimintaa ylläpitää ja kehittää Luonnonvarakeskuksen Tilastopalvelut. Saatavaa aineistoa hyödynnetään muun muassa tutkimuksessa, hallinnossa, neuvonnassa ja tiloilla.

Kannattavuuskirjanpidosta saatavia tuloksia on saatavilla Taloustohtorin Maa- ja puutarhatalous - verkkopalvelussa (www.luke.fi/taloustohtori/maatalous). Taloustohtori on raportointi- ja analyysijärjestelmä, jossa on erilaisia maatalouden tunnuslukuja eri luokittelijoittain. Taloustohtorissa näytetään tunnuslukuja muun muassa keskiarvolukuina. Niiden välinen vertailu on hankalaa, koska pelkkien keskiarvojen perusteella ei voida luotettavasti päätellä, ovatko havaittavat erot tilastollisesti merkitseviä. Ryhmien välillä voi näyttää olevan paljonkin eroa keskiarvoissa, mutta jos lukujen taustalla oleva hajonta on suurta, erot saattavat olla vain normaalia vaihtelevuutta. Ja päinvastoin, pienikin ero saattaa olla merkitsevä, jos luvut normaalisti poikkeavat toisistaan hyvin vähän eli ryhmä on hyvin homogeeninen. Keskiarvo on poikkeaville arvoille herkkä keskiluku ja yksittäinenkin poikkeava arvo saattaa vaikuttaa siihen paljon. Jos tutkittavia ryhmiä on useampia, ei myöskään varmuudella voida päätellä, mitkä ryhmistä eroavat muista ja kuinka paljon.

Materiaali ja menetelmät

Ryhmien keskilukujen eroja voidaan tutkia tilastollisilla testeillä, jotta saadaan selville, ovatko erot merkitseviä. Taloustohtorin Maa- ja puutarhatalous -verkkopalveluun on rakennettu analyysijärjestelmä, joka testaa eroavatko eri ryhmien tulokset toisistaan tilastollisesti merkitsevästi. Näin voidaan esimerkiksi tuotantokustannusraportista tarkastella, missä kustannuserissä on todellisia eroja eri tukialueiden välillä.

Keskilukujen tilastollinen testaus vaatii sarjan eri tarkistuksia ja testisuureiden laskemisia. Testaukseen käytetään joko parametrisia tai epäparametrisia testejä, riippuen muun muassa tutkittavan muuttujan jakaumaominaisuuksista. Sopivien testien valintaan vaikuttaa myös ovatko ryhmät keskenään riippumattomia vai riippuvia. Käytännössä eri vuosien väliset ryhmät oletetaan keskenään riippuviksi, koska sama havainto voi esiintyä eri vuosien aineistoissa. Vuosien sisällä eri ryhmät oletetaan riippumattomiksi. Oletuksia tutkitaan ensin omilla testeillään ja edetään sitten sopiviin testeihin tulosten perusteella. Tarkemmin käytetyistä testeistä on kerrottu kirjallisuusviitteissä.

Lähtökohtana erojen merkitsevyyden arvioinnissa on oletus, jonka mukaan ryhmien keskilukujen välillä ei ole eroa. Tilastollisin keinoin lasketaan todennäköisyys lähtöoletuksen hylkäämisen riskistä. Tätä tunnuslukua kutsutaan p-arvoksi. Sitä tarvitaan, koska johtopäätöksen varmuutta ei pystytä sanomaan varmasti, sillä kyseessä on vain otos pääjoukosta. P-arvo on tekninen luku, jolle on yleisesti sovittuja raja-arvoja. Kun p-arvo on pienempi kuin 0,05, mutta suurempi kuin 0,01 sanotaan, että tulos on tilastollisesti melkein merkitsevä. Jos tuolla p-arvolla tulos hyväksytään, eli uskotaan, että tutkittu keskiarvojen välinen ero on todellinen, otetaan korkeintaan 5% riski siitä, että tehdään väärä johtopäätös. Kun p-arvo on pienempi kuin 0,01, mutta suurempi kuin 0,001, puhutaan tilastollisesti merkitsevästä tuloksesta. Pienemmillä p-arvoilla tulos on erittäin merkitsevä ja riski tehdä väärä johtopäätös on korkeintaan 0,1%.

Testaus suoritetaan automaattisesti ja reaaliaikaisesti Taloustohtori-analyysijärjestelmän taustalla toimivalla SAS-tilasto-ohjelmistolla käyttäjän valintojen mukaan. Koodi tutkii tarvittavat taustatiedot, joiden perusteella valitaan kulloiseenkin tilanteeseen sopiva tilastollinen testi. Testeistä toiseen etenemisessä käytetään p-arvojen raja-arvona 0,05.

Tulokset

Testien automaattinen suorittaminen tuottaa käyttäjälle taulukon. Siinä rivit, joiden ryhmien välillä on tilastollisesti merkitsevää eroa, on merkitty tähdin. Erittäin merkitsevä tulos on (***) eli p-arvo on pienempi kuin 0,001, merkitsevä tulos on (**), jolloin p-arvo on edellistä suurempi, mutta pienempi kuin 0,01 ja melkein merkitsevä tulos (*), kun p-arvo on edellisiä suurempi, mutta pienempi kuin 0,05. Jos tulos on suurempi kuin 0,05 tulos ei ole tilastollisesti merkitsevä ().

Esimerkkiraportti on Taulukossa 1 ja parittaistestin tulos Taulukossa 2. Taulukossa 1 on kuvattu lypsykarjatilojen kustannuksia eri alueilla vuonna 2015. Taulukossa jokainen rivi on testattu erikseen ja saatu merkitsevyytasot. Kun osoitin viedään riville, saadaan esiin käytetyn testin nimi, tarkka p-arvo ja sen sanallinen selitys. Taulukossa 1 rivit saavat tähtiä nolasta kolmeen. Monella rivillä C4-alue näyttäisi eroavan silmämääräisesti muista alueista, mutta tulos ei silti ole välttämättä merkitsevä. Tämä kertoo C4-alueen vaihtelusta. Keskiarvotulos pitää sisällään melko vähän lukuja, joiden jakauma saattaa kuitenkin olla hyvinkin laaja, ja siksi ryhmän keskiarvon eroavaisuus muista ryhmistä ei ole tilastollisesti merkitsevä. Osa riveistä saa tuloksen, että alueet eroavat merkitsevästi toisistaan kyseistä muuttujaa tarkasteltaessa. Nähdään esimerkiksi, että pellon vuokra hehtaaria kohden eroaa alueittain tilastollisesti erittäin merkitsevästi. Vielä tästä taulukosta ei kuitenkaan saada tarkemmin selville, minkä alueiden välillä eroa on. Kiinnostuksen kohteena olevan rivin valitsemalla saadaan vielä esiin parittaisen testin tuottama taulukko (Taulukko 2), josta saadaan lisäinformaatiota kyseisestä rivistä.

Taulukko 1. Kustannuksia lypsykarjatiloilta eri alueilla vuonna 2015

() $p \geq 0,05$, (*) $p < 0,05$, (**) $p < 0,01$, (***) $p < 0,001$

Kustannukset	2015						
	Lypsykarja						
	A_alue	B_alue	C1_alue	C2_alue	C2p_alue	C3_alue	C4_alue
Tiloja otoksessa	11<n<20	20<n<30	60<n<70	120<n<130	11<n<20	20<n<30	5<n<10
Viljelyala	101,8	81,9	94,7	85,1	109,9	92,5	28,4
Eläinyksiköt	78,8	64,7	94,6	77,1	77,6	84,9	27,2
TUOTANTOKUSTANNUS (*)	427 562	359 248	483 044	408 030	445 947	503 047	196 072
Tarvikekustannus (*)	87 471	71 902	103 484	91 248	98 436	119 195	46 816
Lannoitteet. kalkitus ()	12 221	10 585	12 340	11 544	7 690	14 671	5 462
Muut kasvinviljelykust. ()	11 595	9 439	11 730	11 029	14 119	14 347	3 206
Polttoaineet ()	12 664	10 272	11 382	9 516	11 163	12 744	4 292
Sähkö ()	7 318	6 310	8 296	7 107	9 971	8 693	4 717
Ostorehukustannukset (*)	43 673	35 295	59 736	52 051	55 492	68 741	29 139
...							
Muu kustannus ()	46 123	42 073	52 487	43 439	46 240	49 075	20 971
Vakuutukset ()	12 946				462	14 717	8 020
Kiinteät vuokrat (**)	12 423	8 475	15 218	7 057	4 691	6 218	628
(josta pellon vuokria) (***)	11 316	7 861	11 049	5 826	4 245	3 452	628
(pellon vuokra hehtaaria kohti) (***)	291	245	273	191	86	87	53
Muut poistot (***)	2 756	1 231	2 624	2 165	1 705	1 181	210
Muut kustannukset ()	17 998	19 366	23 288	20 384	23 382	26 960	12 113

Taulukossa 2 on lypsykarjatilojen pellon vuokra hehtaaria kohden vuonna 2015, siten että alueita on verrattu pareittain. Tilastollista merkitsevyyttä on korostettu eri värein. Vaaleilla alueilla on suurin p-arvo eli kyseiset alue-parit eivät eroa merkitsevästi toisistaan. Ne ovat siis kyseisen muuttujan suhteen samankaltaisia. Punaisilla alueilla on pienin p-arvo eli parin välinen ero on selvän. Esimerkiksi C3-alue eroaa erittäin merkitsevästi A-, B-, C1- ja C2-alueista, mutta ei C2p- ja C4-alueista. Mitä pienempi p-arvo, sitä pienempi väärän johtopäätöksen riski otetaan. Vaikka Taulukosta 1 nähdään, että C4-alueella pellon vuokra hehtaaria kohden on keskimäärin pienin, se ei kuitenkaan parittaistestauksen perusteella eroa yhtä merkitsevästi muista kuin esimerkiksi C3-alue. Tämä johtuu C4-alueen havaintojen suuresta vaihtelusta, jolloin keskiarvo ei ole yksinään tarpeeksi informatiivinen.

Taulukko 2: Pellon vuokra hehtaaria kohden vuonna 2015 lypsykarjatiloiilla eri alueilla ja parittaistestin tuottamat p-arvot. Tilastollista merkitsevyyttä on korostettu eri värein (violetti $p < 0,05$ (*), oranssi $p < 0,01$ (**) ja punainen $p < 0,001$ (***)).

(pellon vuokra hehtaaria kohti)		2015					
		Lypsykarja					
		B_alue	C1_alue	C2_alue	C2p_alue	C3_alue	C4_alue
A_alue	0,780	0,999	0,016	0,005	0,000	0,036	
B_alue		1,000	0,118	0,000	0,000	0,018	
C1_alue			0,048	0,001	0,000	0,019	
C2_alue				0,008	0,000	0,055	
C2p_alue					1,000	0,820	
C3_alue						0,711	

Johtopäätökset

Tilastollisen testauksen sisällyttämisellä raportteihin halutaan tuottaa käyttäjille konkreettista apua Taloustohtorin Maa- ja puutarhatalous -verkkopalvelun tulosten analysointiin. Taulukossa näytettävät tähdet antavat hyvän yleiskuvauksen, mihin ominaisuuksiin kannattaa kiinnittää huomiota ja tarkempi tarkastelu pareittaistestauksella antaa käyttäjälle lisätietoja ryhmien eroavaisuuksista. Eri tilaryhmien väliset erot on helpompi havaita ja parittaistestauksen avulla päästä myös erojen syiden jäljille. Vaikka tähdet kiinnittävät huomion merkitseviin tuloksiin, myös se, että jokin muuttuja ei eroa tilastollisesti merkitsevästi tarkastelluissa ryhmissä, voi olla tärkeä tulos. Tulosten tulkinnassa pitää ottaa huomioon, että tilastollinen merkitsevyys ei kuitenkaan ole aina sama asia kuin eron merkittävyys. Joskus asiayhteydessä epärelevantti ero saattaa osoittautua laskennallisesti tilastollisesti merkitseväksi.

Tavoitteena on käyttää vuosittain kerättävää kirjanpitoaineistoa paremmin tuottamalla lisätietoja ja analyysimahdollisuuksia Taloustohtorin käyttäjille. Testausjärjestelmän käytettävyyttä ja raportteja on pyritty kehittämään siten, että ne vastaisivat visuaalisesti ja teknisesti parhaiten käyttäjien tarpeisiin.

Kirjallisuusuuttelo

Sepponen, A.-M. & Latukka, A. 2012. Statistical testing of differences of means: Adding statistical testing to the Economydoctor –internet service. Teoksessa: The 20th Pacioli Workshop, Rome, Italy. 13 s.

Sinisalo, A., Latukka, A. & Sepponen, A.-M. 2015. Testing differences of means. In: 4th Baltic-Nordic conference on survey statistics, 24-28 August 2015, Helsinki, Finland. University of Helsinki. 1 s.

Sinisalo, A., Sepponen, A.-M. & Latukka, A. 2017. Automated testing of differences of means to promote the effective use of public data. In: EMS 2017: European Meeting of Statisticians Helsinki, July 24-28. Program and Book of Abstracts. Bernoulli Society. 133 s.

Sinisalo, A., Sepponen, A-M. & Latukka, A. 2017. Automated statistical test routines for Finnish agricultural holding bookkeeping results. In: Baltic-Nordic-Ukrainian Workshop on Survey Statistics Theory and Methodology 21-24 August, 2017, Vilnius, Lithuania. Statistics Lithuania. p. 107–111.

Latukka, A., Sepponen, A-M. & Sinisalo, A. 2017. Statistical testing in Economy doctor reporting service. In: *The 25th Pacioli Workshop*. Helsingør, Denmark. 13 p.

Luke 2018. Taloustohtori-verkkopalvelu. www.luke.fi/taloustohtori.