

Statistical testing in Economy doctor reporting service

Arto Latukka, Anne-Mari Sepponen and Alina Sinisalo

Contents

- Background and objectives
- Economy Doctor Internet service
- Testing and reporting
- Conclusion

Background and objectives

(FADN farm level data)



- In Finland more than 100 years tradition in farm bookkeeping.
- Our team is responsible for Farm Accountancy Data Network (FADN) farm-level data in Finland.
- Every year data are collected from approximately 1000 voluntary farms.
- Using the weight factors the data are used to describe the results of all Finnish farms.
- The results are available at Economy Doctor online service.
 - The results are presented mainly on average level by region, economic size and type of agricultural holding.
 - No detailed farm level results due to data privacy decrees.



Background and objectives

(ED and mean values)

- On Economy Doctor (ED) Internet service descriptive statistics reported at average level by classes.
www.luke.fi/economydoctor
- The mean values are sensitive to outlier observations, meaning variables having very different distributions may have approximately the same calculated average.
- These reported averages may be different between groups, but user is not able to distinguish whether these differences are statistically significant or not and which groups differ from each other.
- **The objective of this project was** to more effectively utilize the FADN data by implementing automated calculation and statistical test routines.

Economy Doctor Internet service www.luke.fi/economydoctor



Luke
ECONOMYDOCTOR


Language selection: [Suomeksi](#) [På svenska](#) [In English](#)


EU MEMBER STATES
EU Member State FADN-results

FINLAND

**Agriculture and horticulture ▶**
Results 2015 published 22.12.2016
Forecasts 2016E published 05.05.2017

**Reindeer farming ▶**
Results 2014/15 published 06.06.2016
Forecasts 2015/16E published 06.06.2016

**Unit cost of the reindeer husbandry ▶**
Results 2014/15 published 06.06.2016

**Fur farming ▶**
Results of accounting year 2006


**Forecast of Structural Development ▶**
Structure Development 2000-2022E
Results updated 22.12.2016

**The Structural Development of Agriculture ▶**
Structure Development 2000-2015
Results updated 22.12.2016


**Total Calculation ▶**
Results 2015 published 22.12.2016
Forecasts 2016E published 05.05.2017

**Beekeeping ▶**
Results 2013 published 27.04.2015
Results 2014 published 08.07.2016

**Unit Costs of Beekeeping ▶**
Service published 27.08.2015
Results 2014 published 08.07.2016

**FADN Standard Results (SO) ▶**
Results of EU member countries 2004-2013
Results updated 15.03.2016

**FADN Advanced Results (SO) ▶**
Key ratios calculated of EU results by member countries 2004-2013
Results updated 15.03.2016

**FADN Standard Results (SGM) ▶**
Results of EU member countries 1989-2009
Results updated 15.03.2016

**FADN Advanced Results (SGM) ▶**
Key ratios calculated of EU results by member countries 1989-2009
Results updated 15.03.2016

**Unit costs of agricultural products ▶**
Service published 19.12.2014
Results updated 19.01.2015

**Coastal Fishing ▶**
Service published 27.05.2015
Results 2015 published 16.01.2017

**Marine Fishery ▶**
Service published 03.11.2015
Results 2014 published 02.11.2016

**Aquaculture ▶**
Service published 07.10.2016

**Fish processing and fish trade ▶**
Service published 02.11.2016

Statistical testing

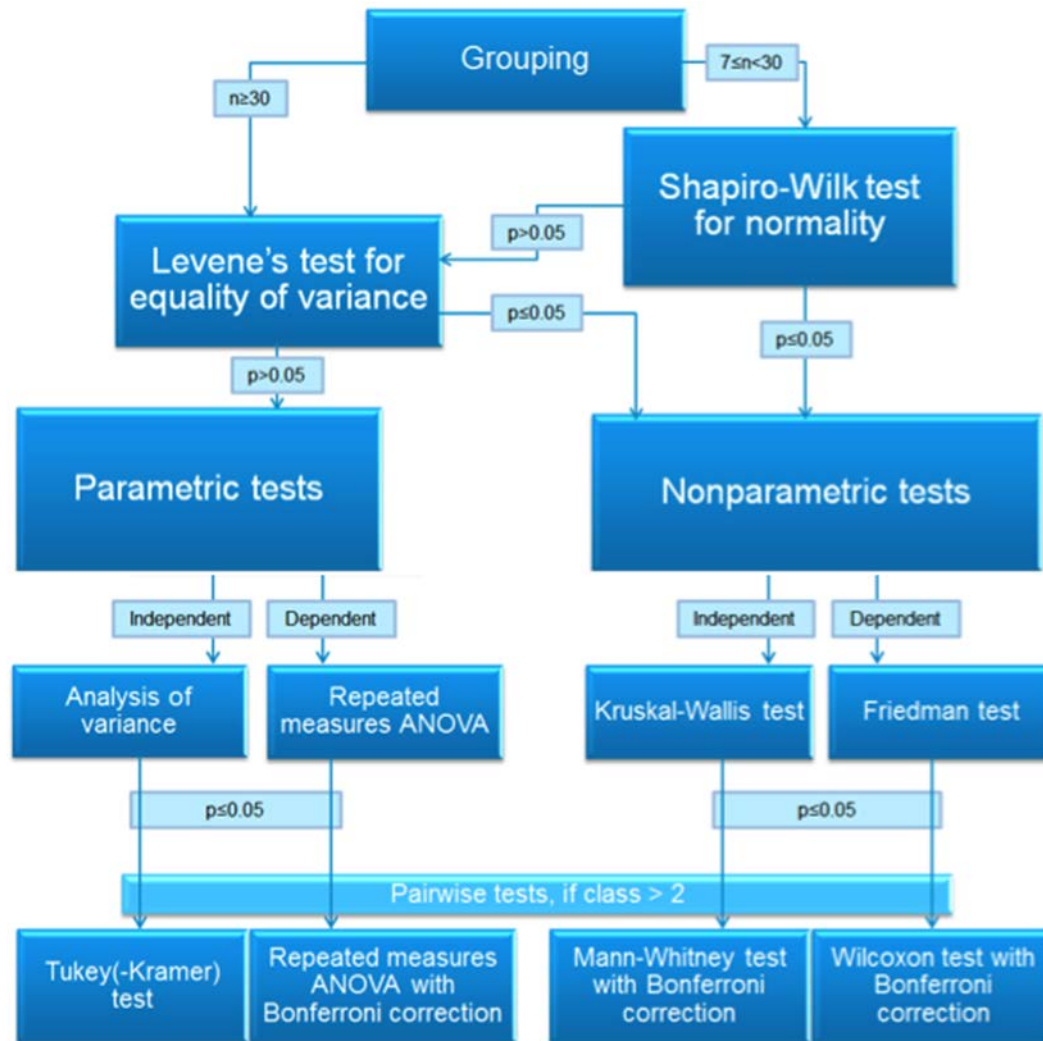
Economy Doctor Internet service www.luke.fi/economydoctor

Production Costs	2014						
	Dairy Farms						
	Subsidy_ Region_A	Subsidy_ Region_B	Subsidy_ Region_C1	Subsidy_ Region_C2	Subsidy_ Region_C2p	Subsidy_ Region_C3	Subsidy_ Region_C4
Farms in sample	11<n<20	30<n<40	70<n<80	130<n<140	11<n<20	30<n<40	5<n<10
Arable land	95,3	70,2	93,6	82,0	95,8	90,3	36,1
Livestock Units	71,9	57,0	91,9	75,3	68,5	79,8	41,2
PRODUCTION COSTS	412 700	321 130	488 035	410 076	403 085	481 551	269 041
Material costs	80 094	62 895	105 214	92 939	92 425	113 296	71 838
Fertilizer. Lime	13 274	9 388	14 234	12 353	8 976	13 721	7 403
Other crop production costs	8 300	8 709	12 617	10 790	14 916	11 728	5 244
Fuel and lubricants	13 006	9 165	13 295	10 364	9 691	13 565	7 733
Electricity	7 443	5 241	7 806	6 979	8 980	8 748	5 952
Forage costs	38 072	30 393	57 263	52 454	49 863	65 533	45 507
Farm use	49 714	36 094	59 025	49 486	49 671	62 090	23 224
Livestock costs	19 088	14 835	23 407	21 218	14 836	22 278	13 547
Livestock purchasing	264	1 172	2 412	2 332	863	2 589	443
Other livestock costs	18 823	13 662	20 996	18 886	13 974	19 689	13 104
Machinery cost	59 656	47 734	82 320	68 300	62 409	89 628	44 021
Depreciation of machines	27 870	25 970	39 071	30 955	26 896	46 531	23 819
Other machinery costs	31 786	21 764	43 249	37 346	35 513	43 097	20 202
Buildings costs	24 425	20 972	33 556	26 047	26 263	33 570	17 611
Depreciation of Buildings	17 231	17 114	28 475	22 046	23 996	30 489	14 222
Other buildings costs	7 193	3 858	5 081	4000	2 267	3 081	3 389
Other cost	44 072	37 423	54 814	43 605	40 357	42 090	22 499
Insurance cost paid	13 485	11 443	13 467	14 111	12 274	13 848	9 418
Fixed rents paid	10 834	8 261	12 436	6 899	3 739	4 502	918
(of which rents paid on land)	9 519	6 654	11 071	5 886	3 481	2 931	729
(rents paid on hectare of land)	288	247	283	205	97	76	42
Other depreciations	2 169	1 098	2 767	2 149	1 509	1 274	195
Other costs	17 585	16 622	26 144	20 445	22 834	22 464	11 969

Testing and reporting

- We developed a system of analyses that automatically examines and reports the differences between groups.
- Test routines compare and test the means of different variable groups to identify if there are statistically significant differences.
- If statistically significant differences are found, test results will be reported on Economy Doctor and annotated as in typical statistical reporting by using asterisks.
- Economy Doctor will run also pairwise testing of differences of means.
- All stages are automated and processed online utilizing micro data.

Flowchart of procedures of obtaining the significance of differences of means



Differences of means by regions in 2014 for dairy farms

(*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$

Production Costs	2014						
	Dairy Farms						
	Subsidy_Region_A	Subsidy_Region_B	Subsidy_Region_C1	Subsidy_Region_C2	Subsidy_Region_C2p	Subsidy_Region_C3	Subsidy_Region_C4
Farms in sample	11<n<20	30<n<40	70<n<80	130<n<140	11<n<20	30<n<40	5<n<10
Arable land	95,3	70,2	93,6	82,0	95,8	90,3	36,1
Livestock Units	71,9	57,0	91,9	75,3	68,5	79,8	41,2
PRODUCTION COSTS (*)	412 700	321 130	488 035	410 076	403 085	481 551	269 041
Material costs (*)	80 094	62 895	105 214	92 939	92 425	113 296	71 838
Fertilizer, Lime ()	13 274	9 388	14 234	12 353	8 976	13 721	7 403
Other crop production costs ()	8 300	8 709	12 617	10 790	14 916	11 728	5 244
Fuel and lubricants ()	13 006	9 165	13 295	10 364	9 691	13 565	7 733
Electricity (*)	7 443	5 241	7 806	6 979	8 980	8 748	5 952
Forage costs (**)	38 072	30 393	57 263	52 454	49 863	65 533	45 507
Farm use (**)							23 224
Livestock costs ()	19 086	14 833	23 407	21 218	14 836	22 278	13 547
Livestock purchasing (*)	264	1 172	2 412	2 332	863	2 589	443
Other livestock costs ()	18 823	13 662	20 996	18 886	13 974	19 689	13 104
Machinery cost (*)	59 656	47 734	82 320	68 300	62 409	89 628	44 021
Depreciation of machines ()	27 870	25 970	39 071	30 955	26 896	46 531	23 819
Other machinery costs (*)	31 786	21 764	43 249	37 346	35 513	43 097	20 202
Buildings costs ()	24 425	20 972	33 556	26 047	26 263	33 570	17 611
Depreciation of Buildings ()	17 231	17 114	28 475	22 046	23 996	30 489	14 222
Other buildings costs ()	7 193	3 858	5 081	4000	2 267	3 081	3 389
Other cost ()	44 072	37 423	54 814	43 605	40 357	42 090	22 499
Insurance cost paid ()	13 485	11 443	13 467	14 111	12 274	13 848	9 418
Fixed rents paid (**)	10 834	8 261	12 436	6 899	3 739	4 502	918
(of which rents paid on land) (***)	9 519	6 654	11 071	5 886	3 481	2 931	729
(rents paid on hectare of land) (***)	288	247	283	205	97	76	42
Other depreciations (***)	2 169	1 098	2 767	2 149	1 509	1 274	195
Other costs ()	17 585	16 622	26 144	20 445	22 834	22 464	11 969

Statistically significant, p-value = 0.0052, test: Kruskal-Wallis

Here is an example of results that users receive as printout in table format.

Extra information can be obtained by moving the mouse over the variable.

The statistical significance, pairwise comparison and the used test criteria are printed out in real time online based on classes defined by user.

Forage costs of dairy farms by regions in 2014, pairwise test results

Forage costs	2014					
	Dairy Farms					
	Subsidy_Region_B	Subsidy_Region_C1	Subsidy_Region_C2	Subsidy_Region_C2p	Subsidy_Region_C3	Subsidy_Region_C4
Subsidy_Region_A	0.998	0.995	0.999	1.000	0.485	1.000
Subsidy_Region_B		0.037	0.030	0.846	0.002	1.000
Subsidy_Region_C1			1.000	1.000	0.958	1.000
Subsidy_Region_C2				1.000	0.659	1.000
Subsidy_Region_C2p					0.999	1.000
Subsidy_Region_C3						0.965

Statistically significant results are reported between regions:

- B and C1 ($p=0.037$),
- B and C2 ($p=0.030$) and
- B and C3 ($p=0.002$).

Conclusion

- The utilization of collected datasets can be promoted by automated calculation routines and statistical reporting without compromising the data privacy.
- This project has resulted in an extension to Agriculture and Horticulture information service in Economy Doctor platform.
- This extension provides to users a means to study the differences of reported accounting data and performance figures at average level within a year or between years by classes (region, economic size, type of holding).
- The aim is to develop Economy Doctor reporting service to produce more concrete help for researchers, decision makers and the public audience.
- The idea is to use the collected data more effectively by producing more information with automated routines.

Related publications and presentations

- Sepponen, A-M., Latukka, A. 2012. Statistical testing of differences of means: Adding statistical testing to the Economydoctor -internet service. In: The 20th Pacioli Workshop, Rome, Italy. 13 p.
- Sinisalo, A., Latukka, A. and 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 p.
- 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. p. 133.
- 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.

Thank you!

arto.latukka@luke.fi

anne-mari.sepponen@luke.fi

alina.sinisalo@luke.fi