

[MTTTA1] TILASTOMENETELMIEN PERUSTEET, KEVÄT 2019

<https://coursepages.uta.fi/mttta1/kevat-2019/>

HARJOITUS 5 viikko 8

RYHMÄT:

ke	12.15–13.45	Is. C6	Leppälä
to	08.30–10.00	Is. C6	Korhonen
to	12.15–13.45	Is. C6	Korhonen
to	14.15–15.45	Is. C8	Leppälä

Nämä ovat kurssin viimeiset laskuharjoitukset. Tentti ti 26.2.2019 klo 12.15- 15.00 Is. A1, ilmoittaudu NettiOpsussa viimeistään 24.2. Tähän tenttiin saa osallistua, jos on tehnyt kurssin aikana vähintään 30 % harjoituksista.

Aihe: Regressioanalyysi, kertausta

1. Ohessa SPSS -harjoitusten 3 tehtävän 1 analyysin liittyviä tuloksia Myynti-aineistosta. Esitä käytetty regressiomalli. Ilmoita estimoinnin tulos. Suorita tarpeelliset testaukset. Määritä selitysprosentti.

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5780,440	2	2890,220	180,355	,000 ^a
	Residual	144,227	9	16,025		
	Total	5924,667	11			

a. Predictors: (Constant), Salesmen, Advertising

b. Dependent Variable: Sale

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	32,277	6,253		5,162	,001
	Advertising	2,506	,329	,452	7,626	,000
	Salesmen	4,759	,410	,688	11,596	,000

a. Dependent Variable: Sale

2. Tutkitaan koululaispojilla, miten pojan ikä (vuosina), paino (kg) ja ponnistusvoima (cm) vaikuttavat Cooperin testin tulokseen. Analysointitulokset ohessa. Esitä käytetty regressiomalli. Ilmoita estimoinnin tulos. Suorita tarpeelliset testaukset. Paljonko keskimäärin juoksee poika, jonka ikä on 15 vuotta, paino 50 kg ja joka ponnistaa 47 cm?

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,581 ^a	,338	,325	284,520

a. Predictors: (Constant), PONNVOIM, PAINO, IKA

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6119720	3	2039906,707	25,199	,000 ^a
	Residual	11980877	148	80951,872		
	Total	18100597	151			

a. Predictors: (Constant), PONNVOIM, PAINO, IKA

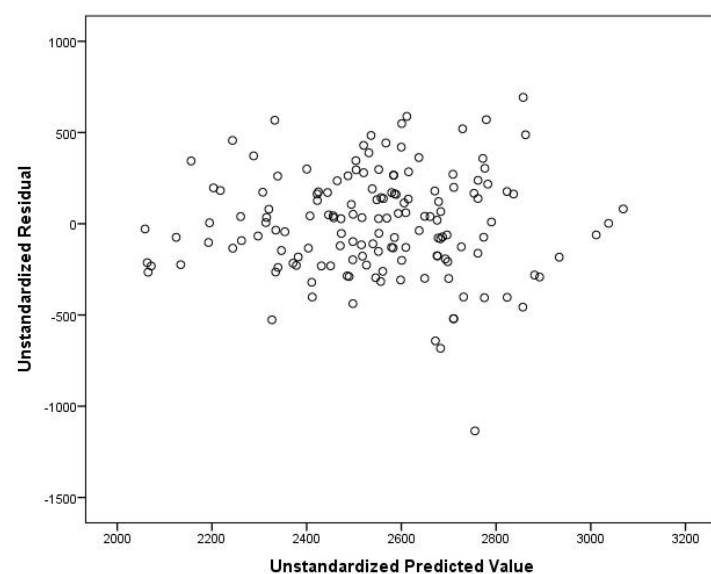
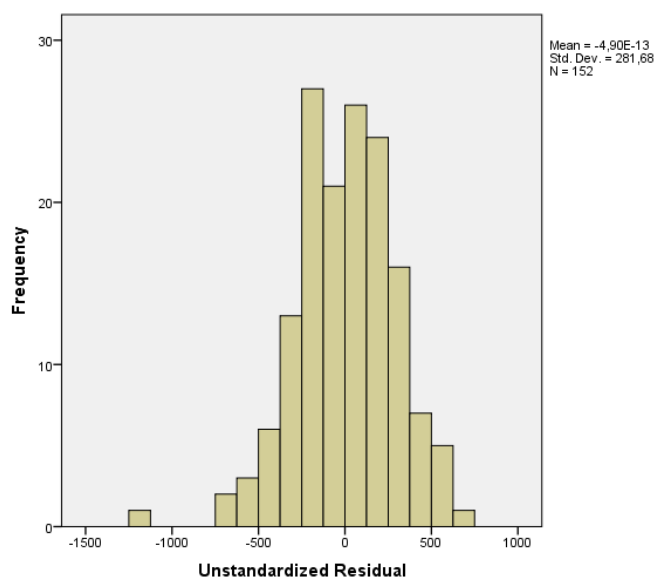
b. Dependent Variable: COOPER

Coefficients^a

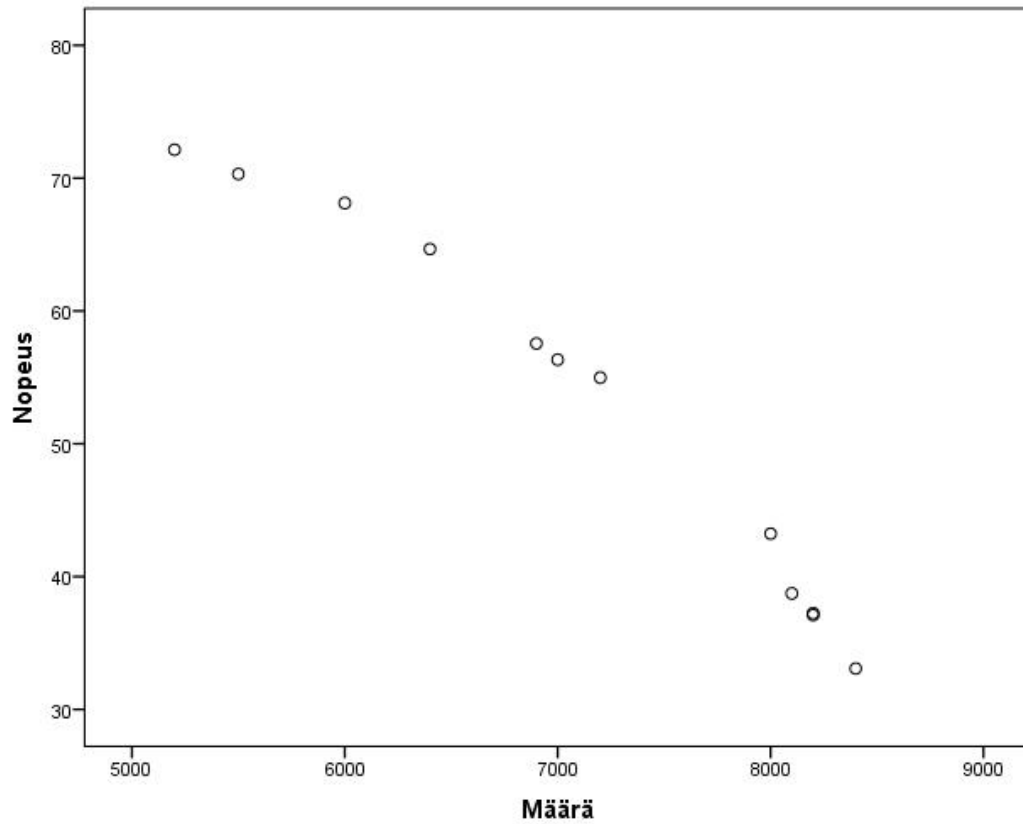
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1642,713	132,788		12,371	,000
	IKA	68,214	16,554	,510	4,121	,000
	PAINO	-13,614	2,633	-,542	-5,170	,000
	PONNVOIM	14,422	3,500	,413	4,121	,000

a. Dependent Variable: COOPER

3. Ohessa on histogrammi ja pisteparvi tehtävän 2 regressioanalyysiin liittyen. Mitä näissä on? Miten niitä käytetään hyväksi regressioanalyysissä? Tulkitse.



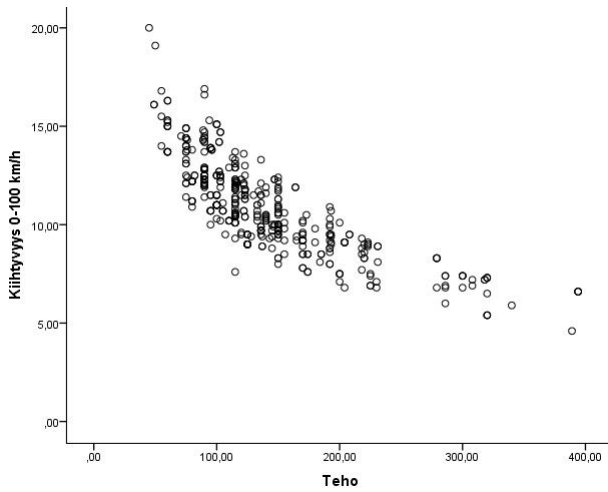
4. Tutkitaan liikennemäärien (Määrä) vaikutusta liikenteen sujuvuuteen (Nopeus). Saadaan oheiset analysointitulokset. Esitä käytetty regressiomalli. Ilmoita estimoinnin tulos. Estimoi nopeus, kun liikennemäärä on 7000.

Coefficients^{a,b}

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	Määrä	,030	,000	3,899	106,161	,000
	Määrä*Määrä	-3,061E-6	,000	-3,005	-81,812	,000

a. Dependent Variable: Nopeus

5. Tutkitaan auton moottorin tehon vaikutusta kiihtyvyyteen. Saadaan oheiset analysointitulokset. Esitä käytetty regressiomalli. Mitkä ovat malliin liittyvät opetukset. Laske F-testisuure. Suorita tarpeelliset testaukset. Voitko olettaa malliin liittyvien oletusten olevan voimassa?

**Model Summary**

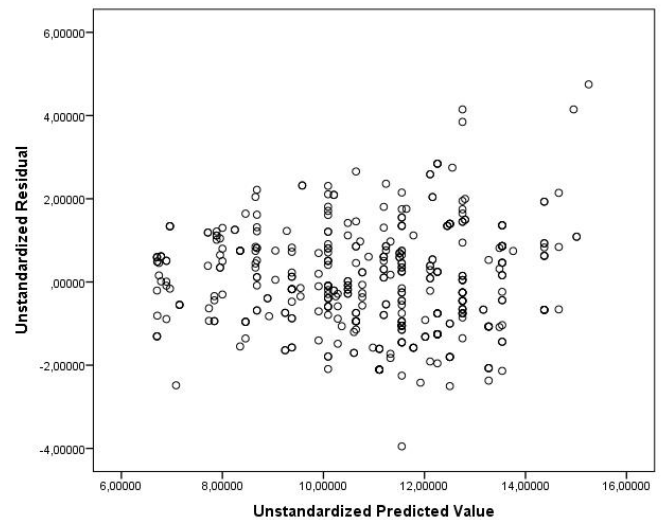
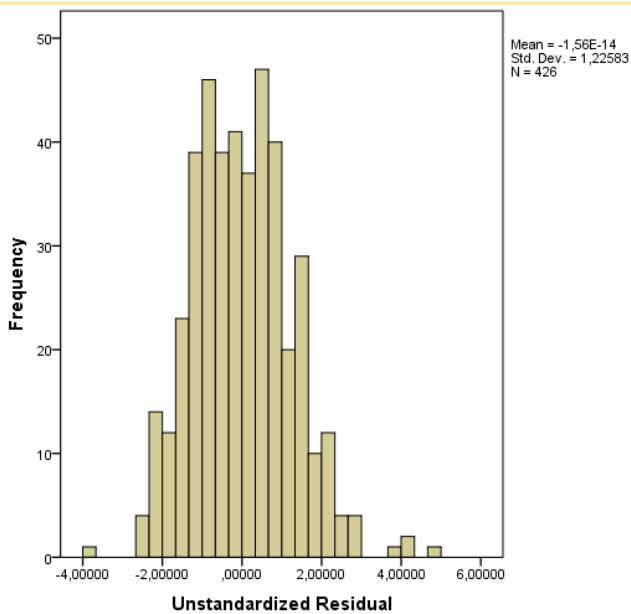
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,853 ^a	,728	,727	1,22872

a. Predictors: (Constant), Teho x Teho, Teho (hv)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	18,17941	,31953		56,894	,000
	Teho	-,06986	,00382	-1,892	-18,274	,000
	Teho*Teho	,00011	,00001	1,118	10,798	,000

a. Dependent Variable: Kiihtyvyys 0-100 km/h



6. Täytä puuttuvat kohdat a) - d) oheisesta regressioanalyysiin liittyvistä estimointituloksista. Suorita testaukset.

Response: y

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	30,736373	3,60818	8,52	0,0000
x1	0,4345565	0,11153	3,90	0,0016
x2	-0,181111	0,07286	-2,49	a) (arvioi)
x4	-0,401229	0,09925	-4,04	0,0012

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	197,85622	65,9521	b)
Error	14	103,25489	7,3753	Prob>F
C Total	17	301,11111		c) (arvioi)

$R^2 = \mathbf{d)}$

7. Tutki, miten kudostiheys muuttuu iän myötä. Käytä varianssianalyysiä ja sivulla <https://coursepages.uta.fi/mttp1/esimerkkiaineistoja/> olevaa Rasvaprosentti-aineistoa http://www.sis.uta.fi/tilasto/tiltp_aineistoja/rasvaprosentti.sav (ks. muuttujien kuvaus http://www.sis.uta.fi/tilasto/tiltp_aineistoja/rasvaprosentti.PDF).

Tee analysointituloksiin liittyvä tehtävä Moodlessa. Tehtävä löytyy kohdasta Harjoitusten 5 palautettavat tehtävät Harjoitus 5, tehtävä 7 (<https://learning2.uta.fi/mod/assign/view.php?id=666475>). Tee tehtävä viimeistään to 21.2. klo 16.

8. Tarkastellaan edellisen tehtävän rasvaprosenttiaineistoa. Muuta pituuden mittayksikkö metreiksi ja painon kilogrammoiksi. Muodosta näiden uusien muuttujien avulla muuttuja painoindeksi = paino/(pituus · pituus). Piirrä painoindeksin histogrammi. Muodosta painon ja pituuden pisteparvi. Huomaat yhden virheellisen havainnon. Jätä se pois jatkotarkasteluista. Tutki, onko painoindeksin ja iän välillä riippuvuutta. Käytä ristiintaulukkoa ja χ^2 -testiä.

Tee analysointituloksiin liittyvä tehtävä Moodlessa. Tehtävä löytyy kohdasta Harjoitusten 5 palautettavat tehtävät Harjoitus 5, tehtävä 8 (<https://learning2.uta.fi/mod/assign/view.php?id=666476>). Tee tehtävä viimeistään to 21.2. klo 16.

9. Tarkastellaan sivulla <https://coursepages.uta.fi/mttp1/esimerkkiaineistoja/> olevaa Audi_A6 – aineistoa. Rakenna hinnalle regressiomalli. Huomioi mallin valinnassa myös residuaalitarkastelut.

Tee analysointituloksiin liittyvä tehtävä Moodlessa. Tehtävä löytyy kohdasta Harjoitusten 5 palautettavat tehtävät Harjoitus 5, tehtävä 9 (<https://learning2.uta.fi/mod/quiz/view.php?id=666477>). Tee tehtävä viimeistään to 21.2. klo 16.