

Korhonen

TENTISSÄ ON LASKIMEN KÄYTTÖ SALLITTU! KYSYMYPAPERISSA ON 2 SIVUA!

1. Vastaa (6p)

- Hahmottele hajontakuviot, joissa muuttujien välinen korrelaatio on i) 1.0, ii) -0.5 ja iii) 0.
- Mitä tarkoitetaan heteroskedastisuudella? Kuinka se havaitaan? Miten heteroskedastisuus vaikuttaa pienimmän neliösumman (PNS) estimointituloksiin?
- Haluat tutkia tupakointia yliopisto-opiskelijoiden keskuudessa binäärisen vastemuuttujan regressiomallin avulla. Hahmottele lyhyesti tutkimuksen kulku erityisesti mallin valinnan osalta.

2. Vastaa lyhyesti (6p)

- Mitkä seikat tulee ottaa huomioon arvioitaessa regressiomallin sisäistä ja ulkoista validiteettia?
- Mihin instrumenttimuuttujilla pyritään? Mitä hyvältä instrumenttimuuttujalta edellytetään?
- Miksi muuttuja(ie)n stationaarisuus (kehityksettömyys) on tärkeä ominaisuus tilastollisen testaamisen kannalta? Anna esimerkki.

3. Tietyn tuotteen kysyntää estimoitiin PNS-menetelmällä ja saatiin regressioyhtälö: (6p)

$$\ln Q = 5.1 - \underset{(0.05)}{0.25} \ln P + \underset{(0.20)}{0.80} \ln Y \quad R^2 = 0.915$$

Q = tuotteen kysyntä, P = tuotteen hinta ja Y = tulotaso

Suluissa olevat luvut kuvaavat regressiokertoimien keskivirheitä. (6p)

- Tulkitse regressiokertoimet ja vakio. Kuinka ne on saatu?
 - Mitkä ovat eksogeenisiä ja mitkä endogeenisiä muuttujia? Mitä eroa niillä on regressiomallin tulkinnassa?
 - Laske hinnan regressiokertoimen 95% luottamusväli. Tulkitse.
 - Kuinka testaat, onko tulomuuttujan regressiokerroin 1? Onko se? Kuinka tulkitset mallia testin jälkeen?
 - Tulkitse selitysaste. Kuinka se on saatu? Mitä on huomioitava selitysastetta tulkittaessa?
 - Mitä voit testata mallissa F-testin avulla? Muodosta nolla- ja vastahypoteesi ja kuvaile F-testin käyttöä (ei tarvitse laskea).
4. Mitä hyötyä vuorovaikutusermien käytöstä on regressiomallien yhteydessä?
- Tehtävän liitteenä on taulukko, jossa on selitettävänä muuttujana työntekijöiden tuntiansiot (AHE) ja selittävinä muuttujina ikä (Age), sukupuoli (Female = 1, jos nainen; 0 = mies) ja tutkinto (Bachelor = 1 jos alempi korkeakoulututkinto, 0 = jos lukio). (12p)
- Miten regressiomallit (1), (2) ja (3) eroavat toisistaan? Mikä niistä on mielestäsi paras malli. Perustele lyhyesti.
 - Onko regressiomalli (4) mielestäsi parempi kuin regressiomallit (1), (2) tai (3). Perustele.
 - Tulkitse regressiomalli (5).
 - Mitä eroa on regressiomalleissa (6), (7) ja (8) muihin malleihin verrattuna? Mihin kysymykseen näillä malleilla pyritään vastaamaan?

- e. Tarkastellaan regressiomallia (6). Jos muuttuja Age nousee arvosta 25 arvoon 26 niin kuinka paljon tuntiansioiden odotetaan muuttuvan? Entä jos se nousee arvosta 33 arvoon 34? Hahmottele vaikutus tuntiansioihin myös muille Age-muuttujan arvoille ja tulkitse tulos.
- f. Tarkastellaan regressiomallia (7). Alexis on 30-vuotias nainen, jolla on alempi korkeakoulututkinto. Mikä on hänen odotettu tuntiansionsa (AHE) mallin mukaan?
- g. Mikä regressiomalleistasi on mielestäsi paras? Tulkitse myös mallisi antamat tulokset? Ovatko sisäinen ja ulkoinen validiteetti valitsemassasi mallissa kunnossa

Data from 2004								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent Variable							
	<i>AHE</i>	$\ln(AHE)$	$\ln(AHE)$	$\ln(AHE)$	$\ln(AHE)$	$\ln(AHE)$	$\ln(AHE)$	$\ln(AHE)$
<i>Age</i>	0.439** (0.030)	0.024** (0.002)		0.147** (0.042)	0.146** (0.042)	0.190** (0.056)	0.117* (0.056)	0.160 (0.064)
<i>Age</i> ²				-0.0021** (0.0007)	-0.0021** (0.0007)	-0.0027** (0.0009)	-0.0017 (0.0009)	-0.0023 (0.0011)
$\ln(Age)$			0.725** (0.052)					
<i>Female</i> × <i>Age</i>						-0.097 (0.084)		-0.123 (0.084)
<i>Female</i> × <i>Age</i> ²						0.0015 (0.0014)		0.0019 (0.0014)
<i>Bachelor</i> × <i>Age</i>							0.064 (0.083)	0.091 (0.084)
<i>Bachelor</i> × <i>Age</i> ²							-0.0009 (0.0014)	-0.0013 (0.0014)
<i>Female</i>	-3.158* * (0.176)	-0.180** (0.010)	-0.180** (0.010)	-0.180** (0.010)	-0.210** (0.014)	1.358* (1.230)	-0.210** (0.014)	1.764 (1.239)
<i>Bachelor</i>	6.865** (0.185)	0.405** (0.010)	0.405** (0.010)	0.405** (0.010)	0.378** (0.014)	0.378** (0.014)	-0.769 (1.228)	-1.186 (1.239)
<i>Female</i> × <i>Bachelor</i>					0.064** (0.021)	0.063** (0.021)	0.066** (0.021)	0.066** (0.021)
Intercept	1.884 (0.897)	1.856** (0.053)	0.128 (0.177)	0.059 (0.613)	0.078 (0.612)	-0.633 (0.819)	0.604 (0.819)	-0.095 (0.945)
<i>F</i> -statistic and <i>p</i> -values on joint hypotheses								
(a) <i>F</i> -statistic on terms involving <i>Age</i>				98.54 (0.00)	100.30 (0.00)	51.42 (0.00)	53.04 (0.00)	36.72 (0.00)
(b) Interaction terms with <i>Age</i> and <i>Age</i> ²						4.12 (0.02)	7.15 (0.00)	6.43 (0.00)
<i>SER</i>	7.884	0.457	0.457	0.457	0.457	0.456	0.456	0.456
\bar{R}^2	0.1897	0.1921	0.1924	0.1929	0.1937	0.1943	0.1950	0.1959

Significant at the *5% and **1% significance level.