

Kommenteret output fra -regress-

Dette notat gennemgår elementerne i outputtet fra Stata regress-kommando. Som eksempel anvendes outputtet fra en multipel regression med variabelen intolerance som afhængig variabel (Y) og v357 (alder), male (køn) og v337 (urbanisering) som uafhængige variable. Regressionen svarer til Opgave 180a i K.M. Sønderskov (2014), *Stata – en praktisk introduktion*, 2. udgave. intolerance er en intervallskaleret variabel (0-100), der tapper respondenternes holdning til fremmede, hvor høje værdier indikerer fremmedskepticisme. v357 indeholder information om respondenternes alder (målt i år), male er dummyvariabel, der antager værdien 1 for mandlige respondenter. v337 er en kategorisk variabel, hvor 1-5 indikerer at respondenter bor hhv. i 1) 'et landdistrikt', 2) 'En by med under 10.000 indbyggere', 3) 'En by med 10.000-50.000 indbyggere', 4) 'En by med 50.001-500.000', 5) Hovedstadsområdet. Regressionen er estimeret med følgende kommando:
 regress intolerance v357 i male i.v337

SS er Sum of Squares, hvor Total SS den er samlede varians $\sum(Y - \bar{Y})^2$, Residual SS er $\sum(Y - \hat{Y})^2$ og Model SS er TotalSS-ResidualSS.

df-kolonnen viser frihedsgraderne for varianskomponenterne. Total df er N-1, Model df er antallet af uafhængige variable (inkl. konstanten) fratrukket 1.

MS viser Mean Squares, der er de respektive SS delt med df. Statistikkerne anvendes i beregningen af F-testen.

. regress intolerance v357 i male i.v337

Source kolonnen viser kilderne til Y's varians: Den del der redegøres for af modellen (Model) og den resterende del (Residual) samt den samlede varians (Total).

Source	SS	df	MS
Model	125415.142	6	20902.5237
Residual	1328644.94	2233	595.004451
Total	1454060.08	2239	649.423886

Number of obs = 2240
 F(6, 2233) = 35.13
 Prob > F = 0.0000
 R-squared = 0.0863
 Adj R-squared = 0.0838
 Root MSE = 24.393

Kolonnen viser, fra toppen, N, F-værdien for hele modellen, P-værdien for F-testen, R², justeret R² og Root Mean Square error (altså kvadratroden af Residual MS)

Navnet på den afhængige variabel

intolerance	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
v357	.2925543	.0305002	9.59	0.000	.2327427 .3523659
male					
Male	3.396117	1.036383	3.28	0.001	1.363742 5.428492
v337					
Town, <10.000	-.7966848	1.499881	-0.53	0.595	-3.737991 2.144621
Town, 10-50.000	-3.769547	1.512256	-2.49	0.013	-6.735122 -.8039723
Town, 50-500.000	-3.827396	1.719994	-2.23	0.026	-7.200351 -.4544408
Metropolitan area	-13.06927	1.636201	-7.99	0.000	-16.27791 -9.860638
_cons	39.61703	1.982949	19.98	0.000	35.72842 43.50565

Listen med modellens uafhængige variable samt konstanten. Da male og v337 er kategoriske (givet i.) vises variabelnavnet og kategorierne (med undtagelse af referencekategorierne)

Fra venstre mod højre vises regressionskoefficienterne, standard fejlene, t-værdien, den tosidede P-værdi og et 95%-konfidensinterval.