

## Outputtet fra regress-kommandoen

Dette notat gennemgår outputtet fra Stata's `regress`-kommando. Outputtet er baseret på en multipel regression med variabelen `fremmedskepticisme` som afhængig variabel (Y) og `v357`, `indfødt`, `mand` og `urban` som uafhængige variable. Regressionen svarer til Opgave 164 i K.M. Sønderkov (2011), *Stata – en praktisk introduktion*, København: Hans Reitzel. Datasættet består af surveydata indsamlet blandt den danske vælgerbefolkning.

`Fremmedskepticisme` 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), `indfødt` og `mand` er dummyvariable der antager værdien 1 for hhv. indfødte og mandlige respondenter. `Urban` 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 estimeres med følgende kommando:

---

```
regress fremmedskepticisme v357 indfødt mand i.urban
```

---

hvoraf det fremgår at der anvendes faktorvariable i forbindelse med variabelen `urban`. Outputtet gennemgås på næste side

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

Sum of Squares, hvor Total SS er den 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.

Kolonnen viser, fra toppen, N, F-værdien for hele modellen, P-værdien for F-testen,  $R^2$ , justeret  $R^2$  og Root Mean Square error (altså kvadratroden af Residual MS)

Source	SS	df	MS
Model	133966.353	7	19138.0504
Residual	1320093.73	2232	591.439842
Total	1454060.08	2239	649.423886

Number of obs = 2240  
 F( 7, 2232) = 32.36  
 Prob > F = 0.0000  
 R-squared = 0.0921  
 Adj R-squared = 0.0893  
 Root MSE = 24.32

Navnet på den afhængige variabel (her en den trunkeret)

fremmedske~e	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
v357	.2883663	.0304286	9.48	0.000	.228695 .3480376
indfødt	10.15026	2.669432	3.80	0.000	4.915435 15.38509
mand	3.487723	1.033555	3.37	0.001	1.460894 5.514552
urban					
2	-.6046456	1.496234	-0.40	0.686	-3.538801 2.329509
3	-3.606408	1.50833	-2.39	0.017	-6.564284 -.6485317
4	-3.436923	1.717906	-2.00	0.046	-6.805784 -.0680614
5	-12.71231	1.633992	-7.78	0.000	-15.91661 -9.508005
_cons	29.82745	3.24607	9.19	0.000	23.46182 36.19308

Listen over uafhængige variable samt konstanten. Reference-kategorien for urban er 1

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