

Projekt nr. 2:

Forudsigelse ud fra regnskabsdata af bankrot og overtagelse for danske virksomheder.

Datasættet til denne opgave er stillet til rådighed af Carmine Gioia, tidligere ph.d. studerende på Institut for International Økonomi og Virksomhedsledelse, som (til brug for sin ph.d. afhandling) med stort besvær har stykket data sammen af rådata fra forskellige kilder.

Datasættet har været brugt før til et MPAS projekt (2004), jeg har haft lidt travlt og har derfor ikke haft tid til at opspore et nyt datasæt. Men da det er et stort og kompliceret datasæt stiller jeg uden tøven opgaven en gang til. Der skal nok vise sig aspekter som er nye for mig, og I kender det jo nok ikke i forvejen.

Data vedrører 13473 danske virksomheders regnskabsdata for perioden ca. 1995-99 samt oplysning om hvorvidt den enkelte på et tidspunkt gik konkurs eller blev overtaget af en anden virksomhed.

Formålet med opgaven er at undersøge i hvor høj grad det er muligt at forudsige konkurs eller overtagelse ud fra regnskabsdata.

Data ligger på filen `07-2-dat.txt`, en flad tekstfil på 31190 linier plus en overskriftsline med variablenavne. Hver linie indeholder 13 oplysninger, som vedrører en given virksomhed et givet år. I listen herunder gengives Carmines korte beskrivelse af de 13 variables betydning. Beskrivelsen vil blive udbygget senere. Til variablene er tilføjet et nyt variabelnavn, som er det der vil blive brugt i undervisningen.

Nyt navn	Opr. navn	Beskrivelse
(Tue)	(Carmine)	
id	kobnumme	ID for each company
kode	nacekode	Industry Code
ansatte	antalans	total number of employees
aar	regnskab	accounting year
oms	oms.tnin	total revenues
brutav	bruttoav	gross profit (brutto avance)
result	result	income (results before tax)
egenkap	egenkapi	equity
balance	samletba	total assets
kapital	selskab2	company capital
respons	target	0 for control group, 1 for target and 2 for bankrupt
slutaar	year	year of acquisition or bankruptcy. If equal to 1 is

because is referred to the control group which has
never been acquired nor has gone bankrupt
respons1 target_1 if sold (1) or bankrupt (2) in the next year.
Control group = 0.

Ved indlæsning kan evt. tilføjes (som i READ.ISU) en variabel

nr Observationsnummer internt for virksomhed (1..5)

Enhed for pengeangivelser er ikke opgivet (formodentlig 1000 kr).

I fortolkningen og den videre transformation af data skal man være opmærksom på følgende:

I 58 tilfælde forekommer der "huller", dvs. overspring af årstal (f.eks. 1995 og 1997 uden 1996). Dette kan eventuelt ignoreres ved konstruktion af laggede variable, da det forekommer så relativt sjældent.

Fordeling af antal observationsår for de 13473 virksomheder:

Antal	Antal
år	virks.
1	2677
2	3951
3	6770
4	74
5	1

I 2 tilfælde forekommer det at en virksomhed går bankrot (respons1=2) men stadig findes året efter. Beskrivelsen af respons1 antyder at dette kan være OK, idet bankrotten så skal placeres året efter.

Til brug for den statistiske analyse er det nødvendigt at danne et "kort" datasæt, dvs. et datasæt hvor der er en linie pr. virksomhed. Et forslag til hvordan det kunne se ud følger her. KORT.ISU dokumenterer mere præcist, hvordan et sådant datasæt kan dannes.

Variabel

navn	beskrivelse
bankrot	1 hvis bankrot, 0 ellers.
solgt	1 hvis overtaget, 0 ellers.
respfac	Faktor på 3 niveauer, 1=kontrol, 2=solgt, 3=bankrot
aar0	sidste observationsår
aar1	næstsidste obsår (oftest aar0-1)
aar2	tredjesidste obsår (oftest aar0-2)
aar0fac	sidste obsår som faktor (1995=1, ... , 1999=5)
lngd	antal år observeret
ansat0	antal ansatte sidste obsår
ansat1	antal ansatte næstsidste obsår
ansat2	antal ansatte tredjesidste obsår
oms0	omsætning
oms1	
oms2	
brutav0	bruttoafkast
brutav1	

```

brutav2
result0      indkomst
result1
result2
egenkap0     egenkapital
egenkap1
egenkap2
balance0     balance
balance1
balance2
kapital0     kapital
kapital1
kapital2
respaar      det år hvor bankrot eller overtagelse
              fandt sted, ellers 1

```

Variabelnavne der ender med et ciffer (0, 1 eller 2) er laggede variable, på samme måde som `ansat0`, `ansat1` og `ansat2`.

Det viser sig, som vi skal se, at det nok ikke er lige præcis på denne måde det “korte” datasæt skal dannes. Men da vi ved undervisningen vil tage udgangspunkt i dette datasæt lægges til dokumentation to ISU-programmer ud på kursets hjemmeside, `READ.ISU` som indlæser det “lange” datasæt og `KORT.ISU` som danner det korte datasæt ud fra det lange.

Projektbesvarelsen afleveres i tre eksemplarer senest mandag d. 17. december kl. 12 på kombi-sekretariatet.