

Support Study

Study description

The *dadossupport.sav* dataset is a random sample of 1000 patients from Phases I & II of SUPPORT Study (Study to Understand Prognoses Preferences Outcomes and Risks of Treatment). See the following reference for more information about the SUPPORT study (funded by the Robert Wood Johnson Foundation): Knaus WA, Harrell FE, Lynn J et al. (1995): The SUPPORT prognostic model: Objective estimates of survival for seriously ill hospitalized adults. *Annals of Internal Medicine* 122:191-203. (http://math.uc.edu/~deddens/yr2005/stat720/support_paper.pdf)

Exercise

For teaching purpose, the main objective of the analysis of SUPPORT data is to investigate which predictors measures during the study is associated with four different response variables (outcomes):

- Binary (hospital death) - *hospdead*
- Ordinal (severe functional disability - see below) - *sfdm2*
- Continuous (total hospital costs) - *charges*
- Continuous (possibly study length of stay) - *slos*

The statistical analysis should include:

1. Estimate a model with main effects.
2. Test whether interaction of interest should be consider in the model.
3. Select covariates using likelihood ratio tests
4. Performance the residual analysis of the final model
5. Write one or two paragraphs including results of the final model. Do not forget to give interpretations for the model coefficients.

You may want to use the following normal values that have been found to be satisfactory in imputing missing baseline physiologic data:

Baseline Variable	Normal Fill-in
Bilirubin	1.01
Creatinine	1.01
Serum albumin	3.5
BUN	6.51
White blood count	9 (thousands)
PaO ₂ /FiO ₂ ratio (pafi)	333.3
Urine output	2502

Challenges

- those who will analyse a binomial response variable should try a probit link function besides the logit link function.
- those who will analyse a continuous outcome should try a generalized linear model with Gamma distribution besides Normal distribution.
- those who will analyse the ordinal outcome should try models with different coefficients for each regression.

Variable dictionary

Name	Labels	Levels	Storage	NAs
age	Age		double	0
death	Death at any time up to NDI date:31DEC94		double	0
sex		2	integer	0
hospdead	Death in Hospital		double	0
slos	Days from Study Entry to Discharge		double	0
d.time	Days of Follow-Up		double	0
dzgroup	Disease groups	8	integer	0
dzclass	Major disease groups	4	integer	0
num.co	number of comorbidities		double	0
edu	Years of Education		double	202
income		4	integer	349
scoma	SUPPORT Coma Score based on Glasgow D3		double	0
charges	Hospital Charges		double	25
totcst	Total RCC cost		double	105
totmcst	Total micro-cost		double	372
avtisst	Average TISS, Days 3-25		double	6
race		5	integer	5
meanbp	Mean Arterial Blood Pressure Day 3		double	0
wblc	White Blood Cell Count Day 3		double	24
hrt	Heart Rate Day 3		double	0
resp	Respiration Rate Day 3		double	0
temp	Temperature (celcius) Day 3		double	0
pafi	PaO2/(.01*FiO2) Day 3		double	253
alb	Serum Albumin Day 3		double	378
bili	Bilirubin Day 3		double	297
crea	Serum creatinine Day 3		double	3
sod	Serum sodium Day 3		double	0
ph	Serum pH (arterial) Day 3		double	250
glucose	Glucose Day 3		double	470
bun	BUN Day 3		double	455
urine	Urine Output Day 3		double	517
adlp	ADL Patient Day 3		double	634
adls	ADL Surrogate Day 3		double	310
sfdm2	functional disability variable	5	integer	159
adlsc	Imputed ADL Calibrated to Surrogate		double	0

Variable levels

Variable	Levels
sex	female male
dzgroup	ARF/MOSF w/Sepsis COPD CHF Cirrhosis Coma Colon Cancer Lung Cancer MOSF w/Malig
dzclass	ARF/MOSF - (acute respiratory failure / multiple organ system failure) Coma Cancer
income	under \$11k \$11-\$25k \$25-\$50k >\$50k
race	white black asian other hispanic
sfdm2	1: no(M2 and SIP pres) - (Patient lived 2 months to be able to get 2 month interview, and from this interview there were no signs of moderate to severe functional disability) 2: adl>=4 (>= 5 if sur) - (Patient was unable to do 4 or more activities of daily living at month 2 after study entry. If the patient was not interviewed but the patient's surrogate was, the cutoff for disability was ADL 5) 3: SIP>=30 - (Sickness Impact Profile total score at 2 months 30) 4: Coma or Intub - (Patient intubated or in coma at 2 months) 5: <2 mo. Follow-up - (Patient died before 2 months after study entry)

To read the data into R proceed as follows:

```
> library(Hmisc)
> load("DadosSupport.sav")
> dim(dados)
```

```
[1] 1000 25
```

```
> head(dados)
```

```

hospdead      sfdm2 slos charges      age d.time      sex
1         0      <NA>   12      NA 85.65594      63   male
2         0      <NA>    8    9914 42.25897     370 female
3         0    SIP>=30  115  706577 43.53998    2022 female
4         0      <NA>    7    8400 45.41800     827   male
5         1 <2 mo. follow-up  14  283303 63.66299      14 female
6         1 <2 mo. follow-up  21  320843 41.52197      21   male

num.co edu   income scoma      race meanbp      wblc hrt resp
1      2  12    <NA>    26   black      97  9.6992188  56   20
2      0  11 $25-$50k    0 hispanic      84 11.2988281  94   20
3      1  NA    <NA>    26   white      67 24.5976562 172   20
4      2  NA    <NA>    0   white      97 10.7988281 100   24
5      0  22 $25-$50k    26   white      69 30.0976562 108   22
6      2  18   >$50k    26   white      66 0.1999817 130   18

      temp      pafi      alb      bili      crea sod      ph
1 36.59375 357.1250      NA 0.3999634 1.0000000 143 7.449219
2 38.19531      NA 4.699219 0.1999817 0.7999268 139      NA
3 38.79688 113.3281      NA      NA 0.5999756 134 7.399414
4 37.29688 585.1250      NA 0.2999878 1.0998535 137 7.489258
5 36.69531 155.5312 2.899902 14.0000000 2.8999023 130 7.449219
6 37.50000 315.0000 1.899902 3.7998047 7.2998047 134 7.359375

glucose dzclass
1      NA   Cancer
2      NA   Cancer
3      NA ARF/MOSF
4      NA   Cancer
5      NA ARF/MOSF
6      NA ARF/MOSF

```

```
> summary(dados)
```

```

hospdead      sfdm2      slos
Min.      :0.000  no(M2 and SIP pres):326  Min.      : 3.00
1st Qu.:0.000  adl>=4 (>=5 if sur):111  1st Qu.: 6.00

```

Median :0.000	SIP>=30	: 59	Median : 11.00
Mean :0.253	Coma or Intub	: 7	Mean : 17.86
3rd Qu.:1.000	<2 mo. follow-up	:338	3rd Qu.: 20.00
Max. :1.000	NA's	:159	Max. :241.00

charges	age	d.time	sex
Min. : 1636	Min. : 18.04	Min. : 3.0	female:438
1st Qu.: 10029	1st Qu.: 51.81	1st Qu.: 27.0	male :562
Median : 26499	Median : 64.90	Median : 256.5	
Mean : 56271	Mean : 62.47	Mean : 475.7	
3rd Qu.: 63622	3rd Qu.: 74.50	3rd Qu.: 725.0	
Max. :740010	Max. :101.85	Max. :2029.0	
NA's : 25			

num.co	edu	income	scoma
Min. :0.000	Min. : 0.00	under \$11k:309	Min. : 0.00
1st Qu.:1.000	1st Qu.: 10.00	\$11-\$25k :161	1st Qu.: 0.00
Median :2.000	Median : 12.00	\$25-\$50k :106	Median : 0.00
Mean :1.886	Mean : 11.78	>\$50k : 75	Mean : 11.74
3rd Qu.:3.000	3rd Qu.: 14.00	NA's :349	3rd Qu.: 9.00
Max. :7.000	Max. : 30.00		Max. :100.00
	NA's :202.00		

race	meanbp	wb1c	hrt
white :781	Min. : 0.00	Min. : 0.050	Min. : 0.00
black :157	1st Qu.: 64.75	1st Qu.: 6.899	1st Qu.: 72.00
asian : 9	Median : 78.00	Median : 10.449	Median :100.00
other : 12	Mean : 84.98	Mean : 12.397	Mean : 97.87
hispanic: 36	3rd Qu.:107.00	3rd Qu.: 15.500	3rd Qu.:120.00
NA's : 5	Max. :180.00	Max. :100.000	Max. :300.00
		NA's : 24.000	

resp	temp	pafi	alb
Min. : 0.00	Min. :32.50	Min. : 34.0	Min. : 1.100
1st Qu.:18.00	1st Qu.:36.20	1st Qu.:156.3	1st Qu.: 2.400
Median :24.00	Median :36.70	Median :226.7	Median : 2.800
Mean :23.49	Mean :37.08	Mean :244.2	Mean : 2.917
3rd Qu.:29.00	3rd Qu.:38.09	3rd Qu.:310.0	3rd Qu.: 3.400
Max. :64.00	Max. :41.20	Max. :869.4	Max. : 4.899
		NA's :253.0	NA's :378.000

bili	crea	sod
Min. : 0.09999	Min. : 0.3000	Min. :118.0
1st Qu.: 0.50000	1st Qu.: 0.8999	1st Qu.:134.0
Median : 0.79993	Median : 1.2000	Median :137.0

```

Mean      : 2.52701   Mean      : 1.8076   Mean      :137.7
3rd Qu.: 1.79980   3rd Qu.: 1.8999   3rd Qu.:141.0
Max.      : 50.09375   Max.      :11.7988   Max.      :175.0
NA's      :297.00000   NA's      : 3.0000

      ph              glucose              dzclass
Min.    : 6.960   Min.    : 11.0   ARF/MOSF              :477
1st Qu.: 7.380   1st Qu.:100.0   COPD/CHF/Cirrhosis:314
Median  : 7.420   Median :128.0   Coma                  : 60
Mean    : 7.416   Mean    :156.4   Cancer                :149
3rd Qu.: 7.470   3rd Qu.:185.0
Max.    : 7.670   Max.    :598.0
NA's    :250.000   NA's    :470.0

```

```
> describe(dados)
```

```
dados
```

```

25 Variables      1000 Observations
-----
hospdead : Death in Hospital
      n missing  unique      Sum      Mean
      1000      0      2      253      0.253
-----
sfadm2
      n missing  unique
      841      159      5

no(M2 and SIP pres) (326, 39%), adl>=4 (>=5 if sur) (111, 13%)
SIP>=30 (59, 7%), Coma or Intub (7, 1%), <2 mo. follow-up (338, 40%)
-----
slos : Days from Study Entry to Discharge
      n missing  unique      Mean      .05      .10      .25      .50
      1000      0      88      17.86      4      4      6      11
      .75      .90      .95
      20      37      53

lowest :   3   4   5   6   7, highest: 145 164 202 236 241
-----
charges : Hospital Charges

```

```

      n missing  unique    Mean    .05    .10    .25    .50
    975         25     967  56271   3757   4688   10029   26499
    .75         .90     .95
  63622  147109  223582

lowest :   1636   1680   1830   2045   2082
highest: 504660 538323 543761 706577 740010
-----
age : Age
      n missing  unique    Mean    .05    .10    .25    .50
   1000         0     970   62.47   33.76   38.91   51.81   64.90
    .75         .90     .95
   74.50   81.87   86.00

lowest :   18.04   18.41   19.76   20.30   20.31
highest:   95.51   96.02   96.71  100.13  101.85
-----
d.time : Days of Follow-Up
      n missing  unique    Mean    .05    .10    .25    .50
   1000         0     582   475.7    5.0    8.0   27.0   256.5
    .75         .90     .95
   725.0  1464.3  1757.1

lowest :     3     4     5     6     7, highest: 2006 2011 2022 2024 2029
-----
sex
      n missing  unique
   1000         0         2

female (438, 44%), male (562, 56%)
-----
num.co : number of comorbidities
      n missing  unique    Mean
   1000         0         8   1.886

      0    1    2    3    4    5    6    7
Frequency 120 337 269 151 76 29 17 1
%         12  34  27  15   8   3   2  0
-----
edu : Years of Education
      n missing  unique    Mean    .05    .10    .25    .50

```


798	202	25	11.78	6	8	10	12
.75	.90	.95					
14	16	18					

lowest : 0 1 2 3 4, highest: 20 21 22 24 30

income

n missing	unique
651	349
	4

under \$11k (309, 47%), \$11-\$25k (161, 25%), \$25-\$50k (106, 16%)
>\$50k (75, 12%)

scoma : SUPPORT Coma Score based on Glasgow D3

n missing	unique	Mean	.05	.10	.25	.50
1000	0	11	11.74	0.0	0.0	0.0
.75	.90	.95				
9.0	44.0	62.4				

	0	9	26	37	41	44	55	61	89	94	100
Frequency	704	83	58	24	19	45	11	6	8	7	35
%	70	8	6	2	2	4	1	1	1	1	4

race

n missing	unique
995	5
	5

	white	black	asian	other	hispanic
Frequency	781	157	9	12	36
%	78	16	1	1	4

meanbp : Mean Arterial Blood Pressure Day 3

n missing	unique	Mean	.05	.10	.25	.50
1000	0	122	84.98	47.00	55.00	64.75
.75	.90	.95				
107.00	120.00	128.05				

lowest : 0 20 27 30 32, highest: 155 158 161 162 180

wblc : White Blood Cell Count Day 3

n missing	unique	Mean	.05	.10	.25	.50
-----------	--------	------	-----	-----	-----	-----

976	24	282	12.4	2.475	4.800	6.899	10.449
.75	.90	.95					
15.500	22.248	27.524					

lowest : 0.05000 0.06999 0.09999 0.14999 0.19998
highest: 60.00000 61.19531 67.09375 79.39062 100.00000

hrt : Heart Rate Day 3

n missing	unique	Mean	.05	.10	.25	.50
1000	0	124	97.87	54.0	60.0	72.0
.75	.90	.95				
120.0	135.0	146.1				

lowest : 0 11 30 35 36, highest: 189 193 199 232 300

resp : Respiration Rate Day 3

n missing	unique	Mean	.05	.10	.25	.50
1000	0	45	23.49	9	10	18
.75	.90	.95				
29	36	40				

lowest : 0 4 6 7 8, highest: 48 49 52 60 64

temp : Temperature (celcius) Day 3

n missing	unique	Mean	.05	.10	.25	.50
1000	0	64	37.08	35.50	35.80	36.20
.75	.90	.95				
38.09	38.80	39.20				

lowest : 32.50 33.70 34.00 34.09 34.40

highest: 40.20 40.59 40.90 41.00 41.20

pafi : PaO2/(.01*FiO2) Day 3

n missing	unique	Mean	.05	.10	.25	.50
747	253	463	244.2	92.61	115.00	156.33
.75	.90	.95				
310.00	400.00	442.81				

lowest : 34.00 39.00 45.00 48.00 53.33

highest: 600.00 623.75 640.00 680.00 869.38

alb : Serum Albumin Day 3

n missing	unique	Mean	.05	.10	.25	.50
622	378	38	2.917	1.800	2.000	2.400 2.800
.75	.90	.95				
3.400	4.000	4.199				

lowest : 1.100 1.200 1.300 1.400 1.500

highest: 4.500 4.600 4.699 4.800 4.899

bili : Bilirubin Day 3

n missing	unique	Mean	.05	.10	.25	.50
703	297	115	2.527	0.3000	0.3000	0.5000 0.7999
.75	.90	.95				
1.7998	5.8594	12.5896				

lowest : 0.09999 0.19998 0.29999 0.39996 0.50000

highest: 32.39844 33.00000 35.00000 39.29688 50.09375

crea : Serum creatinine Day 3

n missing	unique	Mean	.05	.10	.25	.50
997	3	87	1.808	0.6000	0.7000	0.8999 1.2000
.75	.90	.95				
1.8999	3.6396	5.5996				

lowest : 0.3 0.4 0.5 0.6 0.7, highest: 10.4 10.6 11.2 11.6 11.8

sod : Serum sodium Day 3

n missing	unique	Mean	.05	.10	.25	.50
1000	0	42	137.7	129	131	134 137
.75	.90	.95				
141	145	148				

lowest : 118 120 121 122 123, highest: 156 157 158 168 175

ph : Serum pH (arterial) Day 3

n missing	unique	Mean	.05	.10	.25	.50
750	250	53	7.416	7.289	7.319	7.380 7.420
.75	.90	.95				
7.470	7.500	7.520				

lowest : 6.960 6.989 7.069 7.119 7.130

highest: 7.590 7.600 7.609 7.659 7.670

glucose : Glucose Day 3

n	missing	unique	Mean	.05	.10	.25	.50
530	470	226	156.4	74.0	82.0	100.0	128.0
.75	.90	.95					
185.0	269.3	327.5					

lowest : 11 25 30 42 51, highest: 492 512 528 576 598

dzclass

n	missing	unique
1000	0	4

ARF/MOSF (477, 48%), COPD/CHF/Cirrhosis (314, 31%), Coma (60, 6%)
 Cancer (149, 15%)
