

ModelCAI

KS(Kolmogorov-Smirnov):

ks用于模型风险区分能力进行评估，指标衡量的是好坏样本累计分部之间的差值。

好坏样本累计差异越大，ks指标越大，那么模型的风险区分能力越强。

$$ks = \max(ks_i) = \max\left(\frac{Cum.B_i}{Bad_{total}} - \frac{Cum.G_i}{Good_{total}}\right) \quad (1)$$

```
DATA TA;  
SET SASHELP.CARS;  
IF CYLINDERS IN (4,6) THEN K = 0;  
ELSE K = 1;  
KEEP MSRP K;  
RUN;  
  
PROC PRINT DATA = TA(OBS = 20);  
RUN;
```

```
SAS Connection established. Subprocess id is 10751
```

Obs	MSRP	K
1	\$36,945	0
2	\$23,820	0
3	\$26,990	0
4	\$33,195	0
5	\$43,755	0
6	\$46,100	0
7	\$89,765	0
8	\$25,940	0
9	\$35,940	0
10	\$31,840	0
11	\$33,430	0
12	\$34,480	0
13	\$36,640	0
14	\$39,640	0
15	\$42,490	0
16	\$44,240	0
17	\$42,840	0
18	\$49,690	1
19	\$69,190	1
20	\$48,040	1

```

PROC SORT DATA = TA;
  BY MSRP;
RUN;

```

```

49  ods listing close;ods html5 (id=saspy_internal) file=stdout
options(bitmap_mode='inline') device=svg style=HTMLBlue; ods
49 ! graphics on / outputfmt=png;
NOTE: Writing HTML5(SASPY_INTERNAL) Body file: STDOUT
50
51  PROC SORT DATA = TA;
52  BY MSRP;
53  RUN;
NOTE: There were 428 observations read from the data set WORK.TA.
NOTE: The data set WORK.TA has 428 observations and 2 variables.
NOTE: PROCEDURE SORT used (Total process time):
      real time          0.00 seconds
      cpu time           0.00 seconds

54
55  ods html5 (id=saspy_internal) close;ods listing;

56

```

```

PROC RANK DATA = TA OUT = TB GROUPS = 5;
VAR MSRP;
RANKS GMSRP;
RUN;

PROC PRINT DATA = TB(OBS = 20);
RUN;

```

Obs	MSRP	K	GMSRP
1	\$10,280	0	0
2	\$10,539	0	0
3	\$10,760	0	0
4	\$10,995	0	0
5	\$11,155	0	0
6	\$11,290	0	0
7	\$11,560	0	0
8	\$11,690	0	0
9	\$11,839	0	0
10	\$11,905	0	0
11	\$11,939	0	0
12	\$12,269	0	0
13	\$12,360	0	0
14	\$12,585	0	0
15	\$12,740	0	0
16	\$12,800	0	0
17	\$12,884	0	0
18	\$12,965	0	0
19	\$13,270	0	0
20	\$13,270	0	0

```

PROC FREQ DATA = TB NOPRINT;
TABLE  GMSRP * K/OUT = TC;
RUN;

PROC PRINT DATA = TC(OBS = 20);
RUN;

```

Obs	GMSRP	K	COUNT	PERCENT
1	0	0	84	19.6262
2	0	1	1	0.2336
3	1	0	83	19.3925
4	1	1	3	0.7009
5	2	0	75	17.5234
6	2	1	11	2.5701
7	3	0	60	14.0187
8	3	1	26	6.0748
9	4	0	24	5.6075
10	4	1	61	14.2523

```

/* NPARTWAY KS*/
PROC NPARTWAY DATA = TC KS NOPRINT;
CLASS K;
VAR  GMSRP;
FREQ COUNT;
OUTPUT OUT = TD;
RUN;

PROC PRINT DATA = TD;
RUN;

```

The SAS System

Obs	_VAR_	_KS_	_KSA_	_D_	P_KSA	_CM_	_CMA_	_K_	_KA_	P_KA
1	GMSRP	0.25362	5.24690	0.59527	0	0.033333	14.2663	0.59527	5.24690	0

```

DATA T_KS;
SET TD;
KEEP _D_;
RUN;

PROC PRINT DATA = T_KS;
RUN;

```

The SAS System

Obs	_D_
1	0.59527

--