

欠測の種類	主な補完方法	補助変数の利用
項目の欠測	層化平均値代入	標本の層化
	回帰代入	説明変数
	比率補完	比率の計算
	ホットデック法	標本の距離の計算 傾向スコアの算出など
	LOCF	伸び率の計算
ユニット単位の欠測	ウェイト調整	解答確率を計算するクラス、 傾向スコアの算出

比率補完

資本金1,350億円の法人Aが設備投資金額を未提出
 →他の法人のデータから設備投資を見積もる

未提出法人	資本金 (億円)	設備投資 (億円)	設備投資／資本金
A	1,350	?	
A	1,350	188	0.139

③

提出法人	資本金 (億円)	設備投資 (億円)	設備投資／資本金
a	9,000	38	0.004
b	3,200	600	0.188
c	3,100	610	0.197
d	2,100	300	0.143
e	2,000	550	0.275
f	1,400	240	0.171
g	1,050	90	0.086
h	230	22	0.096
i	200	62	0.310
k	140	0	0.001
m	100	0	0.000
o	72	8	0.111
p	71	1	0.014
q	70	2	0.029

4社

②

平均 0.139

5社

- ① 資本金上位から降順に並べて、未提出法人の資本金以下の基準法人を決める
- ② 基準法人を含めて上位4社と下位5社の計10社の(設備投資／資本金)比率の平均値を算出
- ③ 求めた平均比率×資本金＝設備投資を算出する

①
基準法人

ホットデック法

ID	x_1 売上	x_2 仕入	x_3 経費	給料
1	5000	150	3050	0
2	6325	273	4797	1238
3	4173	225	2759	
4	2500	135	882	0
5	985	0	358	0
6	6610	970	3628	900
7	3908	175	1323	0
平均	4214.4	275.4	2399.6	

ID	D^2	D
1	2214760	1,488
2	9014196	3,002
3	1538785	1,240
4	6536349	2,557
5	40587632	6,371
6	23622857	4,860
7	5895810	2,428

$$D = \sqrt{T(x_i - \bar{x})V^{-1}(x_i - \bar{x})}$$

$$= \sqrt{(x_1 - \bar{x} \quad x_2 - \bar{x} \quad x_3 - \bar{x})V^{-1} \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix}}$$

マハラノビス距離Dに近いID1のデータ(給料)を採用

$x_i - \bar{x}$

	$x_1 - \bar{x}$	$x_2 - \bar{x}$	$x_3 - \bar{x}$
1	785.6	-125.4	650.4
2	2110.6	-2.4	2397.4
3	-41.4	-50.4	359.4
4	-1714.4	-140.4	-1517.6
5	-3229.4	-275.4	-2041.6
6	2395.6	694.6	1228.4
7	-306.4	-100.4	-1076.6

$$V^{-1} \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix}$$

	1	2	3	4	5	6	7
1	4613.464	3745.17	-4369	-7203.5	-22592.19	17831	7975.4
2	-2035.64	-2906.1	964.99	3082.56	7873.8488	-5263	-1717.1
3	-2559.5	459.949	3913	3545.57	14794.269	-12567	-7586.4

転置

$$\begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \end{pmatrix}$$

	1	2	3	4	5	6	7
1	785.6	2110.6	-41.4	-1714.4	-3229.4	2395.6	-306.4
2	-125.4	-2.4	-50.4	-140.4	-275.4	694.6	-100.4
3	650.4	2397.4	359.4	-1517.6	-2041.6	1228.4	-1076.6

=CORREL(売上,仕入)

=CORREL(経費,仕入)

	売上	仕入	経費
売上	1	0.71	0.93
仕入	0.71	1	0.53
経費	0.93	0.53	1

逆行列 $V \Rightarrow V^{-1}$

14.4	-4.3	-11.1
-4.3	2.7	2.6
-11.1	2.6	10.0

性質が近いデータをドナーとして選び、そのデータを採用

LOCF (Last Observation Carried Forward)

過去からの変化率を加味して算出

例 過去値 × 比率

ID	売上(2023年)	売上(2024年)
1	5,300	5,000
2	6,705	6,325
3	4,131	
4	2,650	2,500
5	1,044	985
6	7,007	6,610
7	4,142	3,908



ID	売上(2023年)	売上(2024年)
1	5,300	5,000
2	6,705	6,325
3	4,131	3,897
4	2,650	2,500
5	1,044	985
6	7,007	6,610
7	4,142	3,908

$$3,897 = 4,131 \times 0.94$$

$$0.94 \leftarrow \text{比率} = \frac{\text{ID3を除いた2024年の売上合計金額}}{\text{ID3を除いた2023年の売上合計金額}}$$