

数量化 II 類

青木繁伸

1 目的

数量化 II 類を行う。

2 使用法

```
from qt2 import qt2
qt2(dat, verbose=True)
```

2.1 結果の図示

```
from qt2 import qt2_plot
qt2_plot(obj, type="cs", ax1=1, ax2=2, color="black", color2="blue")
```

2.2 引数

<code>dat</code>	群変数が最後の列にあるデータフレーム
<code>verbose</code>	必要最小限のプリント出力をする (デフォルトは <code>True</code>)
<code>mute</code>	サンプルスコアをプリント出力しない (デフォルトは <code>True</code>)
<code>obj</code>	<code>pfa()</code> の戻り値
<code>type</code>	<code>type="cs"</code> (デフォルト) ならカテゴリースコア, <code>type="ss"</code> ならサンプルスコアをプロットする。
<code>ax1</code>	横軸にとる解 (軸) の番号
<code>ax2</code>	縦軸にとる解 (軸) の番号
<code>color</code>	点の色
<code>color2</code>	点のそばに付けるテキストの色

2.3 戻り値の名前

<code>"ndim"</code>	解の個数
<code>"ng"</code>	群の数
<code>"dat2"</code>	ダミー変数行列
<code>"cs"</code>	カテゴリースコア
<code>"partialCor"</code>	偏相関係数
<code>"centroid"</code>	群の重心
<code>"eta"</code>	η
<code>"ss"</code>	サンプルスコア
<code>"name"</code>	カテゴリーの名前

3 使用例

3.1 2 群判別

```
import pandas as pd

dat = pd.read_csv("data/qt2-2.csv")

import sys
sys.path.append("statlib")
from qt2 import qt2

a = qt2(dat)
```

Quantification theory type II

Category score

	Axis1
X1-0	-0.342384
X1-1	-0.162955
X1-2	0.490476
X2-0	0.216705
X2-1	-0.216705
X3-0	-0.335986
X3-1	0.322810
X4-0	-0.099126
X4-1	-0.145571
X4-2	0.226555
X5-0	0.552722
X5-1	-0.194830
X5-2	-0.325379
X6-0	-0.182115
X6-1	0.182115
X7-0	-0.220064
X7-1	-0.110143
X7-2	0.320495
X8-0	0.135827
X8-1	-0.135827
X9-0	-0.110795
X9-1	-0.005234
X9-2	0.112616
X10-0	-0.257227
X10-1	0.178965
X10-2	0.075961
X11-0	-0.017642
X11-1	0.098339
X11-2	-0.085145

X12-0 -0.259156

X12-1 0.259156

Partial correlation coefficient

Axis1

X1 0.351399

X2 0.218060

X3 0.312902

X4 0.171104

X5 0.366614

X6 0.181836

X7 0.244922

X8 0.151114

X9 0.097154

X10 0.192320

X11 0.082377

X12 0.283096

Centroid

Axis1

Y-0 -0.764807

Y-1 0.764807

Eta

Axis1

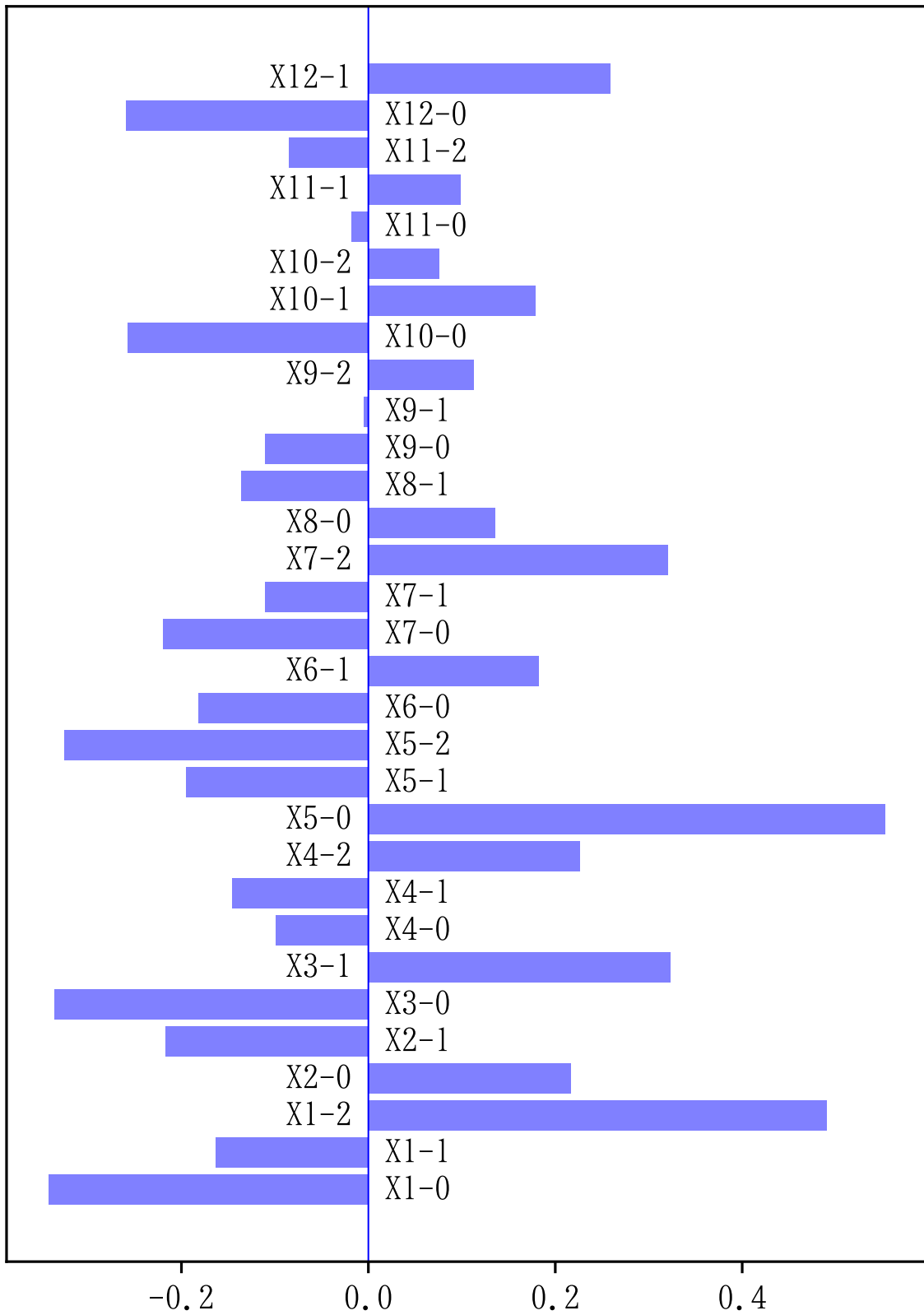
Eta 0.58493

カテゴリースコアの描画

```
from qt2 import qt2_plot
```

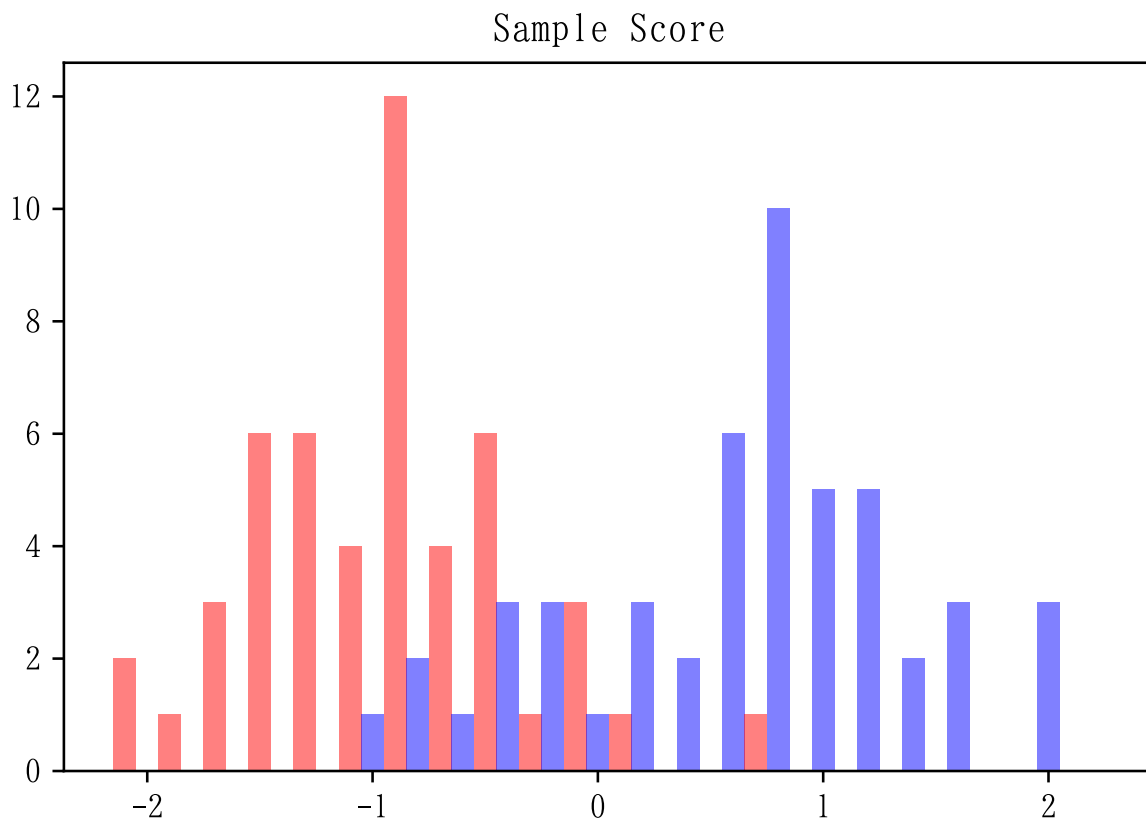
```
qt2_plot(a)
```

Category Score



サンプルスコアの描画

```
qt2_plot(a, type="ss")
```



3.2 3 群以上の判別

```
import pandas as pd

dat = pd.read_csv("data/qt2-4.csv")

import sys
sys.path.append("statlib")
from qt2 import qt2

a = qt2(dat)
```

Quantification theory type II

Category score

	Axis1	Axis2	Axis3
X1-0	0.525172	-0.652550	-0.810571
X1-1	0.008195	0.601504	0.542635
X1-2	-0.517680	0.049544	0.260055
X2-0	-0.199716	-0.141598	-0.349254
X2-1	0.199716	0.141598	0.349254
X3-0	0.381506	-0.105831	-0.088603
X3-1	-0.366545	0.101681	0.085129
X4-0	0.254702	-0.436475	0.283177
X4-1	0.131217	0.051951	-0.108272
X4-2	-0.360118	0.364036	-0.168004

X5-0	-0.327124	-0.850146	-0.294191
X5-1	0.001474	0.663518	0.112568
X5-2	0.306408	0.136620	0.164317
X6-0	0.175197	0.086573	0.234837
X6-1	-0.175197	-0.086573	-0.234837
X7-0	0.119885	0.288439	-0.424764
X7-1	0.129916	-0.084734	-0.254832
X7-2	-0.242454	-0.197713	0.659608
X8-0	-0.102479	-0.092272	0.229884
X8-1	0.102479	0.092272	-0.229884
X9-0	0.214948	-0.346377	-0.209466
X9-1	-0.102444	0.356616	0.084257
X9-2	-0.109195	-0.009938	0.121527
X10-0	0.317666	-0.139486	0.082419
X10-1	-0.121392	-0.203414	0.052118
X10-2	-0.190501	0.332815	-0.130580
X11-0	0.117019	-0.218105	0.623824
X11-1	-0.177660	0.172325	-0.423178
X11-2	0.076191	0.021467	-0.133157
X12-0	0.102985	0.528845	-0.059428
X12-1	-0.102985	-0.528845	0.059428

Partial correlation coefficient

	Axis1	Axis2	Axis3
X1	0.560949	0.208641	0.199812
X2	0.307555	0.059187	0.121880
X3	0.499548	0.040509	0.028083
X4	0.377460	0.124335	0.061158
X5	0.373754	0.249400	0.066295
X6	0.263271	0.033458	0.075698
X7	0.277659	0.083436	0.159488
X8	0.168231	0.037746	0.079318
X9	0.238705	0.119765	0.049747
X10	0.342435	0.099225	0.032048
X11	0.217673	0.065338	0.143488
X12	0.176788	0.228401	0.021943

Centroid

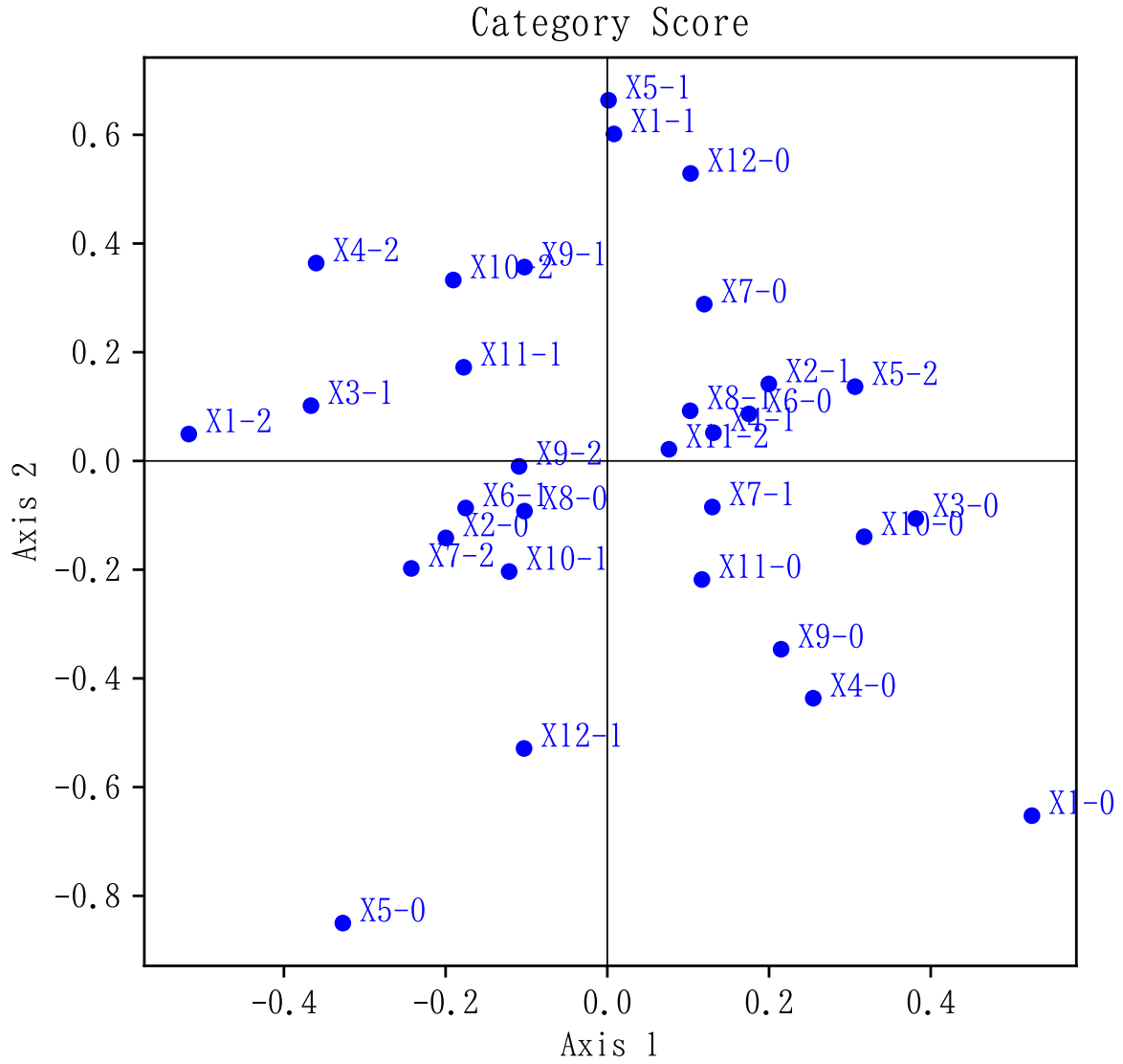
	Axis1	Axis2	Axis3
Y-0	1.210123	-0.190107	-0.368542
Y-1	0.284705	0.615681	0.272748
Y-2	-0.218617	-0.518098	0.425546
Y-3	-1.239194	0.060292	-0.355404

Eta

Axis1 Axis2 Axis3
Eta 0.768379 0.175245 0.12879

カテゴリースコアの描画

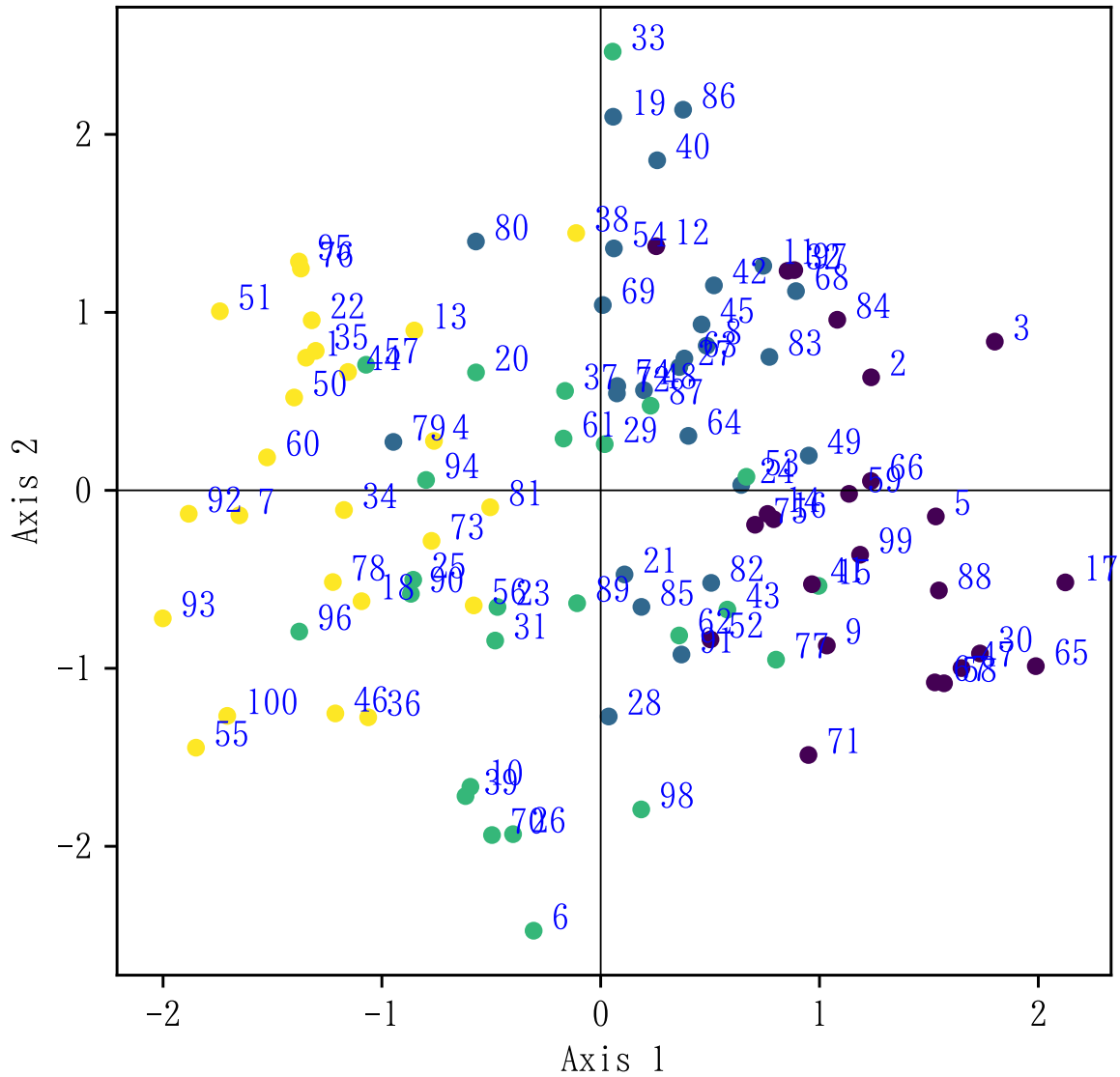
```
from qt2 import qt2_plot  
  
qt2_plot(a)
```



サンプルスコアの描画

```
qt2_plot(a, type="ss")
```

Sample Score



```
import pandas as pd

x1 = [3, 3, 3, 2, 2, 3, 2, 1]
x2 = [2, 1, 2, 1, 1, 1, 1, 2]
group = [1, 2, 2, 1, 2, 3, 3, 3]
dat = pd.DataFrame({"x1": x1, "x2": x2, "group": group})

import sys
sys.path.append("statlib")
from qt2 import qt2

a = qt2(dat, mute=False)
```

Quantification theory type II

Category score

	Axis1	Axis2
x1-1	-3.140993	-0.267697
x1-2	0.917694	-1.195270


```
x1-3 0.096978 0.963377
x2-1 -0.712515 0.655608
x2-2 1.187525 -1.092681
```

Partial correlation coefficient

```
      Axis1    Axis2
x1 0.630249 0.239844
x2 0.513893 0.203824
```

Centroid

```
      Axis1    Axis2
group-1 0.744841 -0.334483
group-2 0.291382 0.316673
group-3 -0.787942 -0.093685
```

Eta

```
      Axis1    Axis2
Eta 0.403355 0.068867
```

Sample score

```
      Axis1    Axis2
0 1.284503 -0.129304
1 -0.615537 1.618985
2 1.284503 -0.129304
3 0.205179 -0.539662
4 0.205179 -0.539662
5 -0.615537 1.618985
6 0.205179 -0.539662
7 -1.953468 -1.360378
```

データは整数値でなくてもよいが、辞書順で割り振られる。

```
x1 = ["hi", "hi", "hi", "med", "med", "hi", "med", "lo"]
x2 = ["female", "male", "female", "male", "male", "male", "male", "female"]
group = ["Treatment2", "Treatment1", "Treatment1", "Treatment2",
         "Treatment1", "Control", "Control", "Control"]
dat = pd.DataFrame({"x1": x1, "x2": x2, "group": group})

a = qt2(dat, mute=False)
```

Quantification theory type II

Category score

```
      Axis1    Axis2
x1-hi -0.096978 0.963377
x1-lo  3.140993 -0.267697
x1-med -0.917694 -1.195270
```

x2-female -1.187525 -1.092681
x2-male 0.712515 0.655608

Partial correlation coefficient

	Axis1	Axis2
x1	0.630249	0.239844
x2	0.513893	0.203824

Centroid

	Axis1	Axis2
group-Control	0.787942	-0.093685
group-Treatment1	-0.291382	0.316673
group-Treatment2	-0.744841	-0.334483

Eta

	Axis1	Axis2
Eta	0.403355	0.068867

Sample score

	Axis1	Axis2
0	-1.284503	-0.129304
1	0.615537	1.618985
2	-1.284503	-0.129304
3	-0.205179	-0.539662
4	-0.205179	-0.539662
5	0.615537	1.618985
6	-0.205179	-0.539662
7	1.953468	-1.360378