

1116即席作業

社會三 109204035 黃楷捷

2022-11-22

Used Packages

```
library(haven)
library(ggplot2)
library(data.table)
```

Tab function

```
Tab <- function(data, var1, code, var2) {
  data <- as.data.frame(data)
  i <- which(names(data) == var1)
  j <- which(names(data) == var2)
  n <- nrow(data[data[, i] == code, ])
  tableA <- table(data[data[, i] == code, j])
  Table1 <- cbind(Freq. = tableA,
                  Percent = round(tableA / n * 100, 3),
                  Cum. = round(cumsum(tableA) / n * 100, 3)) |> data.frame()
  Table1 <- rbind(Table1, Total = c(sum(Table1$Freq.), round(sum(Table1$Percent), 0), ''))
  return(Table1)
}
```

Import data / Data Preprocessing

```
tscs212 <- read_dta(".././../HW/tscs212.dta") |> setDT()
tscs212 <- (tscs212[!(e24a > 5 | e24b > 5 | e24c > 5 | e24d > 5 | e24e > 5 | e24f > 5 | e24g > 5 | e24h > 5), ],
              ":", MH = e24a + e24b + e24c + e24d + e24e)])
data1 <- tscs212[, .(e24a, e24b, e24c, e24d, e24e, e24f, e24g, e24h, MH = as.factor(MH))]
```

Correlation Coefficient

```
cor(data1[, .(e24a, e24b, e24c, e24d, e24e)])

##           e24a           e24b           e24c           e24d           e24e
## e24a 1.0000000 0.6115170 0.4619064 0.4120341 0.4444317
## e24b 0.6115170 1.0000000 0.4601933 0.3485148 0.3591879
## e24c 0.4619064 0.4601933 1.0000000 0.7116856 0.6128380
## e24d 0.4120341 0.3485148 0.7116856 1.0000000 0.6960577
## e24e 0.4444317 0.3591879 0.6128380 0.6960577 1.0000000
```

Item Analysis

MH

```
tableA <- table(data1[, MH])
tableB <- cumsum(tableA)
Table1 <- cbind(Freq. = tableA,
                Percent = round(tableA / nrow(data1) * 100, 2),
                Cum. = round(tableB / nrow(data1) * 100, 2)) |> data.frame()
Table1 <- rbind(Table1, Total = c(sum(Table1$Freq.), sum(Table1$Percent), ''))
Table1
```

##	Freq.	Percent	Cum.
## 5	349	22.07	22.07
## 6	118	7.46	29.54
## 7	152	9.61	39.15
## 8	138	8.73	47.88
## 9	139	8.79	56.67
## 10	159	10.06	66.73
## 11	118	7.46	74.19
## 12	101	6.39	80.58
## 13	100	6.33	86.91
## 14	57	3.61	90.51
## 15	54	3.42	93.93
## 16	21	1.33	95.26
## 17	26	1.64	96.9
## 18	15	0.95	97.85
## 19	8	0.51	98.36
## 20	6	0.38	98.73
## 21	5	0.32	99.05
## 22	5	0.32	99.37
## 23	4	0.25	99.62
## 24	4	0.25	99.87
## 25	2	0.13	100
## Total	1581	100.01	

指數的分佈集中在低分項，尤其5最多。

Internal Validation

e24a

```
Tab(data1, 'e24a', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	349	37.406	37.406
## 6	107	11.468	48.875
## 7	120	12.862	61.736
## 8	105	11.254	72.99
## 9	103	11.04	84.03
## 10	56	6.002	90.032
## 11	37	3.966	93.998
## 12	22	2.358	96.356
## 13	23	2.465	98.821
## 14	7	0.75	99.571
## 15	1	0.107	99.678
## 16	1	0.107	99.786
## 17	1	0.107	99.893
## 18	0	0	99.893
## 19	0	0	99.893
## 20	1	0.107	100
## 21	0	0	100
## 22	0	0	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	933	100	

e24a等於1時的分佈情形與指數一致，對指數建構有貢獻，具有內在鑑別力。

e24b

```
Tab(data1, 'e24b', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	349	56.472	56.472
## 6	62	10.032	66.505
## 7	66	10.68	77.184
## 8	56	9.061	86.246
## 9	25	4.045	90.291
## 10	22	3.56	93.851
## 11	17	2.751	96.602
## 12	11	1.78	98.382
## 13	1	0.162	98.544
## 14	4	0.647	99.191
## 15	1	0.162	99.353
## 16	2	0.324	99.676
## 17	1	0.162	99.838
## 18	1	0.162	100
## 19	0	0	100
## 20	0	0	100
## 21	0	0	100
## 22	0	0	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	618	100	

e24b等於1時的分佈情形與指數一致，對指數建構有貢獻，具有內在鑑別力。

e24c

```
Tab(data1, 'e24c', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	349	56.84	56.84
## 6	88	14.332	71.173
## 7	89	14.495	85.668
## 8	34	5.537	91.205
## 9	25	4.072	95.277
## 10	9	1.466	96.743
## 11	9	1.466	98.208
## 12	6	0.977	99.186
## 13	1	0.163	99.349
## 14	2	0.326	99.674
## 15	1	0.163	99.837
## 16	1	0.163	100
## 17	0	0	100
## 18	0	0	100
## 19	0	0	100
## 20	0	0	100
## 21	0	0	100
## 22	0	0	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	614	100	

e24c等於1時的分佈情形與指數一致，對指數建構有貢獻，具有內在鑑別力。

e24d

```
Tab(data1, 'e24d', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	349	42.15	42.15
## 6	112	13.527	55.676
## 7	133	16.063	71.739
## 8	73	8.816	80.556
## 9	55	6.643	87.198
## 10	42	5.072	92.271
## 11	34	4.106	96.377
## 12	11	1.329	97.705
## 13	11	1.329	99.034
## 14	4	0.483	99.517
## 15	2	0.242	99.758
## 16	1	0.121	99.879
## 17	1	0.121	100
## 18	0	0	100
## 19	0	0	100
## 20	0	0	100
## 21	0	0	100
## 22	0	0	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	828	100	

e24d等於1時的分佈情形與指數一致，對指數建構有貢獻，具有內在鑑別力。

e24e

```
Tab(data1, 'e24e', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	349	44.066	44.066
## 6	103	13.005	57.071
## 7	117	14.773	71.843
## 8	71	8.965	80.808
## 9	57	7.197	88.005
## 10	35	4.419	92.424
## 11	35	4.419	96.843
## 12	10	1.263	98.106
## 13	12	1.515	99.621
## 14	0	0	99.621
## 15	1	0.126	99.747
## 16	0	0	99.747
## 17	1	0.126	99.874
## 18	0	0	99.874
## 19	0	0	99.874
## 20	0	0	99.874
## 21	1	0.126	100
## 22	0	0	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	792	100	

e24e等於1時的分佈情形與指數一致，對指數建構有貢獻，具有內在鑑別力。

External Validation

e24f

```
Tab(data1, 'e24f', 5, 'MH')
```

##	Freq.	Percent	Cum.
## 5	331	44.016	44.016
## 6	90	11.968	55.984
## 7	108	14.362	70.346
## 8	66	8.777	79.122
## 9	56	7.447	86.569
## 10	35	4.654	91.223
## 11	30	3.989	95.213
## 12	16	2.128	97.34
## 13	12	1.596	98.936
## 14	6	0.798	99.734
## 15	1	0.133	99.867
## 16	1	0.133	100
## 17	0	0	100
## 18	0	0	100
## 19	0	0	100
## 20	0	0	100
## 21	0	0	100
## 22	0	0	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	752	100	

e24f等於5時的分佈情形與指數一致，具有外在鑑別力。（本題編碼方向不同）

e24g

```
Tab(data1, 'e24g', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	169	46.301	46.301
## 6	39	10.685	56.986
## 7	56	15.342	72.329
## 8	27	7.397	79.726
## 9	22	6.027	85.753
## 10	17	4.658	90.411
## 11	13	3.562	93.973
## 12	7	1.918	95.89
## 13	10	2.74	98.63
## 14	3	0.822	99.452
## 15	1	0.274	99.726
## 16	0	0	99.726
## 17	0	0	99.726
## 18	0	0	99.726
## 19	0	0	99.726
## 20	0	0	99.726
## 21	0	0	99.726
## 22	1	0.274	100
## 23	0	0	100
## 24	0	0	100
## 25	0	0	100
## Total	365	100	

e24g等於1時的分佈情形與指數一致，具有外在鑑別力。

e24h

```
Tab(data1, 'e24h', 1, 'MH')
```

##	Freq.	Percent	Cum.
## 5	109	52.913	52.913
## 6	19	9.223	62.136
## 7	23	11.165	73.301
## 8	16	7.767	81.068
## 9	11	5.34	86.408
## 10	10	4.854	91.262
## 11	6	2.913	94.175
## 12	7	3.398	97.573
## 13	2	0.971	98.544
## 14	1	0.485	99.029
## 15	1	0.485	99.515
## 16	0	0	99.515
## 17	0	0	99.515
## 18	0	0	99.515
## 19	0	0	99.515
## 20	0	0	99.515
## 21	0	0	99.515
## 22	0	0	99.515
## 23	0	0	99.515
## 24	1	0.485	100
## 25	0	0	100
## Total	206	100	

e24h等於1時的分佈情形與指數一致，具有外在鑑別力。