

## Дисциплина: Анализ данных

ФИО студента: Гаврилина Александра Александровна

### Тема: Факторный анализ

Откройте файл **gmu.dta**

Список и описание переменных:

Variables		⌵ ⌴ ×
Filter variables here		
Name	Label	
name	Наименование администратора средств ...	
grade	Общая оценка (в баллах)	
ind1	Среднесрочное финансовое планирован...	
ind2	Исполнение бюджета в части расходов	
ind3	Исполнение бюджета по доходам	
ind4	Учет и отчетность	
ind5	Контроль и аудит	
ind6	Исполнение судебных актов	
ind7	Кадровый потенциал	
ind8	Управление активами	

Проведите факторный анализ со всеми переменными кроме name и grade.

#### 1) Интерпретируйте результаты расчёта КМО.

Строим факторную модель

**factor ind1 ind2 ind3 ind4 ind5 ind6 ind7 ind8, pcf**

```
. factor ind1 ind2 ind3 ind4 ind5 ind6 ind7 ind8, pcf  
(obs=77)
```

Factor analysis/correlation	Number of obs	=	77
Method: principal-component factors	Retained factors	=	3
Rotation: (unrotated)	Number of params	=	21

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	2.34029	1.07811	0.2925	0.2925
Factor2	1.26217	0.07399	0.1578	0.4503
Factor3	1.18818	0.29163	0.1485	0.5988
Factor4	0.89656	0.15961	0.1121	0.7109
Factor5	0.73694	0.14739	0.0921	0.8030
Factor6	0.58955	0.08989	0.0737	0.8767
Factor7	0.49965	0.01300	0.0625	0.9392
Factor8	0.48666	.	0.0608	1.0000

LR test: independent vs. saturated:  $\chi^2(28) = 80.95$  Prob> $\chi^2 = 0.0000$

# Factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
ind1	0.7241	-0.0458	-0.3241	0.3685
ind2	0.5595	-0.2479	0.3689	0.4894
ind3	0.7067	0.1883	-0.2792	0.3872
ind4	0.7247	-0.2701	0.1122	0.3892
ind5	0.4081	-0.0615	-0.5220	0.5571
ind6	0.2219	0.5844	0.4810	0.3778
ind7	0.4724	0.5531	0.3207	0.3680
ind8	0.1983	-0.6625	0.4998	0.2720

estat kmo, novar

Kaiser-Meyer-Olkin measure of sampling adequacy (overall) = 0.6708

Тест КМО позволяет оценить, в какой мере построенная факторная модель полно описывает структуру значений исходных переменных для рассматриваемых единиц анализа.

В данном случае:  $0.6 < \text{КМО} < 0.7$  – результат **mediocre** (посредственный).

## 2) Укажите процент вариации исходных переменных, объяснённой моделью.

Фактор 1: 0.2925

Фактор 2: 0.1578

Фактор 3: 0.1485

```
. factor ind1 ind2 ind3 ind4 ind5 ind6 ind7 ind8, pcf
(obs=77)
```

Factor analysis/correlation	Number of obs	=	77
Method: principal-component factors	Retained factors	=	3
Rotation: (unrotated)	Number of params	=	21

Factor	Eigenvalue	Difference	Proportion	Cumulative
Factor1	2.34029	1.07811	0.2925	0.2925
Factor2	1.26217	0.07399	0.1578	0.4503
Factor3	1.18818	0.29163	0.1485	0.5988
Factor4	0.89656	0.15961	0.1121	0.7109
Factor5	0.73694	0.14739	0.0921	0.8030
Factor6	0.58955	0.08989	0.0737	0.8767
Factor7	0.49965	0.01300	0.0625	0.9392
Factor8	0.48666	.	0.0608	1.0000

LR test: independent vs. saturated:  $\chi^2(28) = 80.95$  Prob> $\chi^2 = 0.0000$

### 3) Опишите и придумайте названия для факторов.

Выведем матрицу факторных нагрузок после вращения:

**rotate, blanks(.3)**

```
. rotate, blanks(.3)
```

```
Factor analysis/correlation          Number of obs   =          77
Method: principal-component factors   Retained factors =           3
Rotation: orthogonal varimax (Kaiser off) Number of params =          21
```

Factor	Variance	Difference	Proportion	Cumulative
Factor1	2.01039	0.61019	0.2513	0.2513
Factor2	1.40020	0.02015	0.1750	0.4263
Factor3	1.38005	.	0.1725	0.5988

```
LR test: independent vs. saturated:  chi2(28) =    80.95 Prob>chi2 = 0.0000
```

Rotated factor loadings (pattern matrix) and unique variances

Variable	Factor1	Factor2	Factor3	Uniqueness
ind1	0.7850			0.3685
ind2		0.6162		0.4894
ind3	0.7469			0.3872
ind4	0.5514	0.5381		0.3892
ind5	0.6239			0.5571
ind6			0.7857	0.3778
ind7			0.7612	0.3680
ind8		0.8382		0.2720

(blanks represent abs(loading)<.3)

Factor rotation matrix

	Factor1	Factor2	Factor3
Factor1	0.8448	0.3925	0.3637
Factor2	0.0028	-0.6828	0.7306
Factor3	-0.5351	0.6162	0.5779

**Фактор 1:** переменные

ind1 (Среднесрочное финансовое планирование),

ind3 (Исполнение бюджета по доходам),

ind4 (Учет и отчетность),

ind5 (Контроль и аудит)

**Фактор 2:** переменные

ind2 (Исполнение бюджета в части расходов),

ind4 (Учет и отчетность),

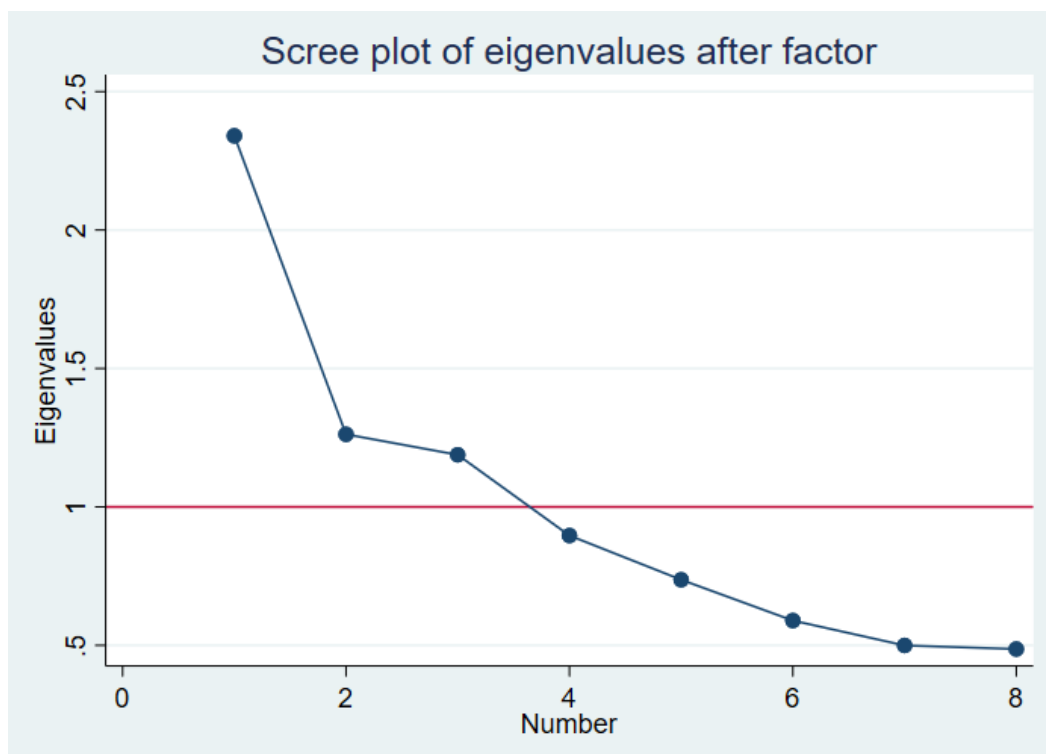
ind8 (Управление активами)

**Фактор 3:** переменные

ind6 (Исполнение судебных актов),

ind7 (Кадровый потенциал)

Графическое изображение факторных нагрузок:  
**screeplot, yline(1)**



4) Сохраните факторы в виде переменных.

**predict factor1 factor2 factor3**

```
. predict factor1 factor2 factor3  
(regression scoring assumed)
```

Scoring coefficients (method = regression; based on varimax rotated factors)

Variable	Factor1	Factor2	Factor3
ind1	0.40725	-0.02185	-0.07162
ind2	0.03530	0.41924	0.12288
ind3	0.38125	-0.12813	0.08298
ind4	0.21047	0.32588	0.01086
ind5	0.38228	-0.16898	-0.22607
ind6	-0.13524	-0.02950	0.60674
ind7	0.02733	-0.05372	0.54956
ind8	-0.15496	0.65085	-0.10955

В результате созданы 3 новые переменные:

#### Variables

Filter variables here	
Name	Label
name	Наименование администратора средств ...
grade	Общая оценка (в баллах)
ind1	Среднесрочное финансовое планирован...
ind2	Исполнение бюджета в части расходов
ind3	Исполнение бюджета по доходам
ind4	Учет и отчетность
ind5	Контроль и аудит
ind6	Исполнение судебных актов
ind7	Кадровый потенциал
ind8	Управление активами
factor1	Scores for factor 1
factor2	Scores for factor 2
factor3	Scores for factor 3