

Экономика-математическая школа
Введение в экономику и анализ данных

~~Домашнее задание~~ “Банковские системы”

Срок сдачи – до 23:59
3 февраля 2021

Экономика-математическая школа
Введение в экономику и анализ данных

Часть 1. Работа с выбросами

Winsorization

На занятии мы обсуждали способы борьбы с выбросами

Если не хочется удалять наблюдения, можно провести винсоризацию:

- Заменить экстремально большие значения на 95-ый процентиль
- А экстремально маленькие – на 5-ый процентиль

- Для индикатора `gfddai01` найдите 95-ый и 5-ый процентиля (используйте функцию `процентиль.вкл`)
- Замените все значения столбца `gfddai01`, превосходящие значение 95-ого процентиля, на 95-ый процентиль. Все значения меньше 5-ого процентиля – на 5-ый процентиль.
- Используйте преобразованный столбец, чтоб посчитать среднее `gfddai01` отдельно по годам (возьмите только 2005-2015 гг.) и регионам (сделайте это в Сводной таблице наподобие того, как на семинаре).
- Постройте сводную линейчатую диаграмму на основе полученной сводной таблицы. Видите ли вы какие-то закономерности на графике?

Экономика-математическая школа
Введение в экономику и анализ данных

Часть 2.

Доверительные интервалы

Финансовая стабильность до и после кризиса 2008

- Индикаторы стабильности - gfdd.si.01, gfdd.si.05
- Объясните, что будет означать увеличение значений этих индикаторов: рост стабильности или, наоборот, дестабилизацию банковской систем? (информацию о показателях смотрите на листе “Definitions and Sources”)
- Для 2007 и 2014 гг. постройте Сводную таблицу со средними, количествами и стандартными отклонениями индикатора gfdd.si.01 (вновь отдельно по годам и регионам). Но по столбцам поставьте 2007 и 2014 гг., по строкам – регионы.
- Добавьте колонку-разность средних за 2014 и 2007 гг. (для этого скопируйте значения из сводной таблицы и дальше работайте с обычным массивом)

Доверительный интервал

- Добавьте колонку со стандартным отклонением для разности

Формула стандартного отклонения разности =

$$\sqrt{\text{ст. отклон}_{2014}^2 + \text{ст. отклон}_{2007}^2}$$

- Используя функцию ДОВЕРИТ, посчитайте ширину доверительного интервала для разности средних (уровень значимости 5%, стандартное отклонение из предыдущего пункта, размер = количество наблюдений в 2014 г.)

Подсказка, как будут выглядеть итоговые таблички

- Сводная таблица

Названия столбцов		2007			2014		
Названия строк		Среднее gfddsi01	Количество gfddsi01	Стандартное отклонение gfddsi01	Среднее gfddsi01	Количество gfddsi01	Стандартное отклонение gfddsi01
East Asia & Pacific		13,64	25	7,33	13,42	23	5,85
Europe & Central Asia		10,06	50	5,85	10,69	50	6,86
Latin America & Caribbean		14,74	36	7,66	14,84	36	8,19
Middle East & North Africa		23,28	20	11,87	23,32	21	11,23
North America		21,96	3	4,37	20,28	3	8,46
South Asia		14,04	8	5,02	15,31	8	7,60
Sub-Saharan Africa		10,38	42	6,49	10,73	41	5,42

- Полное решение (делается не в сводной таблице)

Названия строк	2007			2014			Разность средних	Стандартное отклонение разности	Ширина доверительного интервала
	Среднее gfddsi	Количество gfd	Стандартное о	Среднее gfddsi	Количество gfd	Стандартное о			
East Asia & Pacific	13,63675444	25	7,33	13,42	23	5,85	-0,21	9,379	3,83
Europe & Central Asia	10,0553861	50	5,85	10,69	50	6,86	0,64	9,015	2,50
Latin America & Caribbean	14,73606944	36	7,66	14,84	36	8,19	0,10	11,214	3,66
Middle East & North Africa	23,27595	20	11,87	23,32	21	11,23	0,05	16,344	6,99
North America	21,9642	3	4,37	20,28	3	8,46	-1,69	9,518	10,77
South Asia	14,04176125	8	5,02	15,31	8	7,60	1,27	9,105	6,31
Sub-Saharan Africa	10,37945286	42	6,49	10,73	41	5,42	0,35	8,458	2,59

Интерпретация доверительного интервала

- Доверительный интервал (ДИ) для разности средних выглядит как (разность средних - ширина интервала; разность средних + ширина интервала).
- Интерпретация ДИ уровня значимости 5%: с вероятностью 95% мы можем гарантировать, что реальная разность средних (т.е. для генеральной совокупности – всех стран за все года) лежит в этом интервале.
- Если ДИ включает в себя ноль, то разность средних статистически не значима = уровень финансовой стабильности в 2014 г. почти не отличается от того, что было в 2007 г.

Графическое отображение ДИ

- Выделите столбцы “регион” и “разность средних”, вставьте гистограмму. По оси OY разность средних, по OX – регион.
- На вкладке “конструктор диаграмм” найдите “добавить элемент диаграммы” => “пределы погрешностей” => “дополнительные параметры”.
- На открывшейся вкладке в разделе “величина погрешности” выберите “пользовательская”. В качестве положительных и отрицательных значений ошибки укажите массив, где посчитана ширина интервалов.
- На основе того, входит или нет ноль в ДИ, сделайте выводы: Изменилась ли значимым образом финансовая стабильность после 2008 г? Если да, то финансовые системы стали более стабильными или менее? Одинаковые ли изменения в различных регионах?

[Помощь с построением графика](#), минуты 0:50 – 2:00

Подсказка, какой надо получить график

