***Regular clouds***

SPSS macros by Kirill Orlov

kior@comtv.ru, ttnphns@gmail.com

<https://www.spsstools.net/en/KO-spssmacros>

All rights reserved

*Правильные облака*. Создание многомерных данных с правильной, неслучайной структурой. В частности, такие данные можно понимать как полностью бескластерные, в отличие от данных, порожденных случайно. Полезны как модельные данные при изучении особенностей тех или иных статистических алгоритмов, например кластерного анализа.

# МАКРОС !KO\_UNIFORM: ОБЛАКА НА ОСНОВЕ РАВНОМЕРНОГО РАСПРЕДЕЛЕНИЯ

Version 1, Jan 2007. Tested on SPSS 11.5, 13, 15.

!KO\_uniform n= *400* /\*Число наблюдений данных облака (т е точек облака); можно указ любое, но данные нельзя получить при n<2^p

/p= *2* /\*Число переменных (измерений облака), >=1

/size= *20 20* /\*Размерность облака - p чисел, произведение их должно равняться n;

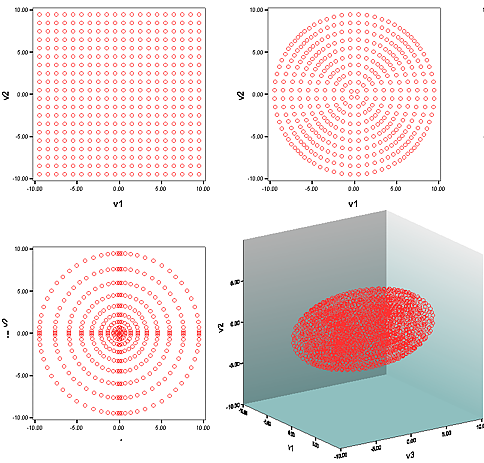
/\*либо ALL - создать все возможные облака разных размерностей

/rndness= *1* /\*Степень окатывания (округлости) - число от 0 (тж п/у, прямоугольник) до 1 (эллипс)

/shpness= /\*Степень заостренности (эксцесса) - число от 0 (тж п/у, равномерность) до 1 (пик в центре).

Минимум надо задать N, P, SIZE.

Макрос порождает данные (создавая новое окно данных), которые имеют правильную структуру на основе равномерного распределения. Речь прежде всего идет о прямоугольных облаках любой размерности (т.е. о *p*-мерных равномерных данных). В дополнение к этому, вы можете «окатать» такой прямоугольник, сгладив его углы вплоть до круга или эллипса. Хотя эта операция несколько нарушает равномерность распределения точек внутри облака, тем не менее окатыш можно практически принять за облако с приблизительно равномерным распределением точек. Также в дополнение, вы можете заострить эксцесс задающих облако переменных; это приводит к тому, что плотность между точками растет от периферии облака к его середине. Ниже, для примера – диаграммы рассеяния некоторых данных, порождаемых макросом:



Слева направо и сверху вниз:

* квадрат 20x20 точек
* он же, полностью окатанный в круг
* он же, полностью окатанный в круг, с добавлением эксцесса умеренной степени
* трехмерный эллипс (вытянутый куб, полностью окатанный)

ПРИМЕР 1.

!KO\_uniform n= 8000 /p= 3 /size= 20 20 20.

* Макрос создаст в новом окне данных 8000 наблюдений и три переменных V1, V2, V3. Геометрически это облако данных – куб 20x20x20 точек.

***Подкоманды***

**N, P**

Укажите, сколько создать наблюдений данных (точек облака), N, и сколько создать переменных (измерений облака), P. Минимум одна переменная. Число наблюдений можно указать любое, но данные невозможно создать, если N<2^P, и в этом случае они просто не появятся в окне данных.

**SIZE**

Подкоманда для задания размерности порождаемого прямоугольного облака (окатыш тоже порождается как прямоугольник). Если вы хотите создать одно облако, укажите P чисел, произведение которых равняется N. Например, 10 20 значит, что вы заказали прямоугольник 10x20 точек данных. С другой стороны, если вы хотите чтобы макрос создал все возможные прямоугольные облака разных размерностей, укажите ключевое слово ALL. Например, если N=100 и P=2, макрос создаст прямоугольники размерностью 2x50, 4x25, 5x20, 10x10 точек.

**RNDNESS**

Воспользуйтесь этой подкомандой, если хотите превратить прямоугольные облака в окатыши с той или иной степенью окатанности, сглаженности углов. Укажите число от 0 до 1. RNDNESS=1 превращает прямоугольник в круг или эллипс. RNDNESS=0, как и умолчание/незадание подкоманды оставляет прямоугольник нетронутым.

**SHPNESS**

Подкоманда позволяет сместить точки облака ближе к его центру, чтобы прибавить эксцесса переменным. Это ведет к неравномерности распределения точек в облаке. Укажите число от 0 до 1. Нуль это то же что умолчать/незадать подкоманду: смещения точек нет. Единица это максимальное усиление эксцесса.

***Особые режимы***

Поскольку макрос не имеет входящих данных, ему безразличны такие вещи как взвешенность, расщепленность или фильтрованность рабочего массива данных.