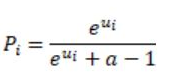
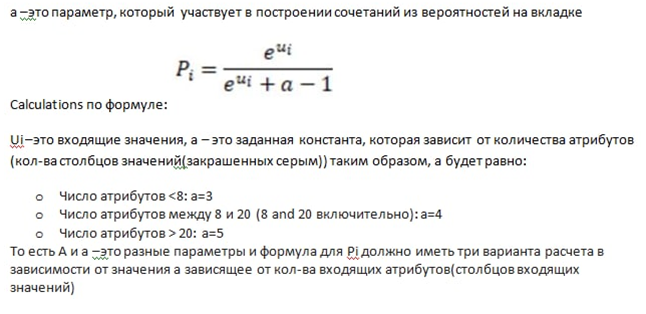
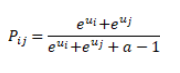
вопрос по возможности корректировки написанного макроса, который строит сочетания из вероятностей

построенных по формуле: , для входящих данных(6 столбцов заголовки, которых выделены цветом)

То есть Ui –это входящие значения, а – это заданная константа, в данном примере она равна 3-м:

Значение А –это подгруппы, количество столбцов стоящих после всех столбцов с атрибутами(в примере выделены жёлтым и состоят их нулей и единиц).

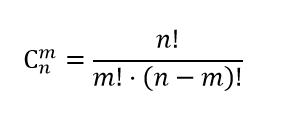
Если подставить первое значение для этой формулы, то получим:  


Для сочетаний из двух значений формула имеет вид:  


Если подставить значения в неё, то получим:  


Для трёх значений формула складывает соответственно 3 элемента: Pijk=(exp(Ui)+ exp(Uj)+ +exp(Uk))/ =(exp(Ui)+ exp(Uj)+ exp(Uk)+2) и так далее для любого количества элементов.

В файле с примером, я ввела эту формулу для сочетаний с одним входящим элементом и всех сочетаний с двумя значениями, также нужно сделать такие сочетания для трех, 4-х,5-ти и 6-ти элементов в приложенном примере. Такие сочетания можно посчитать по комбинаторной формуле для перестановок без повторений:



в коде: (ncombs = Factorial(ncols) / (Factorial(c) \* Factorial(ncols - c)).

Такой алгоритм зашит в онлайн калькуляторе <https://planetcalc.ru/3757/> : то есть для трёх входящих значений будет 20 сочетаний из столбцов:

BCD, BCE, BCF, BCG, BDE, BDF, BDG, BEF, BEG, BFG, CDE, CDF, CDG, CEF, CEG, CFG, DEF, DEG, DFG, EFG.

Для сочетания из всех шести элементов будет только одно: из столбцов BCDEFG.

Таким образом, макрос должен как алгоритм калькулятора рассчитывать комбинации и записывать значения из ячеек по формуле.

Соответственно если всего входящих столбцов будет 30 к примеру, то сочетания нужно последовательно сделать начиная от 1-го элемента заканчивая сочетаниями из 30 элементов. То есть макрос должен перебирать такие сочетания, учитывая формулу Pi.

Нужно учесть, что во входном файле может меняться количество входных данных – обычно таких столбцов как в примере от 6-ти до 30-ти, но может быть меньше или больше, таким образом, макрос должен это учитывать.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_-  
Обрабатываемый файл будет выглядеть как вкладка Input с неизвестным заранее числом входящих значений. Начинаться будет с первого столбца ID(точного названия для него нет) и далее уже сами входящие значения Ui в дальнейшем участвующие в построении комбинаций учитывающих формулу ниже(в примере таких входных столбцов 6), за ними всегда будут следовать столбцы означающие подгруппы в количестве равном вводимому значению а(в примере А=3), то есть при вводе параметра А, макрос будет понимать что с конца нужно отсчитать А столбцов.

Вкладка Results – это то что должно выводиться из расчёта(описание ниже)

Допустим, при запуске макроса будет выводиться диалоговое окно: «Input А». В файле с входными данными число А всегда соответствует последним столбцам(в примере они выделены желтым и содержат либо 0 либо 1). Эти столбцы являются подгруппами, по которым необходимо находить наилучшую комбинацию для каждого набора элементов и считать **Cumulative Reach и Reach** , отдельно для значения 1 и для 0, где **Cumulative Reach** –это максимальное значение вероятности(выбирает со вкладки Calculation), а формула для **Reach** навкладке Results.

Описание вкладки Results:

В первый столбец А выводится название(Title) название подгрупп c вкладки Input (последние столбцы, количество которых равны числу а) – в примере их три, а вверху всегда выводится название Total Sample –для подсчёта общего числа значений без учёта подгрупп( по формуле видно, что считается общее среднее по каждой комбинации- вторая строка на вкладке Calculation)

Затем, выводятся последовательно названия столбцов подгрупп-если есть возможность, то эти названия можно вводить с прибавкой к названию буквы s в конце.

Во второй столбец **Number of Items -** выводится последовательно кол-во наборов комбинаций начиная с первого и заканчивая последним.

В третий столбец **Best Combination –** выводитсяназвание наилучшей комбинации c первой строки (той у которой максимальное среднее значение вероятности, рассчитываемое на вкладке Calculation)

Четвертый и пятый столбцы – это сами значения макс средн значения и Reach – по формуле.

Для отдельных подгрупп **Cumulative Reach и Reach** рассчитывается в соответствии со значением 1. Эти средние для 1 берутся из расчёта со вкладки «Calculation» -первые верхние строки выделенные цветом –они рассчитываются как средние по столбцам построенным на основании формулы, последовательно для всех комбинаций:

2-я строка –общее среднее значение для каждого столбца

3-я строка – среднее для значения 1 по первой подгруппе с вкладки Input

4- я строка - среднее для значения 1 по второй подгруппе с вкладки Input

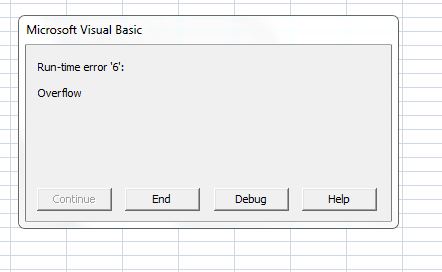
5- я строка - среднее для значения 1 по третей подгруппе с вкладки Input

6-я строка - среднее для значения 0 по первой подгруппе с вкладки Input

7-я и 8- я аналогично для 0 по второй и третей подгруппам.

В уже имеющемся макросе после его отработки создаётся вкладка Calculation, в которой подробно прописываются все сочетания, но поскольку число столбцов Эксель ограничено, то макрос срабатывает только до 15 атрибутов, далее уже не хватает стобцов, к примерудля 30 атрибутов необходимо 1,073,741,824 сочетаний.

При запуске файла с более чем 16-ть колонок(атрибутов) спустя некоторое время обработки(пару часов) выскакивает ошибка:



При дебаге подсвечиватся строка: avgif(i, nc) = avgif\_sum / counter1.