# Домашнее задание №2 по дисциплине "Методы прикладной статистики"

для студентов магистерской образовательной программы "Статистический анализ в экономике" факультета экономических наук

На предварительном этапе выбрать направление статистического исследования социально-экономического явления и информационные ресурсы данных на микроуровне. Например, «*Жилищные условия населения региона и его возможные детерминанты»* или «*Статистическое исследование доходов, расходов и сбережений населения»* на основе данных *Обследования бюджетов домашних хозяйств* (http://obdx.gks.ru/) и *Российского мониторинга экономического положения и здоровья населения НИУ ВШЭ (RLMS-HSE)* (https://www.hse.ru/rlms/). Рекомендуемое программное обеспечение: Microsoft Excel, IBM SPSS Statistics, R.

# Произвести отбор признаков для последующей классификации объектов с целью разделения общей совокупности на однородные подсовокупности. Произвести предварительный анализ этих признаков. Построить ортогональный базис из отобранных признаков методом главных компонент. Проанализировать возможность снижения размерности признакового пространства при сохранении в достаточной мере его информативности.

# Произвести параметрическую классификацию объектов в отсутствие обучения методом расщепления смесей вероятностных распределений по предполагаемому наиболее информативному признаку либо по первой главной компоненте.

# Разбить объекты на кластеры в пространстве главных компонент. Сравнить результаты кластеризации и классификации на основе декомпозиции смеси распределений.

# Создать обучающую выборку из наиболее типичных представителей выделенных однородных групп объектов. Произвести классификацию объектов методом дискриминантного анализа. Произвести проверку правильности классификации методом кросс-валидации.

# Построить в выделенных одним из способов однородных группах регрессионные модели по иным, чем используемые в классификации, признакам. Сравнить полученные модели и сделать выводы по проделанной работе.