

Экзамен по МСПС.

Правила проведения экзамена.

За общение с соседями, за использование мессенджеров, социальных сетей, электронной почты и прочих средств связи, а также мобильных устройств, студент немедленно удаляется с экзамена, а его работа аннулируется. Если будут обнаружены похожие работы (хотя бы в одной задаче), то обе работы будут аннулированы.

Длительность экзамена – 3 часа.

Условия задач.

- 1** (20 баллов) Выдана выборка X_1, \dots, X_n из неизвестного распределения. Предложить методику проверки, что распределение данной выборки принадлежит семейству распределений Парето $\mathbb{F} = \{F_{a,b}(x), a, b > 0\}$, где функция распределения $F_{a,b}(x) = (1 - (b/x)^a)I(x > b)$. Предложить также статистический критерий проверки того, что распределение выборки принадлежит семейству \mathbb{F} , и объяснить, почему он работает. Воспользоваться предложенными процедурами и определить, является ли выданная выборка распределенной по закону Парето.
- 2** (12 баллов) Выдано k выборок $\{X_i^{(1)}\}_{i=1}^n, \dots, \{X_i^{(k)}\}_{i=1}^n$. Проверить гипотезу о независимости выборок методами корреляционного анализа с помощью статистической процедуры, контролирующей FDR на уровне $\alpha = 0.05$. Являются ли выборки независимыми в совокупности? Если нет, укажите пары выборок, которые в результате проведенной статистической процедуры признаны зависимыми.
- 3** (18 баллов) Выданы данные $\{(y_i, x_{ij}), i = 1, \dots, n + q, j = 1, \dots, k\}$, причем y_{n+1}, \dots, y_{n+q} неизвестны. Используя пройденные методы регрессионного анализа, произвести отбор признаков по первым n объектам и предсказать значения откликов объектов с номерами $n + 1, \dots, n + q$. Описать и объяснить проделанные процедуры.