

Занятие № 4. Методы принятия оптимальных решений

Идентификация законов распределения

1. Подобрать закон распределения, наилучшим образом описывающий выборку K1000.dat.

1.1. Построить по выборке модели симметричных законов (нормальный, логистический, двусторонний экспоненциальный), оценить для построенных законов те же вероятности, зафиксировать значения в построенной таблице. Выполнить то же, используя режим «группировать при оценивании».

1.3. Используя 2 режима построить несимметричные модели (минимального значения, Su-Джонсона) и оценить соответствующие вероятности.

1.4. Опираясь на непараметрические критерии, постарайтесь подобрать модель, наилучшим образом описывающую данную выборку.

2. Для выборки model_1.dat подобрать наилучшим образом описывающую её параметрическую модель.

Указание 1. В качестве множества законов, из которого предпочтительнее выбирать модели рекомендуется рассматривать семейства распределений: гамма-распределение, Su- и Sl-Джонсона, бета-распределение II-го рода, бета-распределение III-го рода.

Указание 2. Так как объемы выборок достаточно велики, рекомендуется первоначально использовать режим «группирование при оценивании», а на стадии уточнения модели этот режим отключать.

Указание 3. В случае моделей с большим числом параметров, к сожалению, «удача» в определенной степени зависит от выбора параметра сдвига, с оцениванием которого имеются проблемы. В связи с этим может потребоваться ручная коррекция с последующим исключением из процедуры оценивания этого параметра (иногда и параметра масштаба).

3. Подобрать параметрическую модель, наилучшим образом описывающую выборку model_2.dat. Аналогично п.2.

4. Подобрать параметрическую модель, наилучшим образом описывающую выборку model_3.dat. Аналогично п.2.

5. Подобрать параметрическую модель, наилучшим образом описывающую выборку model_4.dat. Аналогично п.2.

6. Проверьте простые гипотезы о принадлежности выборок «истинным» законам, в соответствии с которыми генерировались соответствующие предложенные выборки.

7. «Считая», что наилучшая построенная Вами модель получена по некоторой другой выборке, проверьте согласие с Вашими моделями. Сравните достигаемые уровни значимости, с полученными в п. 6.

Информация о предложенных выборках

1. Выборка K1000.dat сгенерирована в соответствии с распределением Коши (2,0)

2. Выборка model_1.dat соответствует распределению Колмогорова. «Хорошие» модели:

Гамма (5.6039,0.1087,0.2600),

Бе-II (7.3016,23.3433,1.8971,0.2500)

Бе-III (5.5374,4.9926,2.7374,1.9617,0.2650)

3. Выборка model_2.dat соответствует предельному распределению статистики критерия Крамера-Мизеса-Смирнова $a1$: «Хорошая» модель Бе-III (3.0400,2.1876,20.0985,1.6304,0.0101)

4. Выборка model_3.dat соответствует предельному распределению статистики критерия Андерсона-Дарлинга $a2$: «Хорошая» модель Бе-III (3.7366,2.5535,16.6407,8.2133,0.0894)

5. Выборка model_4.dat соответствует $F_{5,20}$ -распределению Фишера: «Хорошая» модель Бе-II (2.4046,10.3956,4.3591,0.0000)

F-распределением Фишера описываются распределения многих статистик (при выполнении предположений о нормальности).

$F_{k,n}$ -распределение Фишера вычисляется через бета-распределение 2-го рода:

$$F_{k,n} = B_{II}\left(\frac{k}{2}, \frac{n}{2}, \frac{n}{k}, 0\right).$$