В первом файле присутствуют три распределения с соответствующими параметрами. Необходимо для каждого распределения:

1. сгенерировать выборку длины 1000 из данного распределения (см. стр. 19 методички)
2. построить по данной выборке эмпирическую функцию распределения;
3. построить гистограмму частот;
4. сравнить гистограмму частот и реальную плотность данного распределения (вычисление значения плотности в точке в пакете R описано на той же 19 стр.)
5. вычислить следующие выборочные характеристики: выборочное среднее, выборочную дисперсию, выборочную асимметрию, выборочный эксцесс;

(см. стр. 20-22 методички)

1. сравнить результаты пункта e с реальными характеристиками распределения

Второй файл содержит выборку из неизвестного распределения. Следует произвести чтение данных из данного файла (см. стр. 10-11 методички) Далее необходимо проделать пункты a-c, e уже для этих данных. На следующем занятии мы разберем как по форме гистограммы, выборочной асимметрии и выборочному эксцессу выбирать семейство распределений.

PS: файл с данными дан в CSV формате, поэтому вам придется использовать функцию readcsv. Приведу пример:

Пусть есть файл “temp.csv”, в котором данные о температуре чего-то за разные года. В нем температура и год разделены запятой. И есть два заголовка: “year” и “temp”. Прочтем этот файл:

MyData <- read.csv(file="temp.csv", header=TRUE, sep=",")

Получился объект типа dataframe. В нем будут два столбца, соответствующих году и температуре. Нужно вытащить из него только данные о температуре. Это можно сделать следующим образом:

temp<-MyData$temp

Теперь можно спокойно работать с переменной temp.