

ОП «Политология», 2019-20**Математика и статистика, часть 2****Описание выборок – часть 1. (10.04.2020)***А. А. Макаров, А. А. Тамбовцева, Н. А. Василёнок***Задача 1.**

(а) Дана выборка:

0 0 2 2 2 0 0 2

Может ли эта выборка быть правдоподобной (репрезентативной) выборкой из биномиального распределения с параметрами $n = 2$, $p = 0.5$? Обоснуйте свой ответ.

Решение. Биномиальная величина с параметрами $n = 2$, $p = 0.5$ описывает число успехов в серии из двух испытаний Бернулли. Она принимает значения 0, 1 и 2. Найдем вероятности этих значений:

$$P(S = 0) = C_2^0 \times 0.5^0 \times 0.5^2 = 0.25 = 1/4;$$

$$P(S = 1) = C_2^1 \times 0.5^1 \times 0.5^1 = 0.5 = 1/2;$$

$$P(S = 2) = C_2^2 \times 0.5^2 \times 0.5^0 = 0.25 = 1/4.$$

Теперь сравним полученные вероятности с долями 0, 1 и 2 в выборке. Значения 0 и 2 встречаются в выборке 4 раза из 8, значение 1 не встречается в выборке совсем. Получаем доли:

$$p(0) = 4/8 = 1/2 \quad p(1) = 0 \quad p(2) = 4/8 = 1/2.$$

Если бы выборка была репрезентативна, около половины значений составляли бы значения 1. А здесь их нет. Следовательно, выборка не является репрезентативной (правдоподобной) выборкой из указанного распределения.

(b) Дана выборка:

-20 25 5 78 27 16 -11 10 18

Может ли предложенная выборка быть правдоподобной (репрезентативной) выборкой из нормального распределения $N(2, \sigma^2 = 16)$? Обоснуйте свой ответ.

Решение. По правилу трех сигм, 99.8% значений нормальной случайной величины лежат в интервале $[a - 3\sigma; a + 3\sigma]$, в нашем случае $[-10; 14]$. В нашей выборке практически все значения лежат за пределами этого интервала, поэтому выборка не является правдоподобной (репрезентативной) выборкой из указанного распределения.