Лабораторная работа №1.
Генерация случайных величин с различными законами распределения (MathCad, STATISTICA, R)

Цель работы: Познакомиться с различными законами распределения. Научиться генерировать выборки, подчиняющиеся соответствующим законам распределения. Для сгенерированных или уже имевшихся выборок научиться с помощью стат. пакетов находить их выборочные характеристики, строить гистограммы, проводить поверку гипотез о виде распределения.

Теоретическая часть. Законы распределения, их параметры и основные числовые характеристики.

1. Равномерное распределение

Говорят, что случайная величина имеет *непрерывное равномерное распределение* на отрезке $[a,b]$, где $a,b \in R$, если её плотность $f_X(x)$ имеет вид:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, x \in [a,b]; \\ 0, x \notin [a,b]. \end{cases}$$

Обозначают: $X \sim U[a,b]$.

Функция распределения имеет вид:

$$F_X(x) = P\{X < x\} = \begin{cases} 0, x < a; \\ \dfrac{x-a}{b-a}, a \le x < b; \\ 1, x \ge b. \end{cases}$$

Основные числовые характеристики:

$$MX = \frac{a+b}{2},\ DX = \frac{(b-a)^2}{12}.$$

Стандартное равномерное распределение:

Если $a = 0$, $b = 1$, то говорят, что случайная величина $X \sim U[0,1]$ имеет стандартное равномерное распределение.

Если $Y = a + (b-a)X$, $a < b$, где $X \sim U[0,1]$, то $Y \sim U[a,b]$.

2. Нормальное распределение

Говорят, что случайная величина имеет *нормальное распределение* или *распределение Гаусса*, если для $x \in (-\infty,\infty)$ её плотность $f_X(x)$ имеет вид:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right).$$

Обозначают: $X \sim N[a,\sigma]$.

Функция распределения имеет вид:

$$F_X(x) = P\{X < x\} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right) dt .$$

Основные числовые характеристики:

$$MX = a,\ DX = \sigma^2,\ \sqrt{DX} = \sigma \text{ – СКО.}$$

Стандартное нормальное распределение:

Если $a = 1$, $\sigma = 1$, то говорят, что случайная величина $X \sim N[0,1]$ имеет стандартное нормальное распределение.

Если $Y = a + \sigma X$, где $X \sim N[0,1]$, то $Y \sim N[a,\sigma]$.

3. Экспоненциальное распределение.

Говорят, что случайная величина имеет *экспоненциальное* или *показательное распределение*, если её плотность $f_X(x)$ имеет вид:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, x \ge 0; \\ 0, x < 0. \end{cases}$$

где $\lambda > 0$ – параметр распределения.

Обозначают: $X \sim Exp(\lambda)$.

Функция распределения имеет вид:

$$F_X(x) = P\{X < x\} = \begin{cases} 1 - e^{-\lambda x}, x \ge 0; \\ 0, x < 0. \end{cases}$$

Основные числовые характеристики:

$$MX = \frac{1}{\lambda},\ DX = \frac{1}{\lambda^2} .$$

4. χ^2 – распределение

Пусть имеется совокупность k совместно независимых стандартных нормальных величин

$$U_i \sim N[0,1], i = \overline{1,k} .$$

Величина

$$X = U_1^{\ 2} + U_2^{\ 2} + \ldots + U_k^{\ 2}$$

имеет χ^2 *– распределение* с k степенями свободы и плотностью распределения

$$f_{\chi^2(k)} = \frac{(1/2)^{\frac{k}{2}}}{\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}},$$

где $\Gamma(\ldots)$ – гамма-функция,

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt .$$

Обозначают: $X \sim \chi^2(k)$.

Основные числовые характеристики:

$$MX = k,\ DX = 2k .$$

5. Распределение Пуассона

Говорят, что дискретная случайная величина X имеет *распределение Пуассона* с параметром $\lambda > 0$, если

$$p_\lambda(k) = P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0,1,2,\dots .$$

Основные числовые характеристики:

$$MX = \lambda,\ DX = \lambda .$$

6. Биномиальное распределение

Пусть $X_1, X_2, \dots, X_n$ – конечная последовательность случайных величин с *распределением Бернулли*, то есть

$$X_i = \begin{cases} 1, p; \\ 0, q = 1 - p; \end{cases}, i = \overline{1,n} .$$

Тогда дискретная случайная величина

$$X = X_1 + X_2 + \dots + X_n ,$$

определяющая число успехов m в последовательности $X_1, X_2, \dots, X_n$, имеет *биномиальное распределение* с n степенями свободы и вероятностью успеха p, при этом

$$p_X(p,n,m) = P\{X = m\} = C_n^m p^m q^{n-m}, m = \overline{0,n} ,$$

где

$$C_n^m = \frac{n!}{m!(n-m)!} .$$

Обозначают: $X \sim Bin(n, p)$.

Основные числовые характеристики:

$$MX = np,\ DX = npq .$$

7. Геометрическое распределение

Пусть $X_1, X_2, \dots, X_n$ – конечная последовательность случайных величин с *распределением Бернулли*, то есть

$$X_i = \begin{cases} 1, p; \\ 0, q = 1 - p; \end{cases}, i = \overline{1,n} .$$

Тогда дискретная случайная величина

$$X = \min\{i : X_i = 1\} - 1 ,$$

определяющая количество «неудач» до первого «успеха» имеет *геометрическое распределение*.

$$p(k) = P\{X = k\} = q^k p, k = 0,1,2,\dots .$$

Иногда в качестве геометрической случайной величины используют номер первой удачной попытки, то есть

$$X = \min\{i : X_i = 1\}$$

Тогда

$$p(k) = P\{X = k\} = q^{k-1} p, k = 1,2,3,\ldots .$$

Основные числовые характеристики:

$$MX = \frac{q}{p},\ DX = \frac{q}{p^2}.$$

Задание (Mathcad, Statistica, R):

Сгенерировать выборки объема N (параметр N задать произвольно в диапазоне от 200 до 500) следующих распределений:

- Равномерное,
- Нормальное,
- Экспоненциальное,
- χ^2,
- Пуассона,
- Биномиальное,
- Геометрическое.

Для каждого распределения параметры задать самостоятельно. Для полученных выборок найти выборочные средние и дисперсии, построить оценки параметров и проверить гипотезы о виде распределений.

Отчет

По каждому распределению записать:

- Закон распределения;
- Формулы математического ожидания и дисперсии;
- Заданные значения параметров распределения;
- Теоретические значения математического ожидания и дисперсии при выбранных параметрах;
- Выборочные средние и дисперсии, оценки параметров распределения;
- Результат проверки гипотезы о виде распределения.