Лабораторная работа 3, ТВМС

Бочарников Андрей, M3238
Ковешников Глеб, M3238
Шишкин Алексей, M3238

12 апреля 2020 г.

Формулировка

Для случайной величины, распределенной по нормальному закону с параметрами (a,σ^2), выполнить следующие действия:

1. Задать параметры распределения $X \sim N(a, \sigma^2)$.
2. Построить график $F_x(x)$.
3. Построить выборку генеральной совокупности X.
4. По построенной выборке построить график эмпирической функции распределения $F_n(x)$.
5. Построить доверительную полосу надежности.
6. На этом же графике построить $F_n(x)$ и $F_x(x)$.
7. На основе критерия Колмогорова провести проверку гипотез.

Аналогично для $X \sim U(a, b)$ - равномерно распределенной на $[a, b]$ случайной величины.

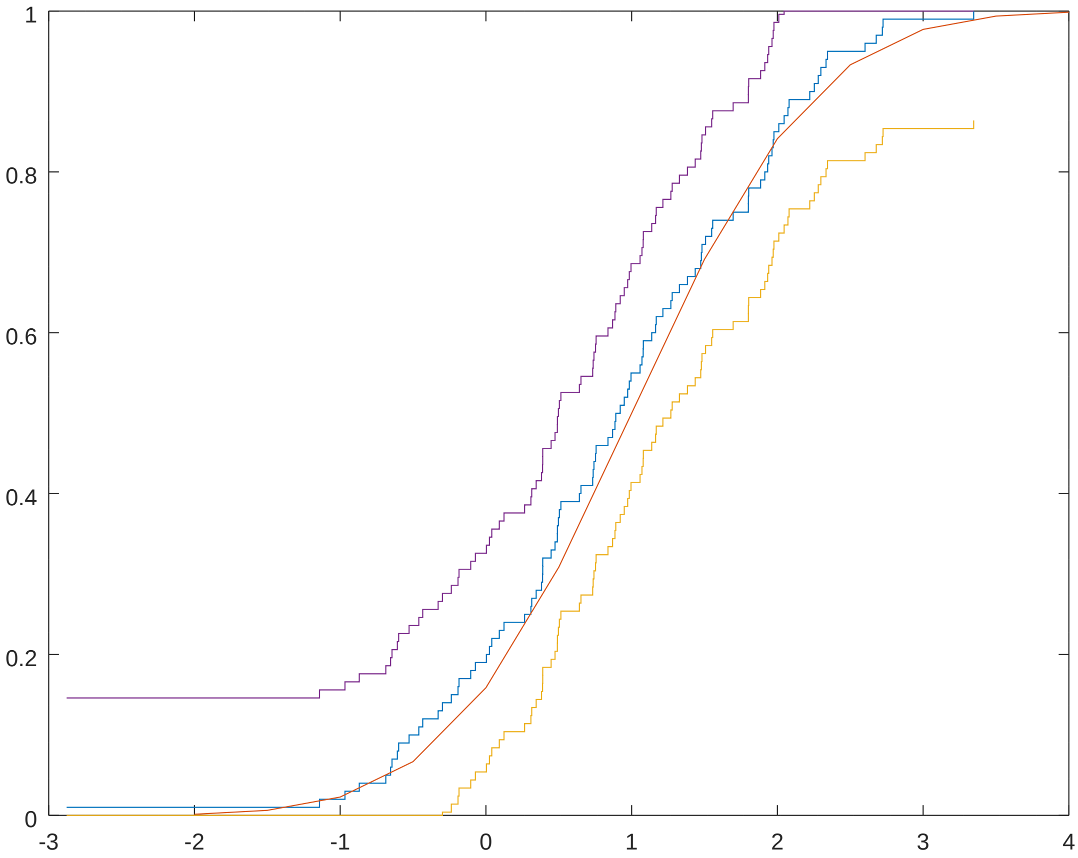
Входные данные

- Выборка генеральной совокупности: $n = 100$
- Доверительная полоса надежности: $\alpha = 0.05$, $u(1 - \alpha) = 1.36$
- Проверка критерия Колмогорова: $n = 10^4$ и 10^6

Программа 1

Нормальное распределение.

```
pkg load statistics

function p = test_Kolmogorov(n, mu, sigma)
    m = 100;
    X = sort(normrnd(mu, sigma, m, n));
    res = -1;
    for i=1:m
        X_i = X(i, :);
        F_X_i = normcdf(X_i, mu, sigma);
        current_val = max(abs(F_X_i - i / m), abs(F_X_i - (i - 1) / m));
        res = max(res, current_val);
    endfor
    gamma = 0.95;
    u_gamma = 1.36;
    p = mean((sqrt(m) * res) > u_gamma);
endfunction

function p = test_Smirnov(n, mu, sigma)
    m = 100;
    X = sort(normrnd(mu, sigma, m, n));
    omega = 1 / (12 * m);
    for i=1:m
        X_i = X(i, :);
        F_X_i = normcdf(X_i, mu, sigma);
        omega += (F_X_i - (2 * i - 1) / (2 * m)) .^ 2;
    endfor
    gamma = 0.99;
    w_gamma = 0.84;
    p = mean(omega > w_gamma ^ 2);
endfunction

n = 100;
mu = 1;
sigma = 1;

t = mu - 3 * sigma : 0.5 : mu + 3 * sigma;
F_x = normcdf(t, mu, sigma);

X = sort(normrnd(mu, sigma, n, 1));
F_n = 1 / n : 1 / n : 1;
[a, b] = stairs(X, F_n);

u = 1.36;
delta = u / sqrt(n);
plot(a, b, t, F_x, a, max(b - delta, 0), a, min(b + delta, 1))

test_Kolmogorov(10000, mu, sigma)
test_Kolmogorov(1000000, mu, sigma)

test_Smirnov(10000, mu, sigma)
test_Smirnov(1000000, mu, sigma)
```

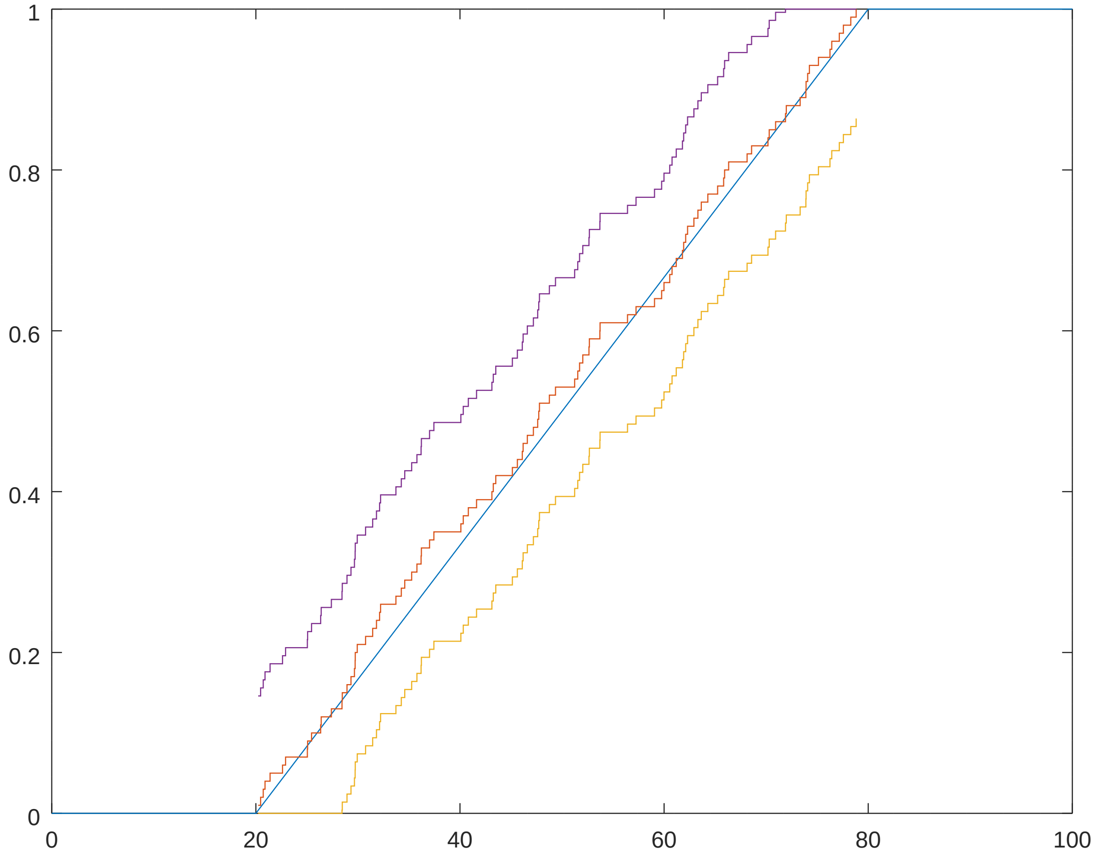


1
0.8
0.6
0.4
0.2
0
-3
-2
-1
0
1
2
3
4

Программа 2

Равномерное распределение.

```
pkg load statistics

function p = test_Kolmogorov(n, a, b)
  m = 100;
  X = unifrnd(a, b, m, n);
  X = sort(X);
  res = -1;
  for i=1:m
    X_i = X(i, :);
    F_X_i = unifcdf(X_i, a, b);
    current_val = max(abs(F_X_i - i / m), abs(F_X_i - (i - 1) / m));
    res = max(res, current_val);
  endfor
  gamma = 0.95;
  u_gamma = 1.36;
  p = mean((sqrt(m) * res) > u_gamma);
endfunction

function p = test_Smirnov(n, mu, sigma)
    m = 100;
    X = sort(unifrnd(mu, sigma, m, n));
    omega = 1 / (12 * m);
    for i=1:m
        X_i = X(i, :);
        F_X_i = unifcdf(X_i, mu, sigma);
        omega += (F_X_i - (2 * i - 1) / (2 * m)) .^ 2;
    endfor
    gamma = 0.99;
    w_gamma = 0.84;
    p = mean(omega > w_gamma ^ 2);
endfunction

n = 100;
a = 20;
b = 80;

df_x = 0:0.01:n;
df_y = unifcdf(df_x, a, b);

edf_x = sort(unifrnd(a, b, 1, n));
edf_y = 1/n:1/n:1;
[st_a, st_b] = stairs(edf_x, edf_y);

gamma = 0.95;
u = 1.36;
delta = u / sqrt(n);
edf_y_minus = max(0, st_b - delta);
edf_y_plus = min(1, st_b + delta);
plot(df_x, df_y, st_a, st_b, st_a, edf_y_minus, st_a, edf_y_plus);

test_Kolmogorov(10000, a, b)
test_Kolmogorov(1000000, a, b)

test_Smirnov(10000, a, b)
test_Smirnov(1000000, a, b)
```


1
0.8
0.6
0.4
0.2
0
0
20
40
60
80
100

Проверка гипотез

И для нормального, и для равномерного распределения результаты получаются одинаковые:
На основе критерия Колмогорова с параметром $\alpha = 0.05$, вероятность ошибки первого рода асимптотически получается какой и должна быть – 0.0448 и 0.45 при $n = 10^4$ и $n = 10^6$ соответственно.
На основе критерия Смирнова с параметром $\alpha = 0.01$, вероятность ошибки первого рода при $n = 10^4$ и $n = 10^6$ получается 0.0148 и 0.12374 соответственно, то есть сходится к 0.01, как и должно быть.

Вывод

В обоих распределениях, видно по графикам, что функция распределения лежит в доверительной полосе. Гипотезы также сходятся к нужным величинам.