Практическая работа 2.. «Эмпирическая функция распределения. Поведение в точке»

Цель работы:

1. ознакомиться с определением ЭФР и ее поведением при фиксированном значении аргумента;
2. аналитически и графически оценить надежность асимптотического интервала;
3. убедиться в том, что асимптотические методы работают при конечном объеме выборки.

Задание и ход работы.

1. Выбрать параметры двух из трех распределений генеральной совокупности X: $X \sim U(a,b), X \sim Exp^{u}$ или $X \sim N(a,\sigma^2)$.
2. Выбрать такую точку t_0, что $0.05 < F_X(t_0) < 0.95$. Вычислить $F_X(t_0)$.
3. Смоделировать m=10^2 выборок объема n=10^4 для каждого из двух выбранных распределений. Для каждой выборки построить $F_n(t_0)$ – значение эмпирической функции распределения в точке t_0 -- оценку значения функции распределения в точке t_0, то есть величины $F_X(t_0)$. Для каждого из распределений получите 100 оценок величины $F_X(t_0)$.
4. Значение функции распределения $F_X(t_0) = P(X \in (-\infty, t_0) = \Delta)$ является вероятностью события $A = \{X \in (-\infty, t_0)\}$. Значение эмпирической функции распределения $F_n(t_0)$ –оценка вероятности события $A = \{X \in (-\infty, t_0)\}$, то есть k($\Delta$)/n - частота попадания значения случайной величины X в интервал Δ. Частота, полученная по серии независимых однотипных испытаний с двумя исходами – A и $\bar{A}$, является состоятельной, несмещенной, асимптотически нормальной оценкой вероятности события. Свойство асимптотической нормальности позволяет строить асимптотический доверительный интервал надежности γ. Фиксировать $\gamma > 0.9$ и построить по 100 асимптотических доверительных интервалов надежности γ для значения $F_X(t_0)$ каждого из выбранных распределений.
5. Построить 2 графика – по оси x - номер выборки, по оси y – соответствующие левый и правый концы асимптотических доверительных интервалов и значение $F_X(t_0)$.
6. Найти количество δ_n асимптотических доверительных интервалов, в которые значение $F_X(t_0)$ не попало. Сравнить среднее количество δ_n для к =100 серий

(mean(δ_n)) с величиной 1- γ (δ_n можно рассматривать как оценку величины 1- γ) для различных $\gamma = 0.9, 0.91, \dots, 0.99$. Составить таблицу результатов.

Выбранные параметры:

m = 10^2;

n = 10^4;

t0 = 0.4;

alpha = 0;

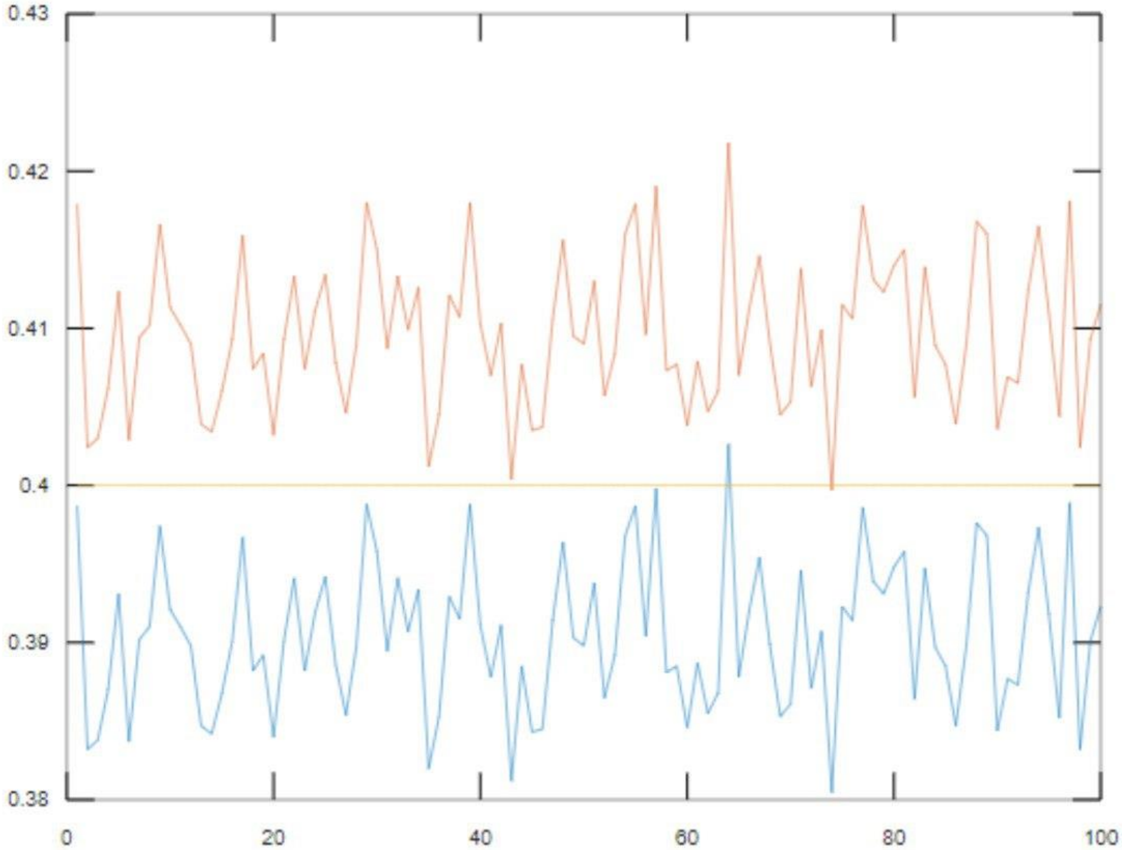
sigma = 1;

a = 0;

b = 1;

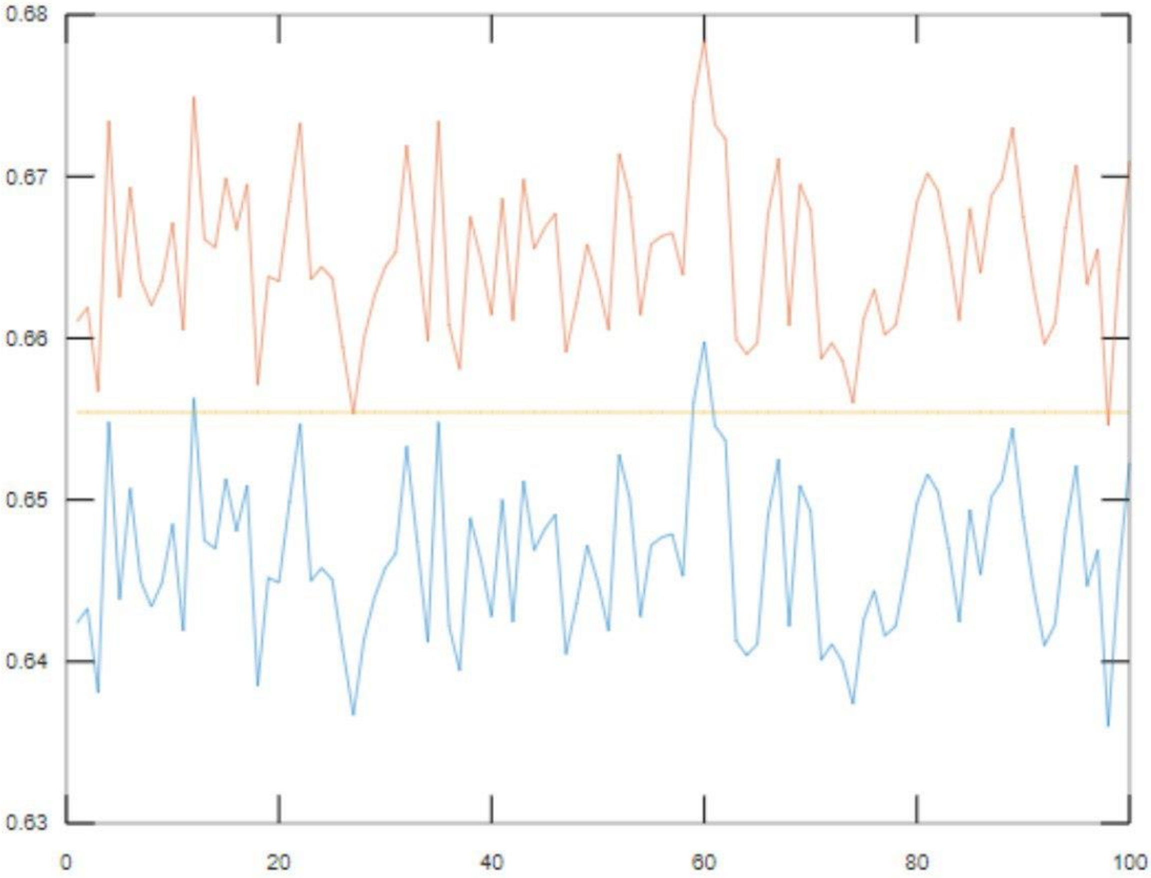
Графики для gamma = 0.95:

Равномерное распределение:




```
gamma = 0.9   and   average = 10.41
gamma = 0.91   and   average = 9.11
gamma = 0.92   and   average = 7.98
gamma = 0.93   and   average = 7.22
gamma = 0.94   and   average = 6.26
gamma = 0.95   and   average = 4.68
gamma = 0.96   and   average = 4.12
gamma = 0.97   and   average = 3.17
gamma = 0.98   and   average = 2.03
gamma = 0.99   and   average = 1.11
```

Нормальное распределение:




```
gamma = 0.9   and   average = 9.62
gamma = 0.91   and   average = 9.06
gamma = 0.92   and   average = 8.34
gamma = 0.93   and   average = 7
gamma = 0.94   and   average = 6.09
gamma = 0.95   and   average = 4.98
gamma = 0.96   and   average = 4.19
gamma = 0.97   and   average = 2.93
gamma = 0.98   and   average = 1.88
gamma = 0.99   and   average = 1.03
```

При увеличении gamma среднее количество δn асимптотических доверительных интервалов, в которые значение Fx(t0) не вошло, уменьшается. Заметим, что значение average примерно в k раз больше (k – количество серий), чем (1 – gamma). Получается, что при приближении gamma к единице увеличивается точность доверительного интервала.