Практическая работа 1.

«Ознакомление с пакетом Octave. Применение метода Монте-Карло для нахождения оценок объемов и интегралов.»

Работа выполняется индивидуально согласно вариантам, распределенным в начале семестра.

Цель работы:

1. ознакомиться с пакетом Octave (Matlab). Часто применяемыми для задач математической статистики и различных ее приложений (экономика, эконометрия, финансы…);
2. научиться использовать вероятностные методы в некоторых вычислительных задачах;
3. научиться применять метод Монте-Карло для оценок вероятностей и моментов в различных задачах.

Практическое задание

Методом Монте-Карло оценить объем части тела $\{F(\overline{x}) \le c\}$, заключённой в k-мерном кубе с ребром [0, 1]. Функция имеет вид $F(\bar{x}) = f(x_1) + f(x_2) + \dots + f(x_k)$. Для выбранной надежности $\gamma \ge 0.95$ указать асимптотическую точность оценивания и построить асимптотический доверительный интервал для истинного значения объёма. Используя объем выборки n=10^4 и n=10^6 оценить скорость сходимости и показать, что доверительные интервалы пересекаются.

Аналогично построить оценку интегралов (представить интеграл как математическое ожидание функции, зависящей от случайной величины с известной плотностью) и для выбранной надежности $\gamma \ge 0.95$ указать асимптотическую точность оценивания и построить асимптотический доверительный интервал для истинного значения интеграла.

Вариант 1

1. $f(x) = x^3$; 2. k = 6; 3. c = 1,4.

Аналогично построить оценку интегралов (представить интеграл как математическое ожидание функции, зависящей от случайной величины с известной плотностью) и для выбранной надежности $\gamma \ge 0.95$ указать асимптотическую точность оценивания и построить асимптотический доверительный интервал для истинного значения интеграла.

a) $\int_2^5 \ln(1 + x^2)\, dx,$ b) $\int_{-\infty}^{\infty} \cos(x) \exp\left(\frac{-(x+3)^2}{4}\right) dx.$

Ход работы:

1. Ознакомится с пакетом Octave используя предоставленные материалы и рекомендованную литературу;

2. изучить методы решения предложенной задачи, используя методические указания к работе и материал лекционных и практических занятий;
3. применить полученные знания к решению предложенных задач;
4. сформировать отчет по практической работе.

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(Δ)/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.

Практическая работа 3. «Эмпирическая функция распределения. Поведение в «целом»

Цель работы:

1. ознакомится с методами и результатами оценивания функции при помощи расстояний Колмогорова и Смирнова;
2. ознакомится теоретически и практически с построением доверительной полосы;
3. научить использовать критерии согласия и исследовать их свойства при конечном n.

Задание и ход работы

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 $F_X(x)$, используя функцию normcdf.
3. При n=100 построить выборку из генеральной совокупности *X*.
4. По построенной выборке построить график эмпирической функции распределения $F_n(x)$, используя при построении встроенную функцию [a,b]=stairs(x,y) для построения кусочно-постоянной функции. Учесть при построении, что $F_n(x)$ изменяется на 1/n в каждой следующей точке выборки.
5. Построить доверительную полосу надежности γ =0.95; u(γ)=1.36.
6. На этом же графике построить $F_n(x)$ и $F_X(x)$. Убедится, что функция распределения попадает (?) в доверительную полосу.
7. На основе критерия Колмогорова и на основе критерия Смирнова провести проверку гипотез согласия с фиксированной функцией распределения при n=10^4 и n=10^6.
8. Оценить ошибки I и II рода каждого из критериев.

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

Практическая работа 4.. « Гистограмма как оценка плотности»

Цель работы:

1. ознакомиться с определением гистограммы и ее поведением при фиксированном значении аргумента;
2. научиться находить значения гистограммы, строить ее график одновременно (в качестве тестового задания) с реальной плотностью генеральной совокупности;
3. убедиться в том, что асимптотические методы работают при конечном объеме выборки при корректном (с дополнительными требованиями) их использовании.

Задание и ход работы

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 $f_X(x)$, используя функцию normpdf.
3. При n=10^6 построить выборку из генеральной совокупности X.
4. По построенной выборке вычислить значения и построить график гистограммы, используя при построении встроенную функцию [a,b]=stairs(x,y) для построения кусочно-постоянной функции.
5. Совместить графики плотности и гистограммы на одном рисунке
6. На основе хи-квадрат критерия Пирсона провести проверку гипотез согласия с семейством распределения генеральной совокупности
7. Оценить ошибки I и II рода критерия.

Сравнить с аналогичной обработкой выборки из равномерного распределения.

Практическая работа 5. «Линейные статистические модели или модели регрессии»

Цель работы:

1. ознакомление с линейными статистическими моделями;
2. ознакомится с встроенным в пакет при помощи функций polyfit, polyval матричным методом;
3. убедиться в том, что матричный метод в координатной форме приводит к задачам регрессии.

Задание и ход работы

Построить по соответствующим варианту данным квадратичный P2 и линейный P1 многочлены на промежутке delta= [xmin, xmax]. Добавить к значениям многочлена n независимых значений случайной величины Z~N(0,s^2). Выбрать на промежутке delta n точек: b1=xmin, b2=xmin+h, b3=xmin+2h,..., bn=xmax; h=(xmax-xmin)/(n-1). Найти в этих точках значения зашумленных многочленов (Y). По этим исходным данным оценить коэффициенты исходных многочленов P2 и P1 и значения X=Y-Z (с получением оценки значений - Yn) матричным методом и через функции Matlab или Octave: polyfit, polyval для квадратичного многочлена. Для линейного многочлена использовать также уравнение выборочной линейной регрессии. Проверить ортогональность Yn-Y и Yn (проецирующего вектора и проекции). Найти оценку уровня шума s. В качестве результата вывести исходные данные и все возможные их оценки. Привести графики исходных многочленов и полученных оценок (значения многочленов в выбранных точках, полученные различными методами должны совпасть).

Конкретный вариант

Смоделировать выборку значений линейной или квадратичной функции в нормальном шуме:

$$y = 3{,}1x + 2{,}4, x \in (-1{,}5), y = 2{,}2x^2 + 1{,}8x + \sigma = 1{,}4, m = 80,$$

где σ – уровень шума, систематическая ошибка отсутствует; m – количество точек измерения функции, зашумленной нормальным шумом с уровнем σ (точки измерения дискретно-равномерно распределены на (-1, 5)). Используя модели линейной и квадратичной простой регрессии оценить коэффициенты линейной или квадратичной зависимости, построить оценку уровня шума, проверить ортогональность «остатка» и базисов соответствующих линейных пространств.

Ход работы:

1. изучить линейную статистическую модель и модель простой регрессии, используя лекционный материал и рекомендованную литературу;
2. изучить методы решения предложенной задачи, используя методические указания к работе и материал лекционных и практических занятий;
3. применить полученные знания к решению предложенных задач;
4. сформировать отчет по практической работе.