Федеральное агентство по образованию
Воронежский государственный университет

Компьютерный практикум по теории вероятностей в среде MATHCAD

Специальность 010501

ОПД.Ф.04

ВОРОНЕЖ
2008

Утверждено научно-методическим советом факультета прикладной математики, информатики и механики: протокол № 9 от 11 июня 2008 г.

Пособие подготовлено на кафедре технической кибернетики и автоматического регулирования факультета прикладной математики, информатики и механики Воронежского государственного университета.
Рекомендуется для студентов 3 курса дневного отделения.
Для специальности 010501 ОПД. Ф.04

Введение

Настоящий практикум предлагает цикл работ для выполнения на персональном компьютере. При решении задач по теории вероятностей используется пакет Mathcad, выбор данного пакета не случаен. Он представляет собой интегрированную многофункциональную систему, предназначенную для проведения разнообразных вычислений. Этот пакет получил широкое распространение благодаря чрезвычайной простоте интерфейса. Mathcad содержит:

- обширную библиотеку встроенных математических функций;
- инструменты построения графиков различных типов;
- удобно организованную интерактивную систему получения справки;
- используется традиционный для математики способ записи функций и выражений
- Mathcad представляет пользователю оперативные возможности электронной таблицы текстового процессора.

1. Функции и инструменты MATHCAD

Прежде чем приступать к решению задач теории вероятностей в Mathcad, познакомимся с инструментами, которые предоставляет пакет для их решения.

Напомним, что дискретная случайная величина ξ, принимающая значения $x_1 < x_2 < \ldots < x_i < \ldots$ с вероятностями $p_1, p_2, \ldots, p_i, \ldots,$ может быть задана *распределением* – таблицей вида

ξ	x_1	x_2	…	x_i	…	x_n
p	p_1	p_2	…	p_i	…	p_n

Такие таблицы в среде Mathcad удобно хранить в виде матрицы размерности 2×n. Функция распределения случайной величины, имеющей приведённое выше распределение, имеет вид:

$$F(x) = \begin{cases} 0 & , x < x_1 \\ p_1 & , x_1 \le x < x_2 \\ p_1 + p_2 & , x_2 \le x < x_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots & \\ p_1 + p_2 + \ldots + p_{n-1} & , x_{n-1} \le x < x_n \\ 1 & , x_n \le x \end{cases}$$

Ниже приведён фрагмент рабочего документа с определением распределения случайной величины, её функции распределения и графиком функции распределения для случайной величины, имеющей следующее распределение:

ξ	1	0	7	4	-2
P	0.1	0.5	0.1	0.1	0.2

Распределение случайной величины

$$A := \begin{pmatrix} -2 & 0 & 1 & 4 & 7 \\ 0.2 & 0.5 & 0.1 & 0.1 & 0.1 \end{pmatrix}$$

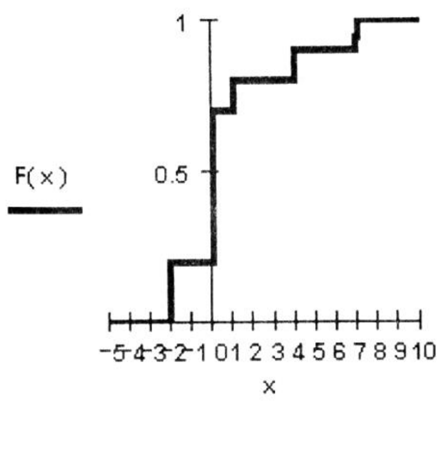
Функция распределения случайной величины

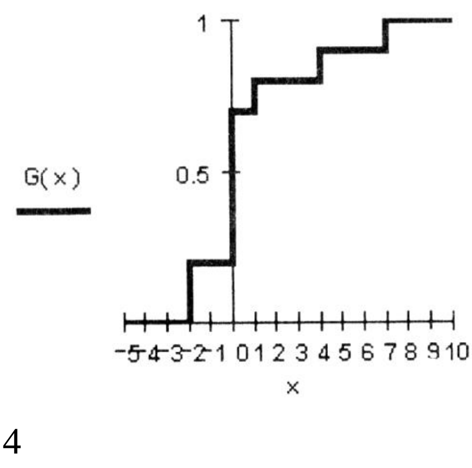
$$F(x) := \begin{cases} 0 & ,-\infty < x < A_{1,1} \\ A_{2,1} & ,A_{1,1} \le x < A_{1,2} \\ (A_{2,1} + A_{2,2}) & ,A_{1,2} \le x < A_{1,3} \\ (A_{2,1} + A_{2,2} + A_{2,3}) & ,A_{1,4} \le x < A_{1,5} \\ (A_{2,1} + A_{2,2} + A_{2,3} + A_{2,4}) & ,A_{1,4} \le x < A_{1,5} \\ 1 & ,A_{1,5} \le x < \infty \end{cases}$$

Функция распределения случайной величины, определённая другим способом:

$$G(x) := \begin{cases} 0 & ,-\infty < x < -2 \\ 0.2 & ,-2 \le x < 0 \\ 0.2 + 0.5 & ,0 \le x < 1 \\ 0.2 + 0.5 + 0.1 & ,1 \le x < 4 \\ 0.2 + 0.5 + 0.1 + 0.1 & ,4 \le x \le 7 \\ 1 & ,7 \le x < \infty \end{cases}$$

График функции распределения случайной величины





Указание. Распределение случайной величины сохранено в матрице А: $A_{1,i}$ – значения случайной величины; $A_{2,i}$ – соответствующие вероятности; i = 1,2,3,4,5. Функцию распределения, заданную разными выражениями на разных интервалах изменения аргументов, можно определить следующим образом: разверните панель программных элементов щелчком по кнопке и панель знаков отношений – щелчком по кнопке и не закрывайте их, пока не закончите определение функции. Введите имя функции переменной ***x*** и знак присваивания, щелкните в панели программных элементов по кнопке Add Line, введите в помеченной позиции нуль, щелкните по кнопке IF и введите неравенства, определяющие первый интервал изменения аргумента (символ ∞ можно ввести щелчком по соответствующей кнопке в панели); затем перейдите во вторую строку определения функции, введите $A_{2,1}$ – имя переменной, содержащей значение p_1, или число ***0.2*** – значение p_1 , выделите, нажимая клавишу <***SPACE***>, выражение для функции, щелкните по кнопке IF, введите неравенства, определяющие второй интервал изменения аргумента (знак можно ввести щелчком по соответствующей кнопке в панели отношений); выделите, нажимая клавишу <***SPACE***>, вторую строку определения функции, щелкните по кнопке Add Line и введите, действуя, как описано выше, определение функции на следующем интервале. В рабочем документе приведены два способа определения функции – с использованием имен переменных и с использованием конкретных значений этих переменных. Графики функций распределений построены стандартным для декартовых графиков способом. Следует помнить, что MathCad не совсем корректно строит графики ступенчатых функций, соединяя отрезками прямых значения функции в точке скачка. Более точный график функции распределения представляет собой отрезки, параллельные оси абсцисс, с “выколотым” правым концом.

Распределение дискретного случайного вектора также удобно хранить в матрице размерности (m+1)x(n+1). Первоначальному элементу первой строки этой матрицы присваивается нулевое значение, остальные элементы первой строки содержат значения случайной компоненты η, элементы первого столбца – значения случайной компоненты ξ, а остальные элементы – соответствующие вероятности: элемент, расположенный в (j+1)-м столбце (i+1)-й строки содержит значение вероятности p_{ij} того, что случайный вектор (ξ,η) принимает значение (x_i, y_i).

Для проведения вычислений со случайными величинами (непрерывными и дискретными) в MathCad есть богатая библиотека встроенных функций наиболее распространенных стандартных распределений. Каждое распределение представлено в библиотеке тремя функциями — плотностью вероятностей, функцией распределения и функцией, обратной к функции распределения.

Например, для работы с нормальным распределением предназначены функции dnorm(x,η,σ), pnorm(x,η,σ) и qnorm(x,η,σ). Значением функции dnorm(x,η,σ) является значение в точке x плотности вероятностей случайной величины ξ, имеющей нормальное распределение с математическим ожиданием $M\xi = \eta$ и дисперсией $D\xi = \sigma^2$; значение функции pnorm(x,η,σ) – значение функции распределения этой же случайной величины ξ; значением функции qnorm(x,η,σ) служит решение уравнения F(x) = p, где F(x) – функция распределения, определенная функцией pnorm(x,η,σ), т.е. значением qnorm(x,η,σ) является квантиль уровня p нормально распределенной случайной величины. Имена всех встроенных функций, определяющих плотности вероятностей, начинаются с буквы d, определяющих функции распределения – с буквы p, определяющих квантили – с буквы q.

Ниже приведены список всех распределений, представленных в библиотеке MathCad, и имена соответствующих функций:

бета-распределение — dbeta(x,s_1,s_2), pbeta(x,s_1,s_2), qbeta(p,s_1,s_2);
биномиальное распределение — dbinom(k,n,p), pbinom(k,n,p), qbinom(p,n,r);
распределение Коши — dcauchy(x,l,s), pcauchy(x,l,s), dcauchy(p,l,s);
χ^2 - распределение — dchisq(x,d), pchisq(x,d), qchisq(p,d);
экспоненциальное распределение — dexp(x,r), pexp(x,r), qexp(p,r);
распределение Фишера (F-распределение) — dF(x,d_1,d_2), pF(x,d_1,d_2), qF(x,d_1,d_2);
гамма-распределение — dgamma(x,s), pgamma(x,s), qgamma(p,s);
геометрическое распределение — dgeom(x,p), pgeom(x,p), qgeom(p,r);
логнормальное распределение — dlnorm(x,η,σ), plnorm(x,η,σ), qlnorm(p,η,σ);
логистическое распределение — dlogis(x,l,s), plogis(x,l,s), qlogis(p,l,s);
отрицательное биномиальное распределение — dnbinom(k,n,p), pnbinom(k,n,p), qnbinom(p,n,r);
нормальное распределение — dnorm(x,η,σ), рпогт(x,η,σ), qnorm(p,η,σ);
распределение Пуассона — dpois(x,λ), ppois(x, λ), qpois(p, λ);
распределение Стьюдента — dt(x,d), pt(x,d), qt(p,d);
равномерное распределение — dunif(x,a,b), punif(x,a,b), qunif(p,a,b);
распределение Вейбулла — dweibull(x,s), pweibull(x,s), qweibull(p,s).

Кроме того, в библиотеке встроенных функций Mathcad, естественно, есть функция Лапласа $erf\, x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$.

Для вычисления числовых характеристик дискретных и непрерывных случайных величин применяются операторы интегрирования и дифференцирования, вычисления конечных сумм и суммирования рядов, которые вызываются щелчком мыши по кнопке в панели и заполнением соответствующих помеченных полей. Примеры использования этих операций при решении задач теории вероятностей приведены в последующих разделах.

2. Случайные величины. Функции распределения

Теория вероятностей изучает математические модели случайных явлений окружающего нас мира. Одно из центральных понятий теории вероятностей – понятие случайной величины. *Случайной величиной* называется числовая функция, заданная на множестве случайных событий.

Например, случайной величиной является число очков, выпавших при бросании игральной кости, или рост случайно выбранного из учебной группы студента. В первом случае мы имеем дело с *дискретной* случайной величиной (она принимает значения из дискретного числового множества) M = {1, 2, 3, 4, 5, 6}; во втором случае – с *непрерывной* случайной величиной (она принимает значения из непрерывного числового множества – из промежутка числовой прямой I = [100, 230]). В дальнейшем случайные величины будем обозначать греческими буквами.

Функция распределения случайной величины

Каждая случайная величина полностью определяется своей функцией распределения.

Если ξ - случайная величина, то функция $F(x) = F_{\xi}(x) = P(\xi < x)$ называется *функцией распределения* случайной величины ξ. Здесь $P(\xi < x)$ – вероятность того, что случайная величина ξ принимает значение, меньшее x.

Функция распределения любой случайной величины обладает следующими свойствами:

- $F(x)$ определена на всей числовой прямой **R;**
- $F(x)$ не убывает, т.е. если $x_1 \le x_2$, то $F(x_1) \le F(x_2)$;
- $F(-\infty) = 0$ и $F(+\infty) = 1$, т.е. $\lim\limits_{x \to -\infty} F(x) = 0$ и $\lim\limits_{x \to +\infty} F(x) = 1$
- $F(x)$ непрерывна слева, т.е. $\lim\limits_{x \to x_0 - 0} F(x) = F(x_0)$

Важно понимать, что функция распределения является «паспортом» случайной величины: она содержит всю информацию об этой случайной величине, и поэтому изучение случайной величины заключается в исследовании ее *функции распределения*, которую часто называют просто *распределением*. Таким образом, когда говорят о *нормальном распределении*, то подразумевают случайную величину, имеющую *нормальную функцию распределения*.

Указание. Следует отметить, что Mathcad, изображая ступенчатые функции, соединяет отрезком прямой значения функций в точках разрыва. Разрывные функции принято изображать иначе – помечая стрелкой направление разрыва (стрелка вправо – функция непрерывна в точке справа, стрелка влево – для точек, где функция непрерывна слева). На рис.1 приведен график функции распределения в общепринятом виде.

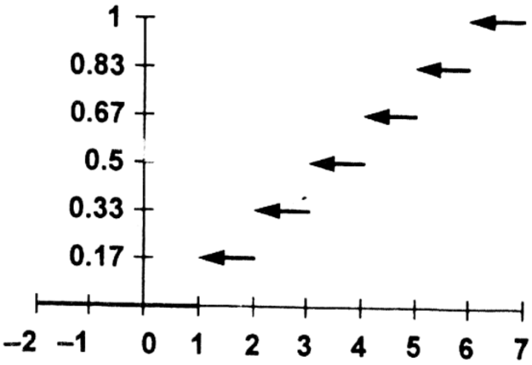


Рис 1.

Если функция распределения $F_\xi(x)$ непрерывна, то случайная величина ξ называется *непрерывной случайной величиной*. Если функция распределения $F_\xi(x)$ непрерывно дифференцируема, то более наглядное представление о случайной величине дает *плотность вероятности случайной величины* $p_\xi(x)$, которая связана с функцией распределения $F_\xi(x)$ формулами:

$$F_\xi := \int_{-\infty}^{x} p_\xi(t)\,dt$$

$$p_\xi(x) := \frac{d}{dx} F_\xi(x)$$

Отсюда, в частности, следует, что для любой случайной величины интеграл

$$\int_{-\infty}^{\infty} p_\xi(x)dx = 1$$

Если ξ - дискретная случайная величина, принимающая значения x_1, $x_2,\ldots, x_i,\ldots$ с вероятностями $p_1, p_2,\ldots, p_i,\ldots$, то таблица вида

ξ	x_1	x_2	…	x_i	…
p	p_1	P_2	…	p_i	…

называется *рядом распределения дискретной случайной величины* или просто *распределением дискретной случайной величины*. В практических задачах именно такая форма представления распределения наиболее удобна.

Вероятность того, что значение случайной величины ξ попадет в интервал (a,b), вычисляется для непрерывной случайной величины по формуле

$$P(a<\xi<b)=F_\xi(b)-F_\xi(a)=\int_a^b p_\xi(t)dt,$$

а для дискретной случайной величины – по формуле

$$P(a<\xi<b)=\sum_{x_i\in(a,b)} p_i$$

Наиболее распространенные распределения дискретных случайных величин

Познакомимся с дискретными случайными величинами, которые чаще всего используются при решении практических задач. Эти случайные величины имеют биномиальное, геометрическое и пуассоновское распределения.

Биномиальное распределение (схема Бернулли). Пусть проводится серия из n независимых испытаний, каждое из которых заканчивается либо «успехом», либо «неуспехом». Пусть в каждом испытании (опыте) вероятность успеха p, а вероятность неудачи – $q = 1 - p$. С таким испытанием можно связать случайную величину ξ, равную числу успехов в серии из n испытаний. Эта величина принимает целые значения от 0 до n. Ее распределение называется *биномиальным* и определяется формулой Бернулли

$$p_k = P(\xi=k)=C_n^k p^k q^{n-k},$$

где $0<p<1,\ q=1-p,\qquad k=0,1,\dots,n,\qquad C_n^k=\frac{n!}{k!(n-k)!}$

Нетрудно убедиться, что $\sum_{k=0}^{n} p_k = 1.$

В Mathcad для вычисления плотности вероятности и функции распределения случайной величины, имеющей биномиальное распределение, предназначены функции dbinom(k, n, p) и pbinom(k, n, p), значения которых – соответственно p_k и $F(k)$.

Геометрическое распределение. Со схемой испытаний Бернулли можно связать еще одну случайную величину: η - число испытаний до первого успеха. Эта величина принимает бесконечное множество значений от 0 до $+\infty$, и ее распределение определяется формулой

$$p_k = P(\eta = k) = q^k p, \qquad k = 0, 1, \ldots, \qquad 0 < p < 1, \qquad q = 1 - p.$$

Используя формулу суммы бесконечно убывающей геометрической прогрессии, легко показать, что $\sum_{k=0}^{\infty} p_k = 1$.

В Mathcad для вычисления плотности вероятности и функции распределения случайной величины, имеющей геометрическое распределение, предназначены функции dgeom(k, p) и pgeom(k, p), значения которых – соответственно p_k и $F(k)$.

Пуассоновское распределение. Пуассоновское распределение имеет случайная величина μ, принимающая значения k = 0, 1, 2,… с вероятностями

$$p_k = P(\mu = k) = \frac{\lambda^k}{k!} e^{-\lambda}, k = 0,1,2,\ldots,$$

где $\lambda > 0$ - параметр пуассоновского распределения.

При любых $\lambda > 0$ $\sum_{k=0}^{\infty} p_k = 1$.

В Mathcad для вычисления вероятности и функции распределения случайной величины, имеющей пуассоновское распределение, предназначены функции dpois(k, λ) и ppois(k, λ), значения которых – соответственно p_k и $F(k)$.

Задание 2.1

- Для указанных значений параметров вычислите и постройте графически биномиальное, геометрическое и пуассоновское распределения. Проверьте для них равенство $\sum_{k=0}^{\infty} p_k = 1$.
- Постройте графики функций распределения. Вычислите вероятность попадания значений случайной величины в указанный интервал. Для каждого распределения найдите значение k, для которого величина $P(\xi=k)$ максимальна. Исследуйте зависимость этой вероятности от параметров распределения.

Порядок выполнения задания

1. Введите параметры распределения.
2. Определите интервал изменения значений случайной величины.

3. Определите вектор, номера компонент которого равны значениям случайной величины, и присвойте компонентам вектора значения вероятности соответствующих значений.
4. Определите функцию распределения случайной величины.
5. Постройте графики распределения и функции распределения случайной величины.
6. Найдите по графику наиболее вероятное значение случайной величины.
7. Введите в рабочий документ наибольшее значение вероятности (значение вероятности в точке, вычисленной в предыдущем пункте).
8. Вычислите сумму всех значений вероятностей.
9. Вычислите вероятность попадания значения случайной величины в указанный интервал как разность соответствующих значений функции распределения.
10. Измените значения параметров распределения и повторите вычисления. Сравните полученные результаты.
11. Выполните вычисления пп. 1-10 для всех приведенных в задании распределений.

Указание: Для того чтобы определить по графику распределения наиболее вероятное значение случайной величины, щёлкните в меню **Format** (Формат) в пункте **Graph** (График) по строке **Trace** (Следование), установите перекрестье маркера на точке максимума распределения и выведите в рабочий документ вероятность значения, указанного в окне X-**Value** (Величина X).

3. Непрерывные случайные величины

Наиболее распространённые распределения непрерывных случайных величин

Равномерное распределение. Непрерывная случайная величина ξ, принимающая значение на отрезке $[a,b]$, распределена равномерно на $[a,b]$, если плотность распределения $p_\xi(x)$ и функция распределения случайной величины ξ имеют соответственно вид

$$p_\xi(x) = \begin{cases} 0, & x \notin [a,b], \\ \dfrac{1}{b-a}, & x \in [a,b], \end{cases} \qquad F_\xi(x) = \begin{cases} 0, & x \le a, \\ \dfrac{x-a}{b-a}, & a < x \le b, \\ 1, & x > b. \end{cases}$$

В Mathcad значения в точке x плотности распределения и функции распределения случайной величины, имеющей равномерное распределение на отрезке *[a,b]*, вычисляются встроенными функциями соответственно dunif(x,a,b) и punif(x,a,b).

Экспоненциальное (показательное) распределение. Непрерывная случайная величина ξ имеет показательное распределение с параметром λ>0, если плотность распределения имеет вид

$$p_{\xi}(x)=\begin{cases}0, x<0\\ \lambda e^{-\lambda x}, x\ge 0\end{cases}$$

Отсюда видно, что показательно распределённая случайная величина принимает только неотрицательные значения. Функция распределения такой случайной величины имеет вид

$$F_{\xi}(x)=\begin{cases}0, x\le 0\\ 1-\lambda e^{-\lambda x}, x>0\end{cases}$$

В Mathcad значения в точке x плотности распределения и функции распределения случайной величины, имеющей экспоненциальное распределение с параметром λ, вычисляются встроенными функциями соответственно dexp(x, λ) и pexp(x, λ).

Нормальное распределение. Это распределение играет исключительно важную роль в теории вероятностей и математической статистике. Случайная величина ξ нормально распределена с параметрами a и σ, $\sigma>0$, если её плотность распределения имеет вид

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

Если случайная величина ξ имеет нормальное распределение с параметрами a и σ, то будем записывать это в виде ξ~N(a,σ). Случайная величина ξ имеет стандартное нормальное распределение, если a=0 и σ=1, ξ~N(0,1). Плотность стандартного нормального распределения имеет вид

$$p_{\xi}(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right),$$

а его функция распределения $F_{\xi}(x)=\Phi(x),\, где\, \Phi(x)$ – функция Лапласа:

$$\Phi_{\xi}(ч)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}\exp\left(-\frac{z^2}{2}\right)dz.$$

Функция распределения нормальной величины η~ N(a,σ) также выражается через функцию Лапласа: $F_{\eta}(x)=\Phi\left(\frac{x-a}{\sigma}\right)$.

В MathCad значения в точке x плотности распределения и функции распределения нормальной случайной величины с параметрами a, σ вычисляются встроенными функциями соответственно dnorm(x,a,σ) и pnorm(x,a,σ).

Распределение хи - квадрат (χ^2 - распределение) Пусть $\xi_1, \xi_2,\dots \xi_n$ – независимые случайные величины, каждая из которых имеет стандартное нормальное распределение N(0,1). Составим случайную величину $\chi^2_n = \xi^2_1 + \xi^2_2 + \dots + \xi^2_n$. Её распределение называется χ^2 - распределением с n степенями свободы. Для справочных целей приведём здесь выражение плотности распределения этой случайной величины:

$$P_n(z) = \begin{cases} 0, & z < 0 \\ \dfrac{1}{\Gamma(\frac{\pi}{2})2^{\frac{\pi}{2}}} z^{\frac{n-2}{2}} e^{-\frac{z}{2}}, & z \ge 0 \end{cases}$$ где Γ(x) – гамма–функция Эйлера.

$$\Gamma(x) = \int_0^{\infty} x^{z-1} e^{-z} dz$$

В Mathcad значения в точке *x* плотности распределения и функции χ^2 - распределения с *n* степенями свободы вычисляется встроенными функциями соответственно dchisq(*x*,*n*) и pchisq(*x*,*n*).

Задание 3.2

Постройте графики плотности распределения и функции распределения χ^2 с указанным числом степеней свободы.

Порядок выполнения задания

1. Постройте графики плотности распределения χ^2 с указанным числом степеней свободы.
2. Постройте графики функции распределения χ^2 с указанным числом степеней свободы.

Распределение Стьюдента. Пусть случайная величина ξ имеет стандартное нормальное распределение, а случайная величина χ_n^2 □ χ^2 - распределение с n степенями свободы. Если ξ и χ_n^2 независимы, то про случайную величину $\tau_n = \dfrac{\xi}{\sqrt{\chi_n^2 / n}}$ говорят, что она имеет распределение Стьюдента с числом степеней свободы *n*. Доказано, что плотность вероятности этой величины вычисляется по формуле

$$p_{\tau}n(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} (1 + \frac{x^2}{n})^{-\frac{n+1}{2}}, x \in R$$

При больших *n* распределение Стьюдента практически не отличается от N(0,1).

В Mathcad значения в точке x плотности распределения и функции Стьюдента с n степенями свободы вычисляются встроенными функциями соответственно dt(x,n) и pt(x,n).

Задание 3.3

Постройте графики плотности распределения и функции распределения Стьюдента с указанным числом степеней свободы.

Порядок выполнения задания

1. Постройте графики плотности распределения Стьюдента с указанным числом степеней свободы.
2. Постройте графики функции распределения Стьюдента с указанным числом степеней свободы.

F-распределение Фишера. Пусть случайные величины χ_n^2 и χ_m^2 независимы и имеют распределение χ^2 с n и m степенями свободы соответственно. Тогда случайная величина $F_{n,m} = \dfrac{x_n^2/n}{x_m^2/m}$ имеет F - распределение с плотностью вероятности

$$p_F(x) = \frac{\Gamma(\frac{n+m}{2})}{\Gamma(\frac{n}{2})\Gamma(\frac{m}{2})}(\frac{n}{m})^{n/2}\frac{x^{\frac{n-2}{2}}}{(1+\frac{nx}{m})^{\frac{n+m}{2}}},\ x>0.$$

Задание 3.4

Постройте графики плотности распределения и функции распределения Фишера для указанных значений n и m.

Порядок выполнения задания

1. Постройте графики плотности распределения Фишера для указанных значений n и m.
2. Постройте графики функции распределения Фишера для указанных значений n и m.

4. Квантили

При решении практических задач часто требуется найти значение x, при котором функция распределения принимает заданное значение, т.е. требуется решить уравнение $F_\xi(x) = p$. Решения такого уравнения в теории вероятностей называются квантилями.

Квантилью x_p (p-квантилью, квантилью уровня p) случайной величины ξ, имеющей функцию распределения $F_\xi(x)$, называют решение x_p уравнения $F_\xi(x) = p$, $p\in(0,1)$.

Для некоторых p уравнение $F_\xi(x) = p$ может иметь несколько решений, для некоторых – ни одного. Это означает, что для соответствующей случайной величины некоторые квантили определены неоднозначно, а некоторые квантили не существуют.

Квантили, наиболее часто встречающиеся в практических задачах, имеют свои названия:

медиана – квантиль уровня 0.5;

нижняя квартиль – квантиль уровня 0.25;

верхняя квартиль – квантиль уровня 0.75;

децили – квантили уровней 0.1, 0.2, …, 0.9;

процентили – квантили уровней 0.01, 0.02, …, 0.99.

Для тех распределений, для которых в Mathcad представлены встроенные функции плотности распределения и функции распределения, определены и встроенные функции вычисления квантилей.

Ниже приведены вычисленные в Mathcad медиана, верхняя и нижняя квартили и 0.95-квантиль для стандартного нормального распределения *N(0,1)*.

qnorm(0.5, 0, 1)=0 медиана

qnorm(0.25, 0, 1)=-0.67 нижняя квартиль

qnorm(0.75, 0, 1)=0.674 верхняя квартиль

qnorm(0.95, 0, 1)=1.645 0.95-квантиль

Задание 4.5

Найдите медиану, верхнюю и нижнюю квартили и *p* квартиль для заданного уровня *p* и для заданного распределения.

Порядок выполнения задания

1. Постройте графики плотности распределения для заданного распределения с указанными значениями параметров.
2. Постройте графики функции распределения для заданного распределения с указанными значениями параметров.

5. Числовые характеристики случайных величин

Каждая случайная величина полностью определяется своей функцией распределения. В то же время при решении практических задач достаточно знать несколько числовых параметров, которые позволяют представить основные особенности случайной величины в сжатой форме. К таким величинам относятся, в первую очередь, математическое ожидание и дисперсия.

Математическое ожидание случайной величины

Математическое ожидание – число, вокруг которого сосредоточены значения случайной величины.

Если ξ – дискретная случайная величина с распределением,

ξ	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

то ее математическим ожиданием – оно обозначается $M\xi$ – называется величина

$$M\xi = \sum_{i=1}^{n} p_i x_i,$$

если число значений случайной величины конечно. Если число значений случайной величины счетно, то

$$M\xi = \sum_{i=1}^{\infty} p_i x_i$$

При этом если ряд в правой части равенства расходится или сходится условно, то говорят, что случайная величина ξ не имеет математического ожидания.

Математическое ожидание непрерывной случайной величины с плотностью вероятностей $p_\xi(x)$ вычисляется по формуле

$$M\xi = \int_{-\infty}^{\infty} x\, p_\xi(x)\, dx$$

При этом если интеграл в правой части равенства расходится, то говорят, что случайная величина ξ не имеет математического ожидания.

При вычислении математического ожидания случайной величины полезны следующие его свойства:

- математическое ожидание константы равно этой константе, т.е. $Mc = c$;
- математическое ожидание – линейный функционал случайной величины, т.е. при произвольных постоянных a и b верно равенство $M(a\xi + b\eta) = aM\xi + bM\eta$;
- математическое ожидание произведения двух *независимых* случайных величин равно произведению их математических ожиданий, т.е. $M(\xi \cdot \eta) = M\xi \cdot M\eta$.

Приведем формулы математических ожиданий для наиболее известных распределений:

- биноминальное распределение: $(P(\xi = k) = C_n^k p^k q^{n-k})$: $M\xi = np$;
- геометрическое распределение: $\left(P(\xi = k) = q^k p\right): M\xi = \dfrac{q}{p}$;
- пуассоновское распределение: $\left(P(\xi = k) = \dfrac{\lambda^k}{k!} e^{-\lambda}\right): M\xi = \lambda$;

• равномерное распределение:

$$\left(p_{\xi}(x)=1/(b-a),\ x\in[a,b]\right): M\xi=\frac{a+b}{2};$$

• экспоненциальное (показательное) распределение:

$$\left(p_{\xi}(x)=\lambda e^{-\lambda x},\ x\geq 1\right): M\xi=\frac{1}{\lambda};$$

• нормальное распределение:

$$N(a,\sigma)\ \left(p_{\xi}(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^{2}\right]\right): M\xi=a;$$

• распределение хи-квадрат (χ^2 - распределение) с n степенями свободы:

$$\left(p_{\chi^2}(z)=\left[\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}\right]^{-1} z^{\frac{n-2}{n}} e^{-\frac{z}{2}},\ z>0\right): M\chi^2=n;$$

• распределение Стьюдента (t-распределение) с n степенями свободы:

$$\left(p_{t_n}(x)=\frac{1}{\sqrt{n\pi}}\Gamma\left(\frac{n+1}{2}\right)\Gamma\left(\frac{n}{2}\right)^{-1}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}\right): Mt_n=0;$$

• F-распределение Фишера с n и m степенями свободы:

$$\left(p_F(x)=\frac{\Gamma((n+m)/2)}{\Gamma(n/2)\Gamma(m/2)}\left(\frac{n}{m}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1}\left(1+\frac{nx}{m}\right)^{-\frac{n+m}{2}},\ x>0\right): MF=\frac{m}{m-2},\ m>2;$$

Дисперсия случайной величины

Дисперсия случайной величины характеризует меру разброса значений случайной величины около ее математического ожидания. Если случайная величина ξ имеет математическое ожидание $M\xi$, то *дисперсией* случайной величины ξ называется величина $D\xi = M(\xi - M\xi)^2$. Легко показать, что $D\xi = M\xi^2 - (M\xi)^2$. Эта универсальная формула одинаково хорошо применима как для дискретных случайных величин, так и для непрерывных. Величина $M\xi^2$ вычисляется по формулам:

$$M\xi^2=\sum_{i=1}^{n} p_i x_i^2, \qquad M\xi^2=\int_{-\infty}^{\infty} x^2 p_{\xi}(x)dx$$

для дискретных и непрерывных случайных величин соответственно.

Еще одним параметром для определения меры разброса значений случайной величины является *среднеквадратическое отклонение* σ_ξ, связанное с дисперсией соотношением $\sigma_\xi = \sqrt{D\xi}$.

Перечислим основные свойства дисперсии:

- дисперсия любой случайной величины неотрицательна: $D\xi \geq 0$;
- дисперсия константы равна нулю: $Dc = 0$;
- для произвольной константы c: $D(c\xi) = c^2 D\xi$;
- дисперсия суммы (разности) двух *независимых* случайных величин равна сумме их дисперсий: $D(\xi \pm \eta) = D\xi + D\eta$.

Приведем формулы для дисперсий наиболее известных стандартных распределений:

- биноминальное распределение: $D\xi = npq$;
- геометрическое распределение: $D\xi = \dfrac{q}{p^2}$;
- пуассоновское распределение: $D\xi = \lambda;$;
- равномерное распределение: $D\xi = \dfrac{(b-a)^2}{12}$;
- экспоненциальное (показательное) распределение: $D\xi = \lambda^{-2}$;
- нормальное распределение: $N(a,\sigma)$: $D\xi = \sigma^2$;
- распределение хи-квадрат (χ^2-распределение) с n степенями свободы: $D\xi^2 = 2n$;
- распределение Стьюдента с n степенями свободы: $D\xi = \dfrac{n}{n-2}$, $n > 2$;
- F-распределение Фишера с n и m степенями свободы:

$$DF = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)}, \ m > 4;$$

Задание 5.6

Вычислите математическое ожидание и дисперсию случайной величины $\xi = S(\eta)$, которая представляет собой площадь указанной в задании геометрической фигуры, для случайной величины η, имеющей заданное распределение.

Порядок выполнения задания

1. Запишите выражение для функции $\xi = S(\eta)$ от случайной величины η, определяющей площадь фигуры.
2. Вычислите математическое ожидание случайной величины ξ.
3. Вычислите математическое ожидание случайной величины ξ^2 .
4. Вычислите дисперсию случайной величины $\xi = S(\eta)$ по формуле $D\xi = M\xi^2\text{-}(M\xi)^2$.

Пример выполнения задания

Случайная величина η распределена равномерно на промежутке [0,1]. Найдем математическое ожидание и дисперсию площади квадрата со стороной η, т.е. характеристики случайной величины $\xi = S(\eta) = \eta^2$.

Математическое ожидание площади квадрата ξ

$$M(\xi) := \int_1^2 x^2 \cdot \frac{1}{2-1}\,dx \qquad M(\xi) \to \frac{7}{3}$$

Математическое ожидание квадрата случайной величины ξ

$$M2(\xi) := \int_1^2 x^4 \cdot \frac{1}{2-1}\,dx \qquad M2(\xi) \to \frac{31}{5}$$

Дисперсия площади квадрата ξ

$$D(\xi) := M2(\xi) - M(\xi)^2 \qquad D(\xi) \to \frac{34}{45}$$

Указание. Математическое ожидание и дисперсию площади квадрата со стороной η вычислите символьно по формулам $M\xi = M\eta^2,\ D\xi = M\eta^4 - (M\xi)^2$. Определите искомые математическое ожидание и дисперсию как функции переменной ξ.

6. Моменты

В теории вероятностей и математической статистике, помимо математического ожидания и дисперсии, используются и другие числовые характеристики случайных величин. В первую очередь это *начальные* и *центральные* моменты.

Начальным моментом k-го порядка случайной величины ξ называется математическое ожидание k-ой степени случайной величины ξ , т.е. $\alpha_k = M\,\xi^k$.

Центральным моментом k-го порядка случайной величины ξ называется величина μ_k , определяемая формулой $\mu_k = M(\xi - M\xi)^k$.

Заметим, что математическое ожидание случайной величины - начальный момент первого порядка, $\alpha_1 = M\xi$, а дисперсия - центральный момент второго порядка: $\mu_2 = M\,(\xi - M\xi)^2 = D\xi$.

Существуют формулы, позволяющие выразить центральные моменты случайной величины через ее начальные моменты. Одна из таких формул приведена выше: $D\xi = M(\xi - M\xi)^2 = \mu_2 - \alpha_1^2$.

В дальнейшем будет использована формула $\mu_3 = \alpha_3 - 3\alpha_2\,\alpha_1 + 2\alpha_1^3$.

Нетрудно понять, что если плотность распределения вероятностей случайной величины симметрична относительно прямой $x = M\xi$, то все ее центральные моменты нечетного порядка равны нулю.

В теории вероятностей и математической статистике в качестве меры асимметрии распределения служит коэффициент асимметрии, который определяется формулой:

$$\beta_\xi = \frac{\mu_3}{\sigma_\xi^3},$$

где μ_3 – центральный момент третьего порядка; $\sigma_\xi = \sqrt{D\xi} = \sqrt{\mu_2}$ – среднеквадратичное отклонение.

Коэффициент асимметрии – безразмерная величина, а по его знаку можно судить о характере асимметрии.

Указание. Для того чтобы вычислить значение коэффициента асимметрии, выделите выражение для него, щелкните в строке **Floating Point** в меню **Symbolics** и укажите в окне диалога число десятичных знаков в выводе.

Задание 5.7

Вычислите коэффициент асимметрии случайной величины ξ с заданным распределением.

Порядок выполнения задания

1. Определите значения параметров распределения случайной величины.
2. Определите центральный момент третьего порядка как функцию параметров распределения случайной величины.
3. Определите центральный момент второго порядка как функцию параметров распределения случайной величины.
4. Вычислите коэффициент асимметрии.
5. Постройте график плотности вероятности.

Эксцесс

Нормальное распределение наиболее часто используется в теории вероятностей и математической статистике, и поэтому график плотности вероятностей нормального распределения стал своего рода эталоном, с которым сравнивают другие распределения. Одним из параметров, определяющих отличие сравниваемого распредления от нормального, является эксцесс.

Эксцесс γ случайной величины ξ определяется равенством

$$\gamma = \frac{\mu_4}{(D\xi)^2} - 3.$$

У нормального распределения, естественно, $\gamma = 0$. Если $\gamma > 0$, то это означает, что график плотности вероятностей $p_\xi(x)$ сильнее "заострен", чем у нормального распределения, если же $\gamma < 0$, то "заостренность" графика $p_\xi(x)$ меньше, чем у нормального распределения.

Задание 5.8

Вычислите эксцесс случайной величины ξ с заданным распределением.

Порядок выполнения задания

1. Определите значения параметров распределения случайной величины.
2. Определите центральный момент четвертого порядка как функцию параметров распределения случайной величины.
3. Определите центральный момент второго порядка как функцию параметров распределения случайной величины.
4. Вычислите коэффициент асимметрии.
5. Постройте график плотности вероятности.

Указание. Mathcad не справляется с вычислением интегралов функций, заданных разными выражениями на разных промежутках. Поэтому при вычислении моментов используйте свойство интеграла по симметричному промежутку от чётной функции. Для того чтобы определить плотность вероятностей равномерного распределения, щёлкните по кнопке Add Line в панели , введите в первой строке выражение для функции, щёлкните по кнопке IF и введите условие; аналогично определите во второй строке функцию на втором промежутке.

Литература

1. Андронов А.М. Теория вероятностей и математическая статистика: Учебник для вузов / А.М. Андронов, Е.А. Копытов, Л.Я. Гринглаз. - СПб.: Питер, 2004. - 461с.

2. Плис А.И. Mathcad: математический практикум для экономистов и инженеров / А.И.Плис, Н.А.Сливина. - М.: Финансы и статистика, 2000. - 656с.

3. Боровиков В. STATISTICA: искусство анализа данных на компьютере / В. Боровиков. - СПб.: Питер, 2001.- 656с.

СОДЕРЖАНИЕ

Составитель: Нелля Михайловна Новикова