

Содержание

I	Теория вероятностей	6
1	Вероятностные пространства	6
1.1	Равновероятные события	6
2	Сложные события	7
2.1	Умножение вероятностей	7
2.2	Сложение вероятностей	8
3	Аксиомы вероятностей	8
4	Примеры	9
4.1	Биномиальное распределение	10
4.2	Полиномиальное распределение	11
5	Случайные числа	12
5.1	Математическое ожидание	12
5.2	Дисперсия	13
5.3	Моменты	13
5.4	Семиинварианты	14
5.4.1	Свойства семиинвариантов	14
6	Дискретные распределения вероятностей	15
6.1	Равномерное распределение	15
6.2	Бинарное распределение	16
6.3	Биномиальное распределение	17
6.4	Распределение Пуассона	17
6.5	Распределение Паскаля	18
6.6	Step-распределение	19
6.7	MultiStep-распределение	20
7	Непрерывные распределения вероятностей	20
7.1	Равномерное распределение	21
7.2	Бета-распределение	22
7.3	Нормальное распределение	23

7.4	Интеграл ошибок	24
7.5	Линейная суперпозиция нормально распределенных величин	26
7.6	Экспоненциальное распределение	26
7.7	Гамма-распределение	27
7.8	Многомерное нормальное распределение	28
8	Функции случайной величины	29
8.1	Квадрат нормальной величины	29
8.2	Распределение Максвелла	30
8.3	Распределение Релея	32
8.4	Распределение χ^2	33
8.5	Распределение Стюдента	35
8.6	Распределение выборочных дисперсий	36
9	Предельные теоремы	37
9.1	Первая предельная теорема	37
9.2	Теорема Муавра-Лапласа	37
9.3	Центральная предельная теорема	38
9.4	Использование ЦПТ	39
II	Математическая статистика	41
10	Основная задача математической статистики	41
11	Теория Байеса	41
11.1	Статистическая зависимость	42
11.2	Теорема Байеса	42
11.3	Методика Байеса	43
12	Параметр бинарного распределения	43
12.1	Априорная вероятность	45
12.2	Прогноз	46

13	Параметры нормального распределения	48
13.1	Распределение дисперсии	50
13.2	Распределение математического ожидания	51
13.3	Прогноз	52
13.4	Различение математических ожиданий	53
14	Регрессия	54
14.1	Регрессия прямой линией	56
14.2	Регрессионный прогноз	57
15	Статистические гипотезы	59
15.1	Критерий χ^2	59
16	Статистическое моделирование	61
16.1	Датчики случайных чисел	61
16.2	Моделирование первой предельной теоремы	61
16.3	Моделирование multistep-распределения	62
16.4	Вычисление интегралов методом Монте-Карло	62
16.5	Пакет расширения “Статистика”	65
III	Задания	66
17	Задания	66
17.1	Распределения	66
17.2	Семиинварианты	66
17.3	Распределение Стьюдента	66
17.4	Моделирование распределений	67
17.5	Статистика параметров	67
17.6	Критерии значимости	67
17.7	Регрессия	68
17.8	Интегрирование методом Монте-Карло	68
	Литература	68

Часть I

Теория вероятностей

1 Вероятностные пространства

1.1 Равновероятные события

- 2 события: орел – решка, любит – не любит, быть – или не быть.
- 6 событий: числа 1 – 6 на игральной кости.
- 36 игральные карт.
- 19 студентов.

Главная характеристика n – число событий.

Вероятность каждого события в этих примерах одинакова

$$p_i = \frac{1}{n}.$$

Суммарная вероятность

$$\sum_{i=1}^n p_i = n \frac{1}{n} = 1. \quad (1)$$

Вопрос: какова вероятность выпадения единицы на игральной кости в виде

- тетраэдра?
- икосаэдра?
- додекаэдра?

2 Сложные события

Бросается сразу две монеты. Сколько может выпасть орлов?

0, 1, 2. Три события. Каковы их вероятности? Д'Аламбер в “Энциклопедии” (1751 – 1780, 35 томов) – $1/3$.

Какова вероятность, выйдя из дома, встретить динозавра?

Ответ: $1/2$ – или встречу или не встречу.

События могут быть и не равновероятными.

2.1 Умножение вероятностей

Итак – две монеты, четыре равновероятных события:

- орел – орел;
- орел – решка;
- решка – орел;
- решка – решка.

Почему четыре? Пространство событий для двух монет есть *прямое произведение* пространств каждой монеты: каждому событию первой монеты соответствуют два события второй.

$$n_2 = n_1 \times n_1; \quad 2 \times 2 = 4.$$

Все события равновероятны. Вероятность любой выбранной пары (например, решка – орел) равна

$$p_{ro} = \frac{1}{n_2} = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = p_r \cdot p_o.$$

Это соотношение между вероятностями отдельных *статистически независимых* событий и вероятностью совместного их появления называется *теоремой об умножении вероятностей*:

$$p_{ab} = p_a \cdot p_b. \quad (2)$$

Сколько событий в пространстве трех монет?

$$n_3 = n_2 \cdot n_1 = n_1 \cdot n_1 \cdot n_1 = 2 \times 2 \times 2 = 2^3 = 8,$$

(а не $2 \times 3 = 6$).

Сколько событий в пространствах

- двух игральных костей?
- монеты и игральной кости?
- двух игральных костей и трех монет?

2.2 Сложение вероятностей

Какова вероятность из колоды в 36 карт вытащить любую пиковую карту?

Вероятность любой карты — $1/36$, пиковых карт — 9. В *множестве* всех 36-и карт пиковые карты образуют *подмножество* из девяти карт.

Вероятность попасть в подмножество равна *сумме вероятностей* элементов, составляющих это подмножество:

$$p_{\spadesuit} = \sum_{i=1}^9 p_{\spadesuit i} = 9 \cdot \frac{1}{36} = \frac{1}{4}.$$

Можно эту проблему решить и другим путем: имеется четыре равновероятные масти, значит вероятность каждой из них равна $1/4$.

Какова вероятность из колоды в 36 карт вытащить даму?

Всего дам — четыре: $p_D = 4 \times 1/36 = 1/9$.

Какова вероятность из колоды в 36 карт вытащить пиковую даму?

$$p_{\spadesuit D} = p_{\spadesuit} \cdot p_D = \frac{1}{4} \cdot \frac{1}{9} = \frac{1}{36},$$

хотя ответ очевиден, так как она — единственная в колоде.

3 Аксиомы вероятностей

Ранее мы стартовали от равновероятных событий. Теперь можно отойти от этого ограничения.

- *Вероятностное пространство* \mathcal{A} – это пространство событий a_i , $1 \leq i \leq n$, где n – *число событий*, каждому элементу которого a_i приписана *вероятность* p_i .
- *Вероятность* события p_i – это неотрицательное число $0 \leq p_i \leq 1$,
- Вероятности всех элементов множества подчиняются *условию нормировки*:

$$\sum_{i=1}^n p_i = 1. \quad (3)$$

В двух предыдущих параграфах мы ознакомились с двумя основными операциями с пространствами событий и вероятностями:

- Прямое произведение пространств событий $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$. Число элементов прямого произведения равно произведению чисел элементов составляющих пространств: $n_C = n_A \cdot n_B$. Вероятность совместного появления пары независимых событий равно *произведению* вероятностей каждого из событий $p_{ab} = p_a \cdot p_b$.
- Выделение подмножества пространства событий в одно событие. Вероятность попадания в это подмножество равна *сумме* вероятностей всех элементов, выделенных в данное подмножество.

Эти операции являются основными в построении пространств сложных событий из некоторых первоначальных с заданными вероятностями.

Сформулируем **основную задачу теории вероятностей**:

Вычисление вероятностей сложных событий по заданным вероятностям элементарных событий.

4 Примеры

Теперь легко ответить, например, на вопрос:

какова вероятность выпадения одного орла и одной решки при выбрасывании двух монет? Подмножество решений содержит два события: (о, р) и (р, о), вероятность каждого $1/4$, поэтому искомая вероятность равна $1/2$. При бросании двух монет возможны четыре комбинации, приводящие к трем значениям числа орлов n_o :

- (р, р): $n_o = 0$; $p_0 = 1/4$.
- (р, о) (о, р): $n_o = 1$; $p_1 = 1/2$.
- (о, о): $n_o = 2$; $p_2 = 1/4$.

Какова вероятность выпадения одного орла и двух решек при выбрасывании трех монет?

$$n_{1/3} = 3 \left(\frac{1}{2} \right)^3 = \frac{3}{8},$$

так как этот орел может выпасть как на первой, так и на второй или третьей монете.

4.1 Биномиальное распределение

Стартует от *бинарного* распределения: пространство событий из двух элементов, которые абстрактно можно именовать + и -. Вероятность плюса — p — вещественное число от 0 до 1, вероятность минуса — $(1 - p)$. При бросании n монет пространство событий — это n - кратное произведение бинарного пространства самого на себя с числом элементов 2^n .

В каждой реализации среди n составляющих элементарных событий выпадает m плюсов и остальные $n - m$ — минусы. Таким образом порождается новое вероятностное пространство, пространство случайного числа m , принимающего значения от нуля до n .

Если в какой-то реализации выпало m плюсов (и $n - m$ минусов), по теореме об умножении вероятностей вероятность такого события равна $p^m (1 - p)^{n-m}$ независимо от порядка выпадения плюсов и минусов. Но m плюсов может выпасть в разных комбинациях, например,;

$$(+ - - + + -), (- + - + - +), \dots$$

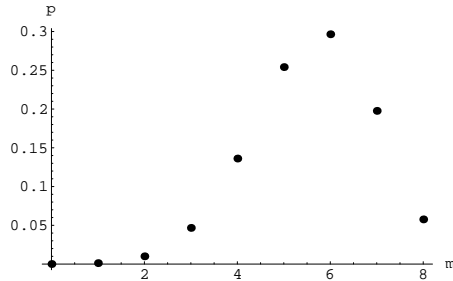
Число таких комбинаций – полное число перестановок, деленное на число перестановок в плюсах и число перестановок в минусах:

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

Вероятности событий с одинаковым m одинаковы $p^m (1-p)^{n-m}$, так что вероятность числа m (выпадения m плюсов при n испытаниях) равна

$$P_n^m = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}. \quad (4)$$

При $p = 0.7$, $n = 8$ это распределение представлено на графике



Теория этого распределения была построена Якобом Бернулли (1654–1705) и нередко именуется *распределением Бернулли*.

Таким образом n экземпляров бинарного пространства с 2^n элементов привели к пространству событий из $n + 1$ элементов – значений числа m – с неравномерным распределением вероятностей, определяемых формулой (4).

4.2 Полиномиальное распределение

Как биномиальное распределение порождается многократным прямым произведением (многократной выборкой) вероятностного пространства с двумя состояниями, так многократная выборка из пространства с l состояниями и вероятностями p_1, p_2, \dots, p_l при условии $\sum_1^n p_i = 1$ порождает *полиномиальное распределение*.

После n -кратной выборки реализации i -х состояний (m_1, \dots, m_l) ; $\sum m_i = n$ определяются вероятностями полиномиального распределения:

$$P_{(m_1, m_2, \dots, m_l)} = \frac{n!}{m_1! m_2! \dots m_l!} p_1^{m_1} p_2^{m_2} \dots p_l^{m_l}. \quad (5)$$

5 Случайные числа

Если случайным событием является значение какого-то числа (например, x), могущего принимать n различных значений в каком-то множестве $(x_i; 1 \leq i \leq n)$ с заданными или вычисленными вероятностями p_i , то такое вероятностное пространство называется *случайным числом* или *случайной величиной*.

Числа можно складывать, вычитать, умножать и делить и это приводит к возможности определять на вероятностных пространствах некоторые числовые характеристики этих пространств.

5.1 Математическое ожидание

Важнейшей числовой характеристикой случайной величины является ее *математическое ожидание*:

$$\langle x \rangle = \sum_{i=1}^n x_i p_i. \quad (6)$$

Значение математического ожидания лежит между минимальным и максимальным значениями случайной величины.

$$x_{min} \leq \langle x \rangle \leq x_{max}.$$

При многократной выборке случайного числа математическое ожидание $\langle x \rangle$ оценивает среднее значение в множестве реализаций.

Можно определить функцию случайной величины $f(x)$ и вычислить математическое ожидание этой функции:

$$\langle f(x) \rangle = \sum_{i=1}^n f(x_i) p_i. \quad (7)$$

Например, математическое ожидание константы равно самой константе $\langle c \rangle = c$.

Математическое ожидание линейной функции:

$$\langle Ax + B \rangle = A \langle x \rangle + B.$$

Математическое ожидание разности случайной величины с ее математическим ожиданием (константой) всегда равно нулю:

$$\langle (x - \langle x \rangle) \rangle = \langle x \rangle - \langle x \rangle = 0. \quad (8)$$

Математическое ожидание суммы двух *статистически независимых* случайных величин равно сумме их математических ожиданий.

$$\langle x + y \rangle = \langle x \rangle + \langle y \rangle.$$

Математическое ожидание случайной величины или ее функции не есть случайная величина – это однозначная константа, определяемая допустимыми значениями случайной величины и их вероятностями.

5.2 Дисперсия

Оценку вероятности отклонения случайной величины от ее математического ожидания дает *дисперсия*:

$$D_x = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 - 2x \langle x \rangle + \langle x \rangle^2 \rangle = \quad (9)$$

$$\langle x^2 \rangle - \langle x \rangle^2 \geq 0.$$

Дисперсия равна нулю только у константы – в случае, если для всех событий случайная величина одна и та же (например, на всех гранях игрального кубика написано одно и то же число). Во всех реальных случаях дисперсия положительна.

5.3 Моменты

Дисперсия в последней формуле вычисляется через математическое ожидание квадрата и первой степени случайной величины. Математическое ожидание l -й степени случайной величины называется l -м

моментом:

$$m_l = \langle x^l \rangle = \sum_{i=1}^n (x_i)^l p_i. \quad (10)$$

Отметим, что во все введенные определения сами вероятности входят линейно.

Моменты — для некоторых аналитических распределений сразу все — могут быть определены через *производящую функцию моментов*:

$$F(t) = \langle e^{tx} \rangle = \sum_{i=1}^n e^{tx_i} p_i = \sum_{l=0}^{\infty} \frac{t^l}{l!} m_l; \quad m_l = \left(\frac{d}{dt} \right)^l F(t)|_{t=0}. \quad (11)$$

Докажите, что второй момент больше квадрата первого.

5.4 Семиинварианты

Наиболее сильным аппаратом теории вероятности являются *семиинварианты*, тесно связанные с моментами. Производящая функция семиинвариантов есть логарифм производящей функции моментов:

$$\chi(t) = \ln(F(t)) = t \chi_1 + \frac{t^2}{2} \chi_2 + \dots + \frac{t^k}{k!} \chi_k + \dots, \quad (12)$$

а коэффициенты разложения ее в ряд Тейлора по степеням параметра t и являются семиинвариантами:

$$\chi_k = \left(\frac{d}{dt} \right)^k \chi(t)|_{t=0}. \quad (13)$$

5.4.1 Свойства семиинвариантов

- Нулевой семиинвариант равен нулю, первый — математическое ожидание случайной величины, а второй — ее дисперсия.
- Любой семиинвариант суммы двух независимых случайных величин равен сумме соответствующих семиинвариантов слагаемых:

$$F_{[x+y]}(t) = \sum_{i=1}^n \sum_{j=1}^m e^{t(x_i+y_j)} p_{[x]_i} p_{[y]_j} = F_{[x]}(t) \cdot F_{[y]}(t);$$

$$\chi_{[x+y]}(t) = \ln(F_{[x+y]}(t)) = \ln(F_{[x]}(t) \cdot F_{[y]}(t)) = \chi_{[x]}(t) + \chi_{[y]}(t).$$

Отсюда вследствие (13)

$$\chi_{[x+y]_k} = \chi_{[x]_k} + \chi_{[y]_k}. \quad (14)$$

Соответствующие семиинварианты при сложении случайных величин складываются.

В частности, *дисперсия суммы независимых случайных величин равна сумме их дисперсий*. Это нетривиально, так как дисперсия – квадратичная структура.

- При масштабном преобразовании

$$\langle e^{t(\lambda x)} \rangle = \langle e^{(t\lambda)x} \rangle; \quad \chi_{[\lambda x]}(t) = \chi_{[x]}(\lambda t); \quad \chi_{[\lambda x]_k} = \lambda^k \chi_{[x]_k} \quad (15)$$

k -й семиинвариант умножается на масштабный множитель в k -й степени. В частности, дисперсия умножается на квадрат масштаба.

Третий семиинвариант χ_3 называют *асимметрией*, так как он равен нулю для распределений, симметричных относительно математического ожидания, а четвертый χ_4 называют *эксцессом*.

6 Дискретные распределения вероятностей

Мы приведем основные характеристики ряда распределений случайных величин – в этом разделе – дискретных, а в следующем – непрерывных.

6.1 Равномерное распределение

Равномерное распределение описывает n равновероятных событий:

$$k = 1, 2, \dots, n; \quad p_k = \frac{1}{n}.$$

Математическое ожидание и дисперсия:

$$\langle k \rangle = \frac{n+1}{2}; \quad D_k = \frac{(n^2-1)}{12}. \quad (16)$$

Для вычисления дисперсии полезно вспомнить некоторые формулы сумм:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}; \quad \sum_{k=1}^n \frac{k(k+1)}{2} = \frac{n(n+1)(n+2)}{6}.$$

Так как можно представить

$$k^2 = \frac{k(k+1)}{2} + \frac{(k-1)k}{2}; \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6},$$

то математическое ожидание k^2

$$\langle k^2 \rangle = \frac{1}{n} \sum_{k=1}^n k^2 = \frac{(n+1)(2n+1)}{6},$$

откуда вычислим дисперсию:

$$D = \langle k^2 \rangle - \langle k \rangle^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2} \right)^2 = \frac{n^2 - 1}{12}.$$

Вычислите математическое ожидание и дисперсию для случайной величины, определяемой обычной игральной костью.

6.2 Бинарное распределение

Бинарное распределение числа k , принимающего всего два значения: 1 или 0 – с вероятностями p и $1-p$ соответственно. Производящая функция семиинвариантов находится просто:

$$F(t) = p e^{t \cdot 1} + (1-p) e^{t \cdot 0} = p e^t + 1 - p; \quad \chi(t) = \ln(p e^t + 1 - p),$$

а из нее и сами семиинварианты:

$$\frac{d\chi(t)}{dt} = \frac{p e^t}{p e^t + 1 - p} = \frac{p}{p + (1-p) e^{-t}}; \quad \chi_1 = \langle k \rangle = p. \quad (17)$$

$$\frac{d^2\chi(t)}{dt^2} = \frac{p(1-p) e^t}{(p e^t + 1 - p)^2}; \quad \chi_2 = D_k = p(1-p). \quad (18)$$

6.3 Биномиальное распределение

Биномиальное распределение (4) числа m , принимающего значения от нуля до n можно представить как сумму n случайных чисел k , принимающих всего два значения: 0 или 1 с вероятностями $1 - p$ и p соответственно, подчиняющихся биномиальному распределению (биномиальному распределению при $n = 1$).

В биномиальном распределении складываются n независимых случайных чисел k с одинаковыми характеристиками $m = \sum_{i=1}^n k_i$, так что соответствующие семинварианты просто умножаются на n :

$$\langle m \rangle = n \langle k \rangle = np; \quad D_m = n D_k = np(1 - p). \quad (19)$$

6.4 Распределение Пуассона

Является пределом биномиального распределения при $n \rightarrow \infty$ с заданным математическим ожиданием $\langle m \rangle = \mu = np$, так что при этом $p = \mu/n \rightarrow 0$. Еще не расписав самого распределения, можно записать дисперсию:

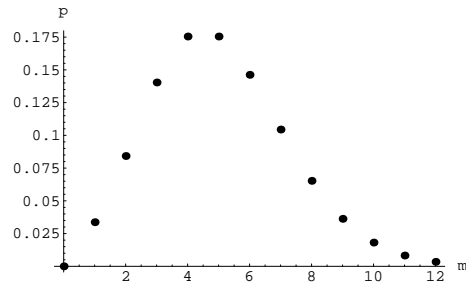
$$D_m = \lim_{n \rightarrow \infty, p = \mu/n} np(1 - p) = \mu. \quad (20)$$

Дисперсия совпадает с математическим ожиданием.

Само распределение целого числа m , принимающего значения от нуля до бесконечности, определяется однопараметрическим выражением (параметр μ):

$$P_m = \lim_{n \rightarrow \infty, p = \mu/n} \frac{n!}{m! (n - m)!} p^m (1 - p)^{n - m} = \frac{\mu^m}{m!} e^{-\mu}. \quad (21)$$

При $\mu = 5$ его представляет график:



Производящая функция моментов

$$F(t) = e^{-\mu} \sum_{m=0}^{\infty} e^{mt} \frac{\mu^m}{m!} = e^{-\mu} e^{\mu e^t} = e^{\mu(e^t - 1)}$$

определяет производящую функцию семиинвариантов

$$\chi(t) = \ln F(t) = \mu(e^t - 1). \quad (22)$$

Отсюда видно, что все семиинварианты, а не только первый и второй, одинаковы и равны μ .

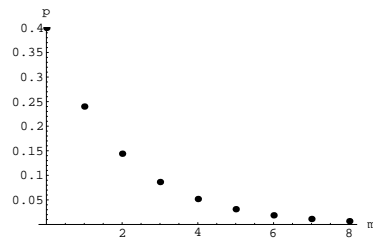
6.5 Распределение Паскаля

Распределение Пуассона может иметь максимум, распределение Паскаля – также вероятность целого числа m от нуля до бесконечности с единственным параметром γ ($0 < \gamma < 1$):

$$p_m = (1 - \gamma) \gamma^m \quad (23)$$

монотонно убывает с ростом m .

Для $\gamma = 0.6$ построен график:



Производящая функция моментов

$$F(t) = \sum_{k=0}^{\infty} e^{t n} p_n = (1 - \gamma) \sum_{k=0}^{\infty} (\gamma e^t)^k = \frac{1 - \gamma}{1 - \gamma e^t}$$

определяет производящую функцию семиинвариантов:

$$\chi(t) = \ln(1 - \gamma) - \ln(1 - \gamma e^t); \quad \frac{d\chi}{dt} = \frac{\gamma}{e^{-t} - \gamma}; \quad \frac{d^2\chi}{dt^2} = \frac{\gamma e^{-t}}{(e^{-t} - \gamma)^2},$$

откуда находятся математическое ожидание и дисперсия случайного целого числа k :

$$\langle k \rangle = \frac{\gamma}{1 - \gamma}; \quad D_k = \frac{\gamma}{(1 - \gamma)^2}.$$

Вместо формального параметра γ можно ввести параметр μ – математическое ожидание (как и в распределении Пуассона $0 < \mu < \infty$). Тогда

$$\langle k \rangle = \mu = \frac{\gamma}{1 - \gamma}; \quad \gamma = \frac{\mu}{\mu + 1};$$

$$1 - \gamma = \frac{1}{\mu + 1}; \quad D_k = \frac{\gamma}{(1 - \gamma)^2} = \mu(\mu + 1),$$

а само распределение вероятности числа k представляется в виде

$$p_k = \frac{\mu^k}{(\mu + 1)^{k+1}}.$$

Распределение Паскаля является базовым в *теории очередей*.

6.6 Step-распределение

Вероятность сдвига по целочисленной цепочке на один шаг вперед равна p , на один шаг назад – q и вероятность остаться на месте $1 - p - q$. Естественно, $p + q \leq 1$. Случайная величина принимает три значения $m_1 = -1$, $m_2 = 0$, $m_3 = 1$ с вероятностями q , $1 - p - q$, p соответственно.

Производящая функция моментов

$$F(t) = e^{-1 \cdot t} q + e^{0 \cdot t} (1 - p - q) + e^{+1 \cdot t} p = 1 + p(e^t - 1) + q(e^{-t} - 1)$$

определяет производящую функцию семиинвариантов

$$\chi(t) = \ln(1 + p(e^t - 1) + q(e^{-t} - 1)).$$

Разложение ее в ряд по степеням t определяет семиинварианты, в частности, математическое ожидание и дисперсию:

$$\langle m \rangle = p - q; \quad D_m = p + q - (p - q)^2. \quad (24)$$

6.7 MultiStep-распределение

Если проводится n единичных шагов со step-распределением, то в результате появляется сдвиг в положительном или отрицательном направлении на величину k , где $-n \leq k \leq n$.

Явный вид распределения вероятности случайной величины k

$$P_{k/n,p,q} = \sum_l \frac{n!}{l!(k+l)!(n-2l-k)!} q^l p^{k+l} (1-p-q)^{n-2l-k}$$

мало полезен. Однако, математическое ожидание и дисперсия (и другие семиинварианты) вычисляются легко: они равны семиинвариантам step-распределения, умноженным на n :

$$\langle m \rangle = n(p - q); \quad D_m = n(p + q - (p - q)^2). \quad (25)$$

7 Непрерывные распределения вероятностей

Для непрерывно распределенных величин x определяется вероятность попадания в бесконечно малый интервал заданной величины dx :

$$dp(x) = \rho(x) dx; \quad \rho \geq 0. \quad (26)$$

где функция $\rho(x)$ называется *плотностью распределения вероятностей*.

Интеграл в некотором интервале

$$\int_a^b \rho(x) dx = \int_a^b dp(x)$$

с точки зрения теории вероятностей определяет вероятность нахождения случайной величины в интервале (a, b) , а с точки зрения геометрии определяет площадь под кривой плотности вероятности на участке (a, b) .

На всем интервале изменения случайной величины эта вероятность равна единице, а это значит, что площадь под всей кривой плотности вероятности равна единице.

Если интервал изменения случайной величины разделен на n частей, то вероятность попадания в ту или иную часть равна площади под кривой плотности вероятности в этой части.

В многомерном случае (n переменных) определяется вероятность попадания в n -мерный объем:

$$dp(x_1, x_2, \dots, x_n) = \rho(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n.$$

7.1 Равномерное распределение

Непрерывная случайная величина x принимает значения на интервале (a, b) равновероятно: вероятность попасть в любой интервал между a и b пропорциональна величине интервала: $p(\Delta x) = c \Delta x$, а так как вероятность попадания в произвольную точку интервала равна единице

$$c(b - a) = 1; \quad c = \frac{1}{b - a}; \quad dp(x) = \frac{1}{b - a} dx.$$

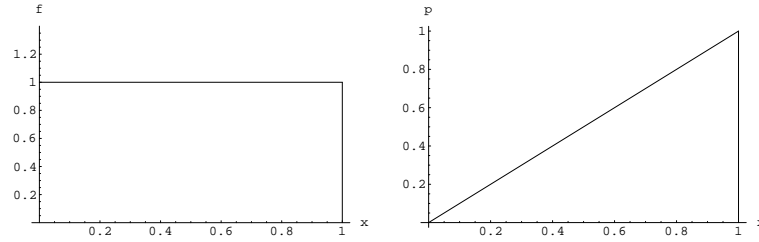
Это распределение имеет два параметра a и b .

Линейным преобразованием любой интервал (a, b) может быть приведен к единичному $(0, 1)$ и распределение становится *стандартным* без параметров.

$$dp(x) = 1 \cdot dx; \quad p(x) = x; \tag{27}$$

Плотность вероятности на интервале от нуля до единицы постоянна и равна единице, а сама вероятность попадания в интервал от нуля

до x растет линейно, достигая значения полной вероятности единица при $x = 1$:



7.2 Бета-распределение

Переменная x изменяется на интервале $(0, 1)$ с дифференциалом вероятности, определяемом двумя параметрами α и β :

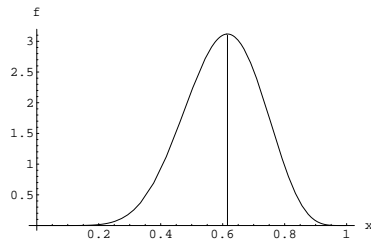
$$dp(x/\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} dx \quad (28)$$

Порождается β -интегралом Эйлера:

$$\int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

В частности, равномерное распределение является частным случаем бета-распределения с $\alpha = \beta = 1$.

На графике представлено распределение с $\alpha = 9$, $\beta = 6$:



Математическое ожидание

$$\langle x \rangle = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^\alpha (1-x)^{\beta-1} dx =$$

$$\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + 1)\Gamma(\beta)}{\Gamma(\alpha + \beta + 1)} = \frac{\alpha}{\alpha + \beta}.$$

Аналогично

$$\langle x^2 \rangle = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + 2)\Gamma(\beta)}{\Gamma(\alpha + \beta + 2)} = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)},$$

откуда дисперсия

$$D = \langle x^2 \rangle - \langle x \rangle^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}. \quad (29)$$

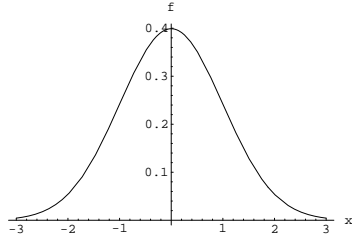
С ростом параметров α и β дисперсия уменьшается.

7.3 Нормальное распределение

Нормальное распределение является наиболее используемым в математической статистике вследствие *центральной предельной теоремы* (см. далее).

Это распределение непрерывной случайной величины на всей вещественной прямой. *Стандартное нормальное распределение* не имеет параметров, математическое ожидание равно нулю, а дисперсия — единице:

$$dp(x) = \frac{dx}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (30)$$



Линейным преобразованием получается нормально распределенная величина с двумя параметрами: μ и σ :

$$dp(x/\mu, \sigma) = \frac{dx}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (31)$$

Практически вся площадь под кривой стандартного распределения лежит в интервале $(-3, +3)$, а для нестандартного распределения в интервале $(\mu - 3\sigma, \mu + 3\sigma)$.

Производящая функция моментов нормального распределения:

$$F(t) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} e^{tx - \frac{(x-\mu)^2}{2\sigma^2}} dx = e^{t\mu + \sigma^2 \frac{t^2}{2}}$$

определяет очень специфическую производящую функцию семиинвариантов нормального распределения:

$$\chi(t) = t\mu + \sigma^2 \frac{t^2}{2}; \quad \chi_1 = \langle x \rangle = \mu; \quad \chi_2 = D_x = \sigma^2. \quad (32)$$

У нормального распределения отличны от нуля только первый и второй семиинварианты.

В частности, стандартно нормально распределенная случайная величина имеет нулевое математическое ожидание и единичную дисперсию, а также нулевые все высшие семиинварианты.

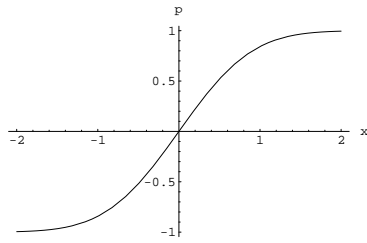
7.4 Интеграл ошибок

Нормально распределенная случайная величина принимает значения от минус до плюс бесконечности. При этом вследствие симметрии распределения вероятность значения меньшего математического ожидания μ равна ровно $1/2$, также как и вероятность значения большего μ : точка $x = \mu$ делит площадь под кривой плотности нормального распределения пополам.

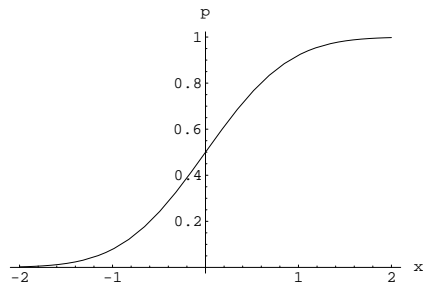
Площадь под кривой стандартного нормального распределения на симметричном участке $(-x, x)$ определяется *интегралом вероятностей* $\Phi(\eta)$ (см. стр. 39), но часто в книгах приводится *интеграл ошибок* – функция $\text{erf}(x)$:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = 2\Phi(\sqrt{2}x). \quad (33)$$

Полная площадь равна $\text{erf}(\infty) = 1$. График функции $\text{erf}(x)$:



Вероятность того, что случайная величина меньше заданного числа a определяется функцией $p(a) = (1 + \operatorname{erf}(a))/2$, принимающей значений 0 на минус бесконечности и 1 на плюс бесконечности:



Исторически первым ввел нормальное распределение Лаплас (1749-1827) в чуть не стандартном виде

$$dp(x) = \frac{dx}{\sqrt{\pi}} e^{-x^2},$$

которое с виду проще, чем стандартное нормальное распределение (30) и потому до сих пор используется довольно часто. Однако в распределении Лапласа дисперсия равна $1/2$, тогда как в стандартном нормальном распределении (30), введенном Гауссом (1777-1855), дисперсия равна единице, что с точки зрения теории вероятностей делает распределение Гаусса более предпочтительным.

7.5 Линейная суперпозиция нормально распределенных величин

Пусть имеется n различных (или одинаковых) нормально распределенных величин x_1, x_2, \dots, x_n с математическими ожиданиями $\mu_1, \mu_2, \dots, \mu_n$ и дисперсиями D_1, D_2, \dots, D_n . Их суперпозиция с постоянными коэффициентами a_i и константой c :

$$z = a_1 x_1 + a_2 x_2 + \dots + a_n x_n + c$$

является также случайной величиной и имеет все семиинварианты выше второго равными нулю, так как они равны нулю у всех составляющих, то есть также является нормально распределенной случайной величиной. Ее математическое ожидание (линейный семиинвариант):

$$\langle z \rangle = a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n + c,$$

а дисперсия (квадратичный семиинвариант):

$$D_z = a_1^2 D_1 + a_2^2 D_2 + \dots + a_n^2 D_n$$

всегда положительна. Распределение этой величины, следовательно,

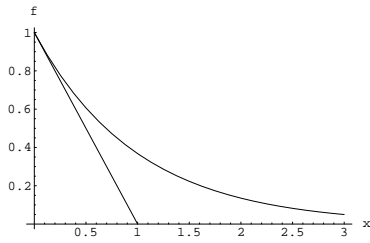
$$dp(z) = \frac{1}{\sqrt{2\pi D_z}} e^{-\frac{(z - \langle z \rangle)^2}{2 D_z}} dz.$$

Это очень важный результат: линейная функция нормально распределенных величин есть нормально распределенная случайная величина.

7.6 Экспоненциальное распределение

Непрерывная случайная величина принимает любые положительные значения. Стандартное и с масштабным параметром k :

$$dp(\xi) = e^{-\xi} d\xi; \quad \xi = kx; \quad x = \xi/k; \quad dp(x/k) = k e^{-kx} dx. \quad (34)$$



Производящая функция моментов

$$F(t) = \int_0^{\infty} e^{-kx+xt} k dx = \frac{1}{1-\frac{t}{k}} = 1 + \frac{t}{k} + \frac{t^2}{k^2} + \dots;$$

$$\langle x \rangle = \frac{1}{k}; \quad \langle x^2 \rangle = \frac{2}{k^2}; \quad D_x = \langle x^2 \rangle - \langle x \rangle^2 = \frac{2}{k^2} - \frac{1}{k^2} = \frac{1}{k^2}.$$

Экспоненциальное распределение нередко имеют интервалы между независимыми событиями.

7.7 Гамма-распределение

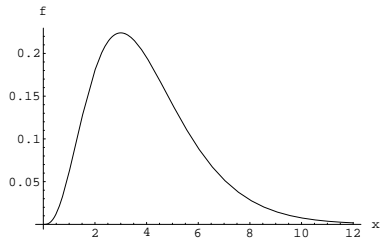
Порождается интегральным представлением Г-функции:

$$\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha-1} dx.$$

Случайная величина принимает любые положительные значения с вероятностью

$$dp(x/\alpha) = \frac{1}{\Gamma(\alpha)} e^{-x} x^{\alpha-1} dx. \quad (35)$$

Приведен график для $\alpha = 4$:



Первые два момента:

$$\begin{aligned} \langle x \rangle &= \frac{1}{\Gamma(\alpha)} \int_0^\infty e^{-x} x^\alpha dx = \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} = \alpha; \\ \langle x^2 \rangle &= \frac{\Gamma(\alpha+2)}{\Gamma(\alpha)} = \alpha(\alpha+1) \end{aligned}$$

определяют дисперсию

$$D_x = \alpha. \quad (36)$$

Не стандартное Γ -распределение получается масштабным преобразованием $x = ky$ с соответствующим изменением математического ожидания и дисперсии:

$$dp(y/k, \alpha) = \frac{k^\alpha}{\Gamma(\alpha)} e^{-ky} y^{\alpha-1} dy; \quad \langle y \rangle = \frac{\alpha}{k}; \quad D_y = \frac{\alpha}{k^2}. \quad (37)$$

7.8 Многомерное нормальное распределение

Совместное распределение n независимых стандартных нормальных величин $\xi_1, \xi_2, \dots, \xi_n$

$$\begin{aligned} dp(\xi_1, \xi_2, \dots, \xi_n) &= dp(\xi_1) dp(\xi_2) \dots dp(\xi_n) = \\ &= \left(\frac{1}{\sqrt{2\pi}} \right)^n e^{-\frac{1}{2}(\xi_1^2 + \xi_2^2 + \dots + \xi_n^2)} d\xi_1 d\xi_2 \dots d\xi_n \end{aligned} \quad (38)$$

линейным преобразованием

$$\xi_i = \sum_{j=1}^n a_{ij} (x_j - \mu_j) \quad (39)$$

приводит к *многомерному нормальному распределению* n случайных величин x_i :

$$\begin{aligned} dp(x_1, x_2, \dots, x_n) &= \frac{\det(a_{ij}) e^{-C}}{(\sqrt{2\pi})^{n/2}} \left(e^{-\frac{1}{2} \sum K_{ij} x_i x_j + \sum b_i x_i} \right) dx_1 dx_2 \dots dx_n = \\ &= \frac{\det(\sqrt{K_{ij}}) e^{-C}}{(\sqrt{2\pi})^{n/2}} \left(e^{-\frac{1}{2} \sum K_{ij} x_i x_j + \sum b_i x_i} \right) dx_1 dx_2 \dots dx_n, \end{aligned} \quad (40)$$

где

$$\begin{aligned} K_{ij} = K_{ji} &= \sum_{s=1}^n a_{is} a_{js}; \quad \det(K_{ij}) = (\det(a_{ij}))^2; \\ b_i &= \sum_{s=1}^n \mu_s K_{si}; \quad C = \frac{1}{2} K_{ij} \mu_i \mu_j. \end{aligned} \quad (41)$$

В распределении (40) два сложных параметра: симметричная $n \times n$ матрица K_{ij} и n -вектор b_i .

Так как математические ожидания всех исходных случайных величин ξ_i равны нулю, то из (39) следует, что математические ожидания

$$\langle x_i \rangle = \mu_i = \sum_{j=1}^n L_{ij} b_j,$$

где L_{ij} – матрица, обратная матрице K_{ij} :

$$\sum_{j=1}^n K_{ij} L_{jk} = \delta_{jk} = \begin{cases} 0 & j \neq k \\ 1 & j = k \end{cases}.$$

Это *ковариационная матрица*, определяющая математические ожидания

$$\langle (x_i - \mu_i)(x_j - \mu_j) \rangle = L_{ij}. \quad (42)$$

В частности, ее диагональные элементы определяют дисперсии случайных величин x_i : $\langle (x_i - \langle x_i \rangle)^2 \rangle = L_{ii}$, а недиагональные – *корреляцию* между различными переменными.

8 Функции случайной величины

Может измениться область определения случайной величины.

8.1 Квадрат нормальной величины

Для нормально распределенной величины x , принимающей значения от минус до плюс бесконечности, величина $y = x^2$ принимает только

положительные значения, в которые дают вклад как положительные, так и отрицательные значения x , что дает в распределении y дополнительный множитель 2:

$$x = \sqrt{y}; \quad dx = \frac{dy}{2\sqrt{y}};$$

$$dp(y) = \frac{2}{\sqrt{2\pi}} e^{-y/2} \frac{dy}{2\sqrt{y}} = \frac{1}{\Gamma(\frac{1}{2})} e^{-y/2} \left(\frac{y}{2}\right)^{1/2-1} d\left(\frac{y}{2}\right).$$

Это – Γ -распределение с $\alpha = 1/2$. Здесь учтено, что $\Gamma(1/2) = \sqrt{\pi}$.

8.2 Распределение Максвелла

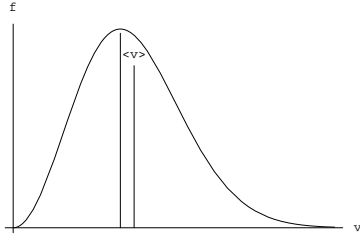
Каждая из трех компонент вектора скорости молекулы имеет нормальное распределение с нулевым математическим ожиданием и дисперсией D , пропорциональной температуре, так что вероятность вектора скорости с компонентами (v_x, v_y, v_z) :

$$dp(v_x, v_y, v_z) = \frac{1}{(\sqrt{2\pi D})^3} e^{-\frac{1}{2D}(v_x^2 + v_y^2 + v_z^2)} dv_x dv_y dv_z$$

Интегрирование ведется по трехмерному объему в пространстве скоростей. Если нас интересует только величина скорости, нужно ввести модуль скорости v . Векторы с одинаковым значением v образуют в пространстве скоростей сферу площадью $4\pi v^2$, а сферический слой скоростей со значениями от v до $v + dv$ имеет объем $4\pi v^2 dv$, так что

$$dp(v) = \frac{4\pi}{(\sqrt{2\pi D})^3} e^{-\frac{v^2}{2D}} v^2 dv = \frac{\sqrt{2\pi}}{D^{3/2}} e^{-\frac{v^2}{2D}} v^2 dv. \quad (43)$$

Это и есть распределение Максвелла



Наиболее вероятная скорость определяется дисперсией $D = kT/m$, где k – постоянная Больцмана, T – абсолютная температура и m – масса молекулы:

$$\frac{d}{dv} \left(-\frac{v^2}{2D} + 2 \ln(v) \right) = -\frac{v}{D} + \frac{2}{v} = 0; \quad v_0 = \sqrt{2D}.$$

Если ввести величину $y = v^2/(2D)$, то распределение для y – это Γ - распределение с параметром $3/2$:

$$dp(y) = \frac{1}{\Gamma(3/2)} e^{-y} \sqrt{y} dy.$$

Скорость $v = \sqrt{2Dy} = v_0 \sqrt{y}$, так что средняя скорость

$$\langle v \rangle = v_0 \langle \sqrt{y} \rangle = v_0 \frac{1}{\Gamma(3/2)} \int_0^\infty e^{-y} y dy = \frac{v_0}{\Gamma(3/2)} = \frac{2}{\sqrt{\pi}} v_0 = 1.13 v_0.$$

Эта скорость также указана на графике и она больше наиболее вероятной.

Математическое ожидание v^2

$$\langle v^2 \rangle = v_0^2 \langle y \rangle = v_0^2 \frac{\Gamma(5/2)}{\Gamma(3/2)} = \frac{3}{2} v_0^2$$

определяет *среднеквадратичную скорость* $v_2 = \sqrt{\langle v^2 \rangle} = \sqrt{1.5} v_0$ и дисперсию модуля скорости

$$D_v = \langle v^2 \rangle - \langle v \rangle^2 = v_0^2 (1.5 - (2/\sqrt{\pi})^2) = 0.227 v_0^2.$$

Оно не равно величине D – дисперсии каждой отдельной компоненты скорости.

Среднеквадратичное отклонение от средней скорости $\langle v \rangle = 1.13 v_0$ равно корню из дисперсии и равно $\Delta v = \sqrt{D_v} = 0.48 v_0$.

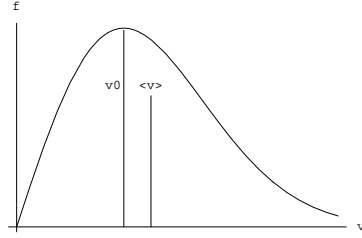
8.3 Распределение Релея

Распределение Релея подобно распределению Максвелла, но не для вектора в трехмерном пространстве, а для вектора на плоскости. Как и у Максвелла каждая компонента распределена нормально с нулевым математическим ожиданием и дисперсией D :

$$dp(v_x, v_y) = \frac{1}{(\sqrt{2\pi D})^2} e^{-\frac{1}{2D}(v_x^2 + v_y^2)} dv_x dv_y.$$

Переходя к модулю, площадь в пространстве скоростей вблизи модуля v равна $2\pi v dv$, так что

$$dp(v) = \frac{1}{D} e^{-\frac{v^2}{2D}} v dv. \quad (44)$$



В окрестности нуля плотность линейна по скорости, в то время как в распределении Максвелла квадратична.

Для величины $y = v^2/(2D)$ это Г-распределение с показателем 1:

$$dp(y) = e^{-y} dy.$$

Наиболее вероятная скорость определяется из условия максимума логарифма плотности вероятности

$$\frac{d}{dv} \left(-\frac{v^2}{2D} + \ln(v) \right) = -\frac{v}{D} + \frac{1}{v} = 0; \quad v_0 = \sqrt{D}.$$

Так как $v = \sqrt{2Dy} = v_0 \sqrt{2y}$ вычисляется средняя скорость

$$\langle v \rangle = \sqrt{2} v_0 \langle \sqrt{y} \rangle = \sqrt{2} v_0 \Gamma\left(\frac{3}{2}\right) = \sqrt{\frac{\pi}{2}} v_0 \approx 1.25 v_0.$$

Средний квадрат скорости

$$\langle v^2 \rangle = 2 v_0^2 \quad \langle y \rangle = 2 v_0^2$$

Определяет среднеквадратичную скорость $v_2 = \sqrt{2} v_0$ и дисперсию модуля скорости (опять не равным D – дисперсии отдельной компоненты):

$$D_v = \langle v^2 \rangle - \langle v \rangle^2 = \left(2 - \frac{\pi}{2}\right) v_0^2 = 0.43 v_0^2 = 0.43 D.$$

Здесь мы называли v вектором скорости, однако теория применима для любого двухкомпонентного вектора и чаще распределение Релея применяется при изучении флуктуаций электрического тока или для волн на воде.

8.4 Распределение χ^2

Распределения Релея и Максвелла фактически являются частным случаем (при $n = 2, 3$) более общего распределения, называемого χ^2 с параметром n – число компонент с нормальным распределением и одинаковой дисперсией D . Так как масштабное преобразование (уменьшение в \sqrt{D} раз) приводит к случайной величине с единичной дисперсией, то достаточно рассмотреть только стандартное χ^2 распределение с n степенями свободы. Правда, в отличие от распределений Релея и Максвелла, в χ^2 - распределении случайной величиной является не модуль n - мерного вектора \mathbf{x} с модулем x , а его квадрат:

$$\chi^2 = x_1^2 + x_2^2 + \dots + x_n^2 = \sum_{i=1}^n x_i^2 = u = r^2.$$

$$dp(x_1, x_2, \dots, x_n) = N_1^{-1} e^{-\frac{1}{2}(x_1^2 + x_2^2 + \dots + x_n^2)} dx_1 dx_2 \dots dx_n.$$

В n -мерном пространстве модуль r (а следовательно и u) постоянен на $(n-1)$ - мерной сфере с $(n-1)$ - мерным объемом $\Omega_{n-1} r^{n-1}$, а сферический слой от r до $r + dr$ имеет n - мерный объем

$$dV_n = \Omega_{n-1} r^{n-1} dr = \Omega_{n-1} u^{(n-1)/2} \frac{du}{2\sqrt{u}},$$

так что распределение по u :

$$dp(\chi^2) = dp(u) = N^{-1} e^{-u/2} u^{n/2-1} du; \quad N^{-1} = N_1^{-1} \Omega_{n-1}/2. \quad (45)$$

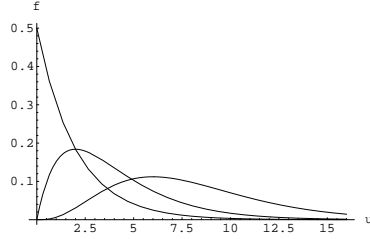
Нормировочный множитель вычислим из условия нормировки вероятности:

$$N = \int_0^\infty e^{-u/2} u^{n/2-1} du = 2^{n/2} \Gamma\left(\frac{n}{2}\right),$$

так что нормированное распределение представлено выражением

$$dp(\chi^2) = dp(u) = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} e^{-u/2} u^{n/2-1} du. \quad (46)$$

На графиках показаны распределения при $n = 2, 4, 8$ (где какое?):



Математическое ожидание

$$\langle u \rangle = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} \int_0^\infty e^{-u/2} u^{n/2} du = \frac{2^{n/2+1} \Gamma(\frac{n}{2} + 1)}{2^{n/2} \Gamma(\frac{n}{2})} = 2 \frac{n}{2} = n.$$

Аналогично, второй момент

$$\langle u^2 \rangle = \frac{2^{n/2+2} \Gamma(\frac{n}{2} + 2)}{2^{n/2} \Gamma(\frac{n}{2})} = 4 \left(\frac{n}{2} + 1 \right) \frac{n}{2} = n(n+2),$$

что определяет дисперсию

$$D_{\chi^2} = n(n+2) - n^2 = 2n.$$

Если исходная дисперсия компонент $D \neq 1$, то

$$\langle u \rangle = nD; \quad D_u = 2nD^2.$$

8.5 Распределение Стьюдента

По Стьюденту (Вильям Госсет, 1876-1937) с m степенями свободы распределена величина

$$S = \frac{x_0}{\sqrt{(x_1^2 + x_2^2 + \dots + x_m^2)/m}},$$

где x_0, x_1, \dots, x_m — нормально распределенные величины с нулевым математическим ожиданием и одинаковой дисперсией. Так как это отношение от дисперсии не зависит, то при вычислении распределения дисперсию можно полагать равной единице:

$$dp(x_0, x_1, \dots, x_m) = \left(\frac{1}{\sqrt{2\pi}} \right)^{m+1} \left(e^{-\frac{1}{2}(x_0^2 + (x_1^2 + \dots + x_m^2))} \right) dx_0 dx_1 \dots dx_m.$$

Обозначим $r \equiv \sqrt{x_1^2 + x_2^2 + \dots + x_m^2}$ и выразим x_0 через S : $x_0 = r S / \sqrt{m}$.

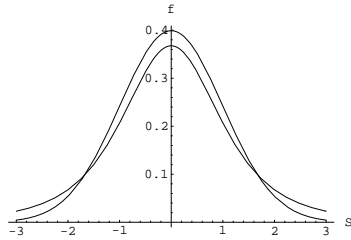
$$dp(x_0, x_1, \dots, x_m) = N^{-1} \left(e^{-\frac{m}{2} \frac{r^2}{m} (1 + S^2/m)} \right) r^m dr d\Omega_{m-1} dS.$$

Интегрирование по x_1, \dots, x_m

Стандартное распределение Стьюдента:

$$dp_m(S) = \frac{1}{\sqrt{\pi m}} \frac{\Gamma(\frac{m+1}{2})}{\Gamma(\frac{m}{2})} \frac{dS}{\left(1 + \frac{S^2}{m}\right)^{\frac{m+1}{2}}} \quad (47)$$

На графике распределение Стьюдента с 3-мя степенями свободы приведено в сравнении со стандартным нормальным распределением:



Какая из этих кривых распределение Стьюдента?

8.6 Распределение выборочных дисперсий

Интеграл

$$\int_0^{\infty} e^{-\frac{r^2}{2}} r^{n-1} dr = 2^{\frac{n}{2}-1} \Gamma\left(\frac{n}{2}\right)$$

определяет распределение модуля суммы квадратов n нормально распределенных случайных величин.

Распределение Фишера – распределение отношения двух средних сумм квадратов нормально распределенных величин

$$s_1^2 = \frac{1}{n_1}(x_1^2 + x_2^2 + \dots + x_{n_1}^2) = \frac{r_1^2}{n_1};$$

$$s_2^2 = \frac{1}{n_2}(x_{n_1+1}^2 + x_{n_1+2}^2 + \dots + x_{n_1+n_2}^2) = \frac{r_2^2}{n_2}.$$

Распределение среднеквадратичных отклонений – величины $F = s_1^2/s_2^2$ не зависит от дисперсии, если она общая в обоих распределениях, поэтому вычисления можно проводить для нормально распределенных величин x_i с единичной дисперсией. Для любой функции этого отношения $f(F)$ математическое ожидание определяется интегралом $\langle f(F) \rangle =$

$$\frac{1}{2^{\frac{n_1+n_2}{2}-2} \Gamma(\frac{n_1}{2}) \Gamma(\frac{n_2}{2})} \int_0^{\infty} dr_1 \int_0^{\infty} dr_2 f(F) e^{-\frac{r_1^2+r_2^2}{2}} r_1^{n_1-1} r_2^{n_2-1}.$$

Обозначим $z = r_1^2/r_2^2 = F n_1/n_2$, выразим $r_1 = \sqrt{z} r_2$; $dr_1/r_1 = dz/(2z)$ (при интегрировании по z величина $r_2 \equiv r$ остается постоянной) и подставим в интеграл

$$N^{-1} \left(\int_0^{\infty} \frac{dr}{r} e^{-\frac{(1+z)r^2}{2}} r^{n_1+n_2} z^{\frac{n_1}{2}} \right) \frac{dz}{2z} = \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2}) \Gamma(\frac{n_2}{2})} \frac{z^{\frac{n_1}{2}}}{(1+z)^{\frac{n_1+n_2}{2}}} \frac{dz}{z}.$$

Возвращаясь теперь от переменной z к переменной $F = z n_2/n_1$, получаем распределение этой величины:

$$dp(F) = n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}} \frac{\Gamma(\frac{n_1+n_2}{2})}{\Gamma(\frac{n_1}{2}) \Gamma(\frac{n_2}{2})} \frac{F^{\frac{n_1}{2}}}{(n_1 + n_2 F)^{\frac{n_1+n_2}{2}}} \frac{dF}{F}. \quad (48)$$

Применяется для сравнения выборочных дисперсий.

9 Пределные теоремы

9.1 Первая предельная теорема

Одна из основных областей приложения теории вероятностей – многократно повторяющиеся события с возможными различными исходами.

В схеме Бернулли – проведение n испытаний с двумя исходами: “+” с вероятностью p и “-” с вероятностью $1 - p$ – случайная величина m (число выпавших плюсов) имеет математическое ожидание $\langle m \rangle = np$ и дисперсию $D_m = np(1 - p)$.

Среднее число плюсов в n испытаниях $x = m/n$ имеет математическое ожидание $\langle x \rangle = p$ и дисперсию

$$D_x = \frac{D_m}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

При увеличении числа испытаний дисперсия уменьшается и в пределе $n \rightarrow \infty$ оказывается равной нулю. В этом и состоит *первая предельная теорема* теории вероятностей: среднее значение $x = m/n$ по вероятности сходится к p – вероятности появления события в одном испытании.

9.2 Теорема Муавра-Лапласа

В той же схеме Бернулли можно построить случайную величину с нулевым математическим ожиданием и единичной дисперсией при любом n :

$$\eta = \frac{m - np}{\sqrt{np(1-p)}}; \quad \langle \eta \rangle = 0; \quad D_\eta = \frac{D_m}{(\sqrt{np(1-p)})^2} = 1.$$

Но все высшие ($k > 2$) семинварианты этой величины стремятся к нулю при $n \rightarrow \infty$:

$$\chi_k^{[\eta]} = \frac{n}{n^{k/2}} \frac{\chi_k^{[+]}}{(p(1-p))^{k/2}} \xrightarrow{n \rightarrow \infty} 0.$$

Здесь $\chi_k^{[+]}$ – семинварианты выпадения плюса в одном бинарном эксперименте – конечные величины, не зависящие от n .

Набор семиинвариантов равных нулю, за исключением второго, равного единице, имеет стандартное нормальное распределение.

При минимальном значении $m = 0$ значение $x_1 = \sqrt{n} \sqrt{p/(1-p)}$ стремится к минус бесконечности, а при максимальном $m = n$ значение $x_2 = \sqrt{n} \sqrt{(1-p)/p}$ стремится к плюс бесконечности, при этом интервалы между соседними значениями x_m , равные $1/\sqrt{np(1-p)}$ стремятся к нулю, то есть в пределе $n \rightarrow \infty$ величина x становится непрерывной, распределенной по стандартному нормальному закону.

Теорема Муавра-Лапласа позволяет значительно упростить решение задач, связанных с биномиальным распределением при больших n (см. далее).

9.3 Центральная предельная теорема

Теорема Муавра-Лапласа является частным случаем более общей *центральной предельной теоремы*. Пусть имеется случайная величина x (дискретная или непрерывная) с математическим ожиданием μ , дисперсией D и конечными высшими семиинвариантами. При n -кратной реализации этой случайной величины (x_1, x_2, \dots, x_n) можно построить

$$\eta = \frac{\sum_{i=1}^n x_i - n \langle x \rangle}{\sqrt{n D_x}}; \quad \langle \eta \rangle = 0; \quad D_\eta = \frac{n D_x}{(\sqrt{n D_x})^2} = 1. \quad (49)$$

Семиинварианты случайной величины η определяются семиинвариантами случайной величины x :

$$\xi_l^{[\eta]} = n \xi_l^{[x]} \left(\frac{1}{n D_x} \right)^{l/2} = \frac{1}{n^{(l-2)/2}} \frac{\xi_l^{[x]}}{D_x^{l/2}}. \quad (50)$$

При стремлении n к бесконечности при $l > 2$ все они стремятся к нулю, величина η становится непрерывной на всей вещественной оси и имеет стандартное нормальное распределение.

9.4 Использование ЦПТ

Пусть мы хотим определить вероятность того, что число m выпадения орла находится в заданном интервале $m_1 < m < m_2$ $p = 0.5$. Применим теорему Муавра-Лапласа.

$$\eta = \frac{m - 0.5n}{0.5\sqrt{n}} = \frac{2m - n}{\sqrt{n}}.$$

$$\eta_1 = \frac{2m_1 - n}{\sqrt{n}}; \quad \eta_2 = \frac{2m_2 - n}{\sqrt{n}}.$$

Пусть, например, $n = 100$, $m_1 = 45$, $m_2 = 55$. Тогда $\eta_1 = -1$; $\eta_2 = 1$ и вероятность того, что при 100 бросаниях $45 < m < 55$ равна площади под кривой нормального распределения от -1 до +1. Она равна интегралу вероятностей $\Phi(\eta_2) - \Phi(\eta_1) = \Phi(1) - \Phi(-1) = 0.68$ (см. таблицу).

Если же $n = 1000$, $m_1 = 4500$, $m_2 = 5500$ — отношение такое же, как и в предыдущем случае, но $\eta_1 = -10$, $\eta_2 = 10$, то вероятность попадания m в этот интервал равна площади от -10 до +10 — почти равна единице с огромной точностью.

Значения интеграла вероятностей

$$\Phi(\eta) = \frac{1}{\sqrt{2\pi}} \int_0^\eta e^{-x^2/2} dx \quad (51)$$

приведены в таблице:

η	$\Phi(\eta)$	η	$\Phi(\eta)$	η	$\Phi(\eta)$
0.0	0.00000	1.0	34134	2.0	47725
0.1	03983	1.1	36433	2.2	48610
0.2	07926	1.2	38493	2.4	49180
0.3	11791	1.3	40320	2.6	49534
0.4	15542	1.4	41924	2.8	49744
0.5	19146	1.5	43319	3.0	49865
0.6	22575	1.6	44520	3.5	4997674
0.7	25804	1.7	45543	4.0	4999683
0.8	28814	1.8	46407	4.5	4999966
0.9	31594	1.9	47128	5.0	4999997

Интеграл вероятностей – нечетная функция: $\Phi(-z) = -\Phi(z)$.

Часть II

Математическая статистика

Я бросаю монету и сразу накрываю ладонью, успев увидеть что на ней выпала “решка”. Однако больше никто не знает результата. Какова для них вероятность того, что под рукой “решка”? $1/2$.

Вероятностно описываются не только события, но и наши знания о них. Какова вероятность выпадения “орла” на гнутой монете. Это вполне определенное число между нулем и единицей, но мы его не знаем и можем делать лишь *вероятностные суждения*.

10 Основная задача математической статистики

Основной задачей теории вероятностей является расчет по заданным вероятностям некоторых (простых) событий вероятностей сложных событий, по определенным правилам, связанным с простыми. Сами вероятности при этом задаются определенными числами и функциями.

Качественное отличие *математической статистики* состоит в том, что эти функции и числа неизвестны заранее, сами являются случайными, и их вероятностные распределения определяются на основе эксперимента.

11 Теория Байеса

Томас Байес (1702-1761) построил теорию развития наших вероятностных знаний о мире на основе эксперимента, используя *статистическую зависимость* между параметрами распределения и результатами эксперимента.

11.1 Статистическая зависимость

Распределение двух случайных величин x и h задается *совместным распределением вероятностей*

$$dp(x, h) = \rho(x, h) dx dh; \quad \int \rho(x, h) dx dh = 1. \quad (52)$$

Для статистически независимых величин $dp(x, h) = dp(x) dp(h) = (\rho_x(x)dx)(\rho_h(h)dh)$. Если это не так, то случайные величины x и h являются *статистически зависимыми*.

Совместное распределение определяет *частные вероятности* x и h :

$$dp(x) = \int_h dp(x, h) = dx \int \rho(x, h) dh;$$

$$dp(h) = \int_x dp(x, h) = dh \int \rho(x, h) dx.$$

Основным инструментом в теории Байеса является *условные вероятности* $\rho(x/h) dx$ и $(\rho(h/x) dh)$, симметрично определяемые для x и h :

$$dp(x, h) = \rho(x, h) dx dh = dp(x/h) \cdot dp(h) = dp(h/x) \cdot dp(x). \quad (53)$$

Отсюда

$$dp(x/h) = \frac{dp(x, h)}{dp(h)}; \quad dp(h/x) = \frac{dp(x, h)}{dp(x)}.$$

11.2 Теорема Байеса

Из (53) можно выразить одну условную вероятность через другую:

$$dp(h/x) = dp(x/h) \frac{dp(h)}{dp(x)}. \quad (54)$$

Это соотношение называется *теоремой Байеса* и является основной формулой статистики Байеса.

11.3 Методика Байеса

Припишем теперь величинам x и h определенный смысл: будем полагать x измеряемой случайной величиной с параметром распределения h , также являющимся случайной величиной вследствие нашего незнания. Пусть до измерения наши знания о параметре h определялись распределением $dp_{apr}(h)$ – *априорным* распределением, – а после проведения опыта, давшего определенное значение $x = x_e$, наши знания о параметре h изменились: после получения результата мы имеем *апостериорное* распределение

$$dp_{post}(h/x_e) = dp(x_e/h) \frac{dp_{apr}(h)}{dp(x_e)} = N^{-1} dp(x_e/h) dp_{apr}(h). \quad (55)$$

В этой формуле x_e – уже не случайная, а определенная, замеренная в эксперименте величина. Входящая в формулу $dp(x_e)$, также является константой и может быть определена после вычисления из условия нормировки вероятности. Случайной величиной является параметр h , и формула Байеса (55) определяет изменение наших знаний об этой случайной величине в результате эксперимента. В случае дискретности x в формуле (55) $dp(x/h)$ заменяется на $p(x/h)$:

$$dp_{post}(h/x_e) = N^{-1} p(x_e/h) dp_{apr}(h). \quad (56)$$

12 Параметр бинарного распределения

Применим методику Байеса к бинарной системе, где h есть вероятность выпадения “+” и соответственно $1 - h$ – вероятность “-”. Мы еще не проводили эксперимент. Знаем ли мы что-либо о параметре h ? Мы знаем, что он может принимать значения от 0 до 1, и если больше нет никакой априорной информации, должны придерживаться *критерия максимального незнания*. В данном случае это предположение о равновероятном распределении этого параметра: $dp_0(h) = dh$; $0 < h < 1$.

При заданном h вероятность выпадения плюса $p(+/h) = h$, а минуса $1 - h$.

Если после проведения эксперимента выпал “+”, вероятность преобразуется по теореме Байеса (56)

$$dp(h/+) = 2h dh \equiv dp(h/0,1).$$

При выпадении минуса

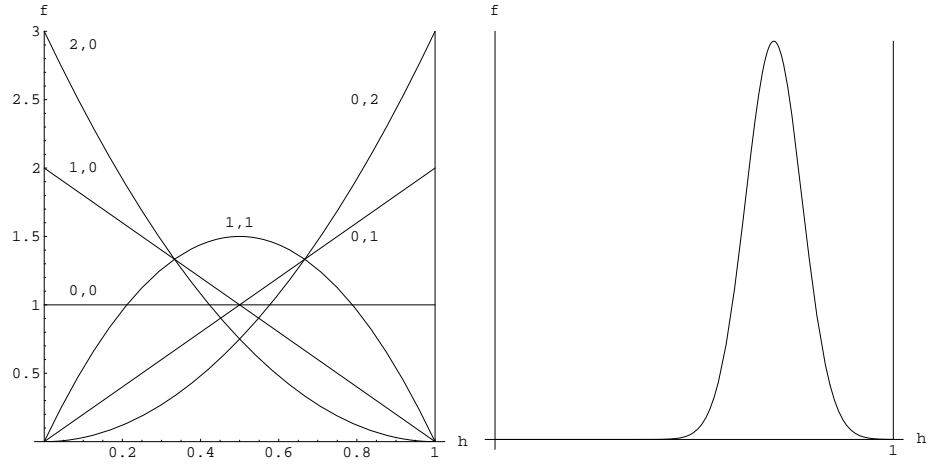
$$dp(h/-) = 2(1-h) dh \equiv dp(h/1,0).$$

В последних выражениях результат эксперимента представляется парой целых чисел (n_-, n_+) с тем, чтобы таким же образом записывать результаты с большим числом испытаний.

Для последующих экспериментов эта вероятность становится априорной. Так, например,

$$dp(h/0,2) = N^{-1} h dp(h/0,1) = 3h^2 dh; \quad dp(h/2,0) = 3(1-h)^2 dh.$$

На левом графике представлены априорная вероятность с меткой $(0,0)$ – горизонтальная прямая $h = 1$ и апостериорные вероятности при одном испытании (исходы $(1,0)$ и $(0,1)$) и двух испытаниях $((n_-, n_+) = (2,0), (1,1), (0,2))$.



При n испытаниях результат $(n-m, m)$, где m – число реализовавшихся плюсов, при заданном h определяется формулой биноми-

ального распределения

$$P((n-m, m)/h) = C h^m (1-h)^{n-m}.$$

По формуле Байеса (56) после реализации $(n-m, m)$ распределение вероятности параметра h определяется β -распределением:

$$dp(h/(n-m, m)) = \frac{(n+1)!}{m!(n-m)!} h^m (1-h)^{n-m}. \quad (57)$$

При некоторых достаточно больших значениях n и m график апостериорного распределения представлен на правом графике.

Таким образом параметр h , определяющий результат статистического испытания, сам стал случайной величиной с распределением (57). Наиболее вероятное значение

$$\frac{d}{dh} \ln(\rho(h)) = \frac{m}{h} - \frac{n-m}{1-h} = 0; \quad h_0 = \frac{m}{n}.$$

Математическое ожидание h

$$\langle h \rangle = \int_0^1 h dp(h) = \frac{m+1}{n+2}; \quad \langle 1-h \rangle = \frac{n-m+1}{n+2}.$$

Дисперсия:

$$\langle h^2 \rangle = \frac{(m+1)(m+2)}{(n+2)(n+3)}; \quad D_h = \langle h^2 \rangle - \langle h \rangle^2 = \frac{(m+1)(n-m+1)}{(n+2)^2(n+3)}$$

уменьшается с ростом n — чем больше объем эксперимента, тем точнее локализуется параметр.

12.1 Априорная вероятность

При определении начальной априорной вероятности мы исходили из принципа максимального незнания о параметре. Однако уже до опыта — или из литературных данных по опытам других исследователей или из предыдущего опыта, — приступая к новой серии опытов, мы кое-что уже знаем о распределении. Как это априорное знание описать математически?

Если мы полагаем, что, например, $h = 0.5 \pm 0.2$, это можно представить априорным распределением (57) с некоторыми (n_0, m_0) :

$$m_0 \approx 0.5n_0; \quad D \approx 0.2^2 = \frac{(m_0 + 1)^2}{(2m_0 + 2)^2 (2m_0 + 3)} = \frac{1}{4(n + 3)},$$

откуда $n \approx 4$, $m \approx 2$, что определяет априорную вероятность

$$dp_{apr}(h) = \frac{5!}{2!2!} h^2 (1 - h)^2 = 30 h^2 (1 - h)^2.$$

Как влияет неточность в задании априорной вероятности? Математическое ожидание h при экспериментальном результате $(n - m, m)$ и априорных параметрах n_0, m_0 равно

$$\langle h \rangle = \frac{m_0 + m + 1}{n_0 + n + 2} \rightarrow \frac{m}{n}$$

при больших n и m и не зависит от n_0 и m_0 , то есть *большие объемы испытаний корректируют неточную априорную информацию*.

12.2 Прогноз

Если о параметре h изначально не было известно ничего, но мы провели n испытаний, в которых реализовалось m плюсов, так что наше апостериорное знание о параметре h определяется формулой (57). Теперь мы хотим рассчитать вероятности появления l плюсов в последующих, еще не проведенных k испытаниях. Если бы h была величина заданная, то

$$p(l/k, h) = \frac{k!}{l!(k-l)!} h^l (1-h)^{k-l}.$$

Теперь по формуле Байеса (53) можно вычислить *совместное распределение* случайных величин l и h , а затем, проинтегрировав по h , получим *частную* вероятность числа l :

$$dp(l, h/k) = \frac{k!}{l!(k-l)!} \frac{(n+1)!}{m!(n-m)!} h^{m+l} (1-h)^{n+k-l-m} dh;$$

$$p(l/k, n, m) = \int_h dp(l, h/k) dp(h/n, m) = \frac{k! (n+1)! (m+l)! (n+k-m-l)!}{l! (k-l)! m! (n-m)! (n+k+1)!}.$$

Это – распределение случайной величины l . Работать с ним – например, вычислить математическое ожидание или дисперсию случайного числа l – непросто. Однако существуют общие теоремы о прогнозе, упрощающие вычисление этих параметров распределения.

- Математическое ожидание случайной величины при вероятностном параметре, также являющимся случайной величиной, равно математическому ожиданию по параметру от условного математического ожидания $\langle x/h \rangle = \int_x x dp(x/h)$, усредненному по распределению случайного параметра:

$$\begin{aligned} dp(x) &= \int_h dp(x/h) dp(h); \\ \langle x \rangle &= \int_x x dp(x) = \int_h \left(\int_x x dp(x/h) \right) dp(h) = \\ &= \int_h \langle x/h \rangle dp(h) = \langle \langle x/h \rangle \rangle_h \equiv \langle \langle x \rangle \rangle. \end{aligned}$$

- Дисперсия случайной величины при вероятностном параметре, являющимся случайной величиной, имеет более сложную конструкцию:

$$\begin{aligned} D_x &= \int_x (x - \langle \langle x \rangle \rangle)^2 dp(x) = \\ &= \int_h \int_x ((x - \langle x/h \rangle) + (\langle x/h \rangle - \langle \langle x \rangle \rangle))^2 dp(x/h) dp(h) = \\ &= \int_h \left(\int_x ((x - \langle x/h \rangle) dp(x/h))^2 dp(h) + \right. \\ &\quad \left. \int_h (\langle x/h \rangle - \langle \langle x \rangle \rangle)^2 dp(h) = \langle D_{x/h} \rangle + D_{\langle x/h \rangle}. \quad (58) \end{aligned}$$

Она равна условной дисперсии, усредненной по параметру, плюс дисперсия по параметру математического ожидания.

Например, в рассмотренном выше случае прогноза на k испытаний после n испытаний, в которых выпало m плюсов для случайной величины l , математическое ожидание

$$\langle\langle l \rangle\rangle = \langle kh \rangle = \frac{k(m+1)}{(n+2)}.$$

Дисперсия в соответствии с (58) ведет себя сложнее:

$$\begin{aligned} D_l &= \langle kh(1-h) \rangle + (\langle (kh)^2 \rangle - \langle kh \rangle^2) = \\ &= \frac{k(m+1)(n-m+1)}{(n+2)(n+3)} + k^2 \frac{(m+1)(n-m+1)}{(n+2)^2(n+3)} = \\ &= \left(k + \frac{k^2}{n+2}\right) \frac{(m+1)(n-m+1)}{(n+2)(n+3)}. \end{aligned}$$

Пока объем прогноза много меньше объема проделанного эксперимента $k \ll n$, дисперсия прогноза растет линейно по k . При объеме прогноза, превышающем объем эксперимента, зависимость от k оказывается квадратичной.

Проследим за относительной погрешностью $\delta = \sigma_l / \langle\langle l \rangle\rangle$, где $\sigma_l = \sqrt{D_l}$ — норма:

$$\delta = \sqrt{\left(\frac{1}{k} + \frac{1}{n+2}\right)} \sqrt{\frac{(m+1)(n-m+1)}{(n+2)(n+3)} \frac{(n+2)}{(m+1)}}.$$

Здесь n и m — объем и результат проведенного эксперимента, а k — объем прогноза. При $k \rightarrow \infty$ относительная ошибка имеет конечный предел в отличие от похожей ситуации в законе больших чисел теории вероятностей.

13 Параметры нормального распределения

Вследствие центральной предельной теоремы многие непрерывные случайные величины распределены нормально.

$$dp(x/\mu, D) = \frac{1}{\sqrt{2\pi D}} e^{-\frac{(x-\mu)^2}{2D}}.$$

Вероятность появления набора случайных величин (x_1, x_2, \dots, x_n) пропорциональна

$$\rho(x_1, x_2, \dots, x_n / \mu, D) = \frac{1}{(2\pi D)^{n/2}} e^{-\frac{(\mu - \bar{x})^2}{2D} - \frac{\Delta^2}{2D}},$$

где показатель экспоненты есть результат раскрытия суммы

$$\begin{aligned} \sum_{i=1}^n \frac{(x_i - \mu)^2}{2D} &= \frac{1}{2D} \sum_{i=1}^n (x_i - \bar{x})^2 + \frac{n}{2D} (\mu - \bar{x})^2 = \\ &= \frac{n(\mu - \bar{x})^2}{2D} + \frac{\Delta^2}{2D}; \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad \sum_{i=1}^n (x_i - \bar{x}) = 0; \quad \Delta^2 = \sum_{i=1}^n (x_i - \bar{x})^2. \end{aligned}$$

Теперь по формуле Байеса

$$\begin{aligned} dp(\mu, D / x_1, x_2, \dots, x_n) &= dp(\mu, D / \bar{x}, \Delta^2) = \\ &= N^{-1} \frac{e^{-\frac{\Delta^2}{2D}}}{D^{n/2}} e^{-\frac{n(\mu - \bar{x})^2}{2D}} dp_0(\mu, D). \end{aligned} \quad (59)$$

Здесь случайными величинами являются μ и D , а экспериментальные данные входят только в виде двух комбинаций: среднего значения \bar{x} и суммарной квадратичной ошибки Δ^2 . Множитель $dp_0(\mu, D)$ есть априорная вероятность распределения μ и D . Максимальная неопределенность для μ – это равновероятность всех μ : $dp_0(\mu) \sim d\mu$. Для D это не так, так как D – положительно определенная величина: $dp_0(D) \sim dD/D$. После такой подстановки распределение (59) принимает вид:

$$dp(\mu, D / \bar{x}, \Delta^2) = N^{-1} \frac{e^{-\frac{\Delta^2}{2D}}}{D^{n/2+1}} e^{-\frac{n(\mu - \bar{x})^2}{2D}} d\mu dD. \quad (60)$$

Нормировочный множитель N находится из условия нормировки вероятности

$$\int dp(\mu, D / \bar{x}, \Delta^2) = 1.$$

13.1 Распределение дисперсии

Распределение (60) несложно проинтегрировать по μ , так как это гауссов интеграл:

$$\int_{-\infty}^{\infty} e^{-\frac{n(\mu-\bar{x})^2}{2D}} d\mu = \sqrt{\frac{2\pi D}{n}},$$

после чего останется распределение дисперсии D

$$dp(D/\Delta^2) = N^{-1} \sqrt{\frac{2\pi}{n}} \frac{e^{-\frac{\Delta^2}{2D}}}{D^{(n-1)/2+1}} dD.$$

Это есть Γ -распределение для величины $u = \frac{\Delta^2}{2D}$. Действительно, представив $D = \frac{\Delta^2}{2u}$, откуда, так как Δ^2 константа,

$$\begin{aligned} dD &= -\frac{\Delta^2}{2u^2} du; \\ dp(u) &= N^{-1} \sqrt{\frac{2\pi}{n}} \left(\frac{2}{\Delta^2}\right)^{\frac{n-1}{2}} u^{(n-1)/2-1} e^{-u} du = \\ &= \frac{1}{\Gamma(\frac{n-1}{2})} u^{(n-1)/2-1} e^{-u} du. \end{aligned}$$

Отсюда определяется нормировочный множитель в (60):

$$N = \sqrt{\frac{2\pi}{n}} \left(\frac{2}{\Delta^2}\right)^{\frac{n-1}{2}} \Gamma\left(\frac{n-1}{2}\right). \quad (61)$$

Математическое ожидание дисперсии

$$\langle D \rangle = \frac{\Delta^2}{2} \left\langle \frac{1}{u} \right\rangle = \frac{\Delta^2}{2} \frac{1}{\Gamma(\frac{n-1}{2})} \int_0^{\infty} u^{(n-3)/2-1} e^{-u} du = \quad (62)$$

$$\frac{\Delta^2}{2} \frac{\Gamma(\frac{n-3}{2})}{\Gamma(\frac{n-1}{2})} = \frac{\Delta^2}{n-3}.$$

Для оценки математического ожидания дисперсии нужно более трех замеров.

13.2 Распределение математического ожидания

Подставив в выражение (60) найденное значение нормировочного множителя (61)

$$dp(\mu, D/\bar{x}, \Delta^2) = \sqrt{\frac{n}{2\pi}} \left(\frac{\Delta^2}{2}\right)^{\frac{n-1}{2}} \frac{1}{\Gamma(\frac{n-1}{2})} \frac{e^{-\frac{\Delta^2}{2D}}}{D^{(n/2+1)}} e^{-\frac{n(\mu-\bar{x})^2}{2D}} d\mu dD, \quad (63)$$

собрав показатель экспоненты в одну функцию w , получаем статистически независимое распределение для случайных величин μ и w :

$$dp(\mu, w) = \left(\sqrt{\frac{n}{\pi \Delta^2}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \frac{d\mu}{\left(1 + \frac{n(\mu-\bar{x})^2}{\Delta^2}\right)^{\frac{n}{2}}} \right) \cdot \left(\frac{e^{-w}}{\Gamma(\frac{n}{2})} w^{\frac{n}{2}-1} dw \right) =$$

$$dp(\mu) \cdot dp(w); \quad w = \frac{1}{2D} (\Delta^2 + n(\mu - \bar{x})^2).$$

Первый сомножитель в этом распределении определяет *распределение Стьюдента* с $m = n - 1$ степенями свободы: формула (47) на странице 35:

$$dp(\mu/\bar{x}, \Delta^2) = \sqrt{\frac{n}{\pi \Delta^2}} \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{n-1}{2})} \frac{d\mu}{\left(1 + \frac{n(\mu-\bar{x})^2}{\Delta^2}\right)^{\frac{n}{2}}}. \quad (64)$$

Из соотношения

$$\mu = \bar{x} + y_{n-1} \sqrt{\frac{\Delta^2}{n(n-1)}}; \quad m = n - 1$$

Можно получить математическое ожидание и дисперсию оцениваемой величины μ :

$$\langle \mu \rangle = \bar{x}; \quad D_\mu = \frac{\Delta^2}{n(n-1)} \frac{n+1}{n-1} = \left(1 + \frac{1}{n}\right) \frac{\Delta^2}{(n-1)^2},$$

а также определить *доверительный интервал*

$$\mu = \bar{x} \pm y_{n-1}(\eta) \sqrt{\frac{\Delta^2}{n(n-1)}},$$

где $y_{n-1}(\eta)$ – *квантиль* надежности η распределения Стюдента с $n - 1$ степенями свободы: $\pm y_{n-1}(\eta)$ отсекают под кривой площадь η (например, 0.9, 0.95), определяющую *надежность вывода*. Квантили заданной надежности распределения Стюдента с m степенями свободы приводятся в таблицах (см., например, [2]).

13.3 Прогноз

При заданных μ и D распределение случайной величины x определяется нормальным распределением

$$dp(x/\mu, D) = \frac{1}{\sqrt{2\pi D}} e^{-\frac{(x-\mu)^2}{2D}} dx.$$

Но сами эти параметры оказались случайными величинами, определяемыми распределением (63), так что с учетом этой вероятности прогнозируемое значение x имеет распределение

$$dp(x) = \int_{\mu, D} dp(x/\mu, D) dp(\mu, D) = N^{-1} dx \int d\mu dD \left(e^{-\frac{f}{2D}} \right) \frac{1}{D^{\frac{n+1}{2}+1}}, \quad (65)$$

где

$$f = \Delta^2 + n(\mu - \bar{x})^2 + (x - \mu)^2 = \Delta^2 + (\mu - \bar{\mu})^2 + \frac{n(x - \bar{x})^2}{n+1};$$

$$\bar{\mu} = \frac{n\bar{x} + x}{n+1}.$$

После интегрирования по μ и D , уменьшающего степень D в знаменателе на $3/2$, распределение прогнозируемой величины x оказывается распределением Стюдента с $m = n - 1$ степенью свободы:

$$dp(x) = N^{-1} \frac{dx}{\left(1 + \frac{n(x-\bar{x})^2}{(n+1)\Delta^2}\right)^{\frac{n}{2}}} \sim \frac{dS}{\left(1 + \frac{S^2}{m}\right)^{\frac{m+1}{2}}}.$$

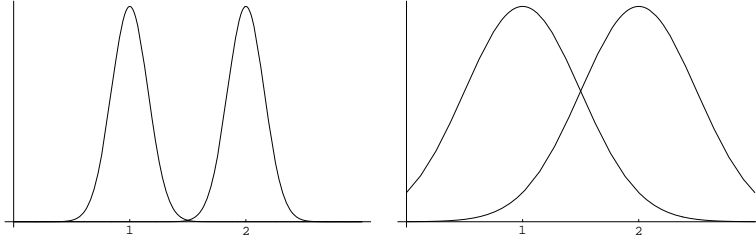
$$x = \bar{x} + S_{n-1} \sqrt{\frac{(n+1)\Delta^2}{n(n-1)}}; \quad \langle x \rangle = \bar{x}; \quad D_x = \frac{(n+1)\Delta^2}{n(n-3)}.$$

В отличие от дисперсии параметра μ , уменьшающейся при увеличении числа замеров, так как Δ^2 растет пропорционально n , дисперсия x с ростом числа замеров стабилизируется. Наши знания о параметрах растут, но объективная неопределенность в системе не уменьшается.

13.4 Различение математических ожиданий

Если в двух различных сериях измерений получены слегка различные значения \bar{x} , являются ли эти различия результатом замера одной и той же величины, или же в каждой серии имелась специфика, приведшая к измерению разных величин?

На рисунках видно, что вывод определяется не разностью, а отношением этой разности к погрешности измерений:



Для установления математического критерия различности средних нужно найти методом Байеса распределение параметров μ_1 и μ_2 в этих сериях с длинами выборок n_1 и n_2 :

$$dp(\mu_1, \mu_2, \sigma) = N^{-1} e^{-\frac{\Delta_1^2}{2\sigma^2} - \frac{\Delta_2^2}{2\sigma^2}} \left(e^{-\frac{n_1(\mu_1 - \bar{x}_1)^2 + n_2(\mu_2 - \bar{x}_2)^2}{2\sigma^2}} \right) \frac{d\mu_1 d\mu_2 d\sigma}{\sigma^{n_1+n_2+1}}. \quad (66)$$

Основную роль играет выражение в показателе экспоненты

$$s = n_1 (\mu_1 - \bar{x}_1)^2 + n_2 (\mu_2 - \bar{x}_2)^2.$$

Выделим из переменных μ_1 и μ_2 их разность

$$\mu_1 = \mu + \frac{n_2}{n_1 + n_2} \lambda; \quad \mu_2 = \mu - \frac{n_1}{n_1 + n_2} \lambda; \quad \lambda = \mu_1 - \mu_2.$$

Тогда

$$s = (n_1 + n_2)(\mu - \bar{\bar{x}})^2 + \frac{n_1 n_2}{n_1 + n_2} (\lambda - \delta)^2,$$

где

$$\bar{\bar{x}} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2}{n_1 + n_2}; \quad \delta = \bar{x}_1 - \bar{x}_2.$$

После интегрирования по μ зависимость от λ определится студентовской функцией с $m = n_1 + n_2 - 2$ степенями свободы

$$\left(1 + \frac{n_1 n_2 (\lambda - \delta)^2}{(n_1 + n_2) (\Delta_1^2 + \Delta_2^2)}\right)^{-(n_1 + n_2 - 1)/2} = \left(1 + \frac{t^2}{m}\right)^{-(m+1)/2},$$

откуда можно определить

$$\langle \mu_1 - \mu_2 \rangle = \langle \lambda \rangle = \bar{x}_1 - \bar{x}_2; \quad D_\lambda = \frac{(n_1 + n_2 - 2)(n_1 + n_2)}{n_1 n_2} \frac{\Delta_1^2 + \Delta_2^2}{(n_1 + n_2 - 4)}.$$

Значимость разности $\bar{x}_1 - \bar{x}_2$ определяется ее отношением к корню из дисперсии.

14 Регрессия

Определению подлежит зависимость

$$y(x) = a_1 u_1(x) + a_2 u_2(x) + \dots + a_m u_m(x) + \delta, \quad (67)$$

где $u_1(x), u_2(x), \dots, u_m(x)$ – заданные *регрессионные функции*, коэффициенты a_1, a_2, \dots, a_m – подлежащие определению *регрессионные коэффициенты*, выступающие с точки зрения математической статистики как случайные величины, для которых необходимо найти распределение вероятностей, математические ожидания, дисперсии. Величина δ является погрешностью и полагается нормально распределенной с нулевым математическим ожиданием и неизвестной дисперсией.

Схема данных: имеется N пар данных (x_i, y_i) ; $i = 1, \dots, N$.

$$\delta_i = y_i - \sum_{j=1}^m a_j u_j(x_i); \quad j = 1, \dots, N.$$

При заданных регрессионных коэффициентах a_j и дисперсии D вероятность появления набора ошибок δ_i определяется произведением гауссовых распределений

$$dp(\delta_1, \dots, \delta_N / D, a_1, \dots, a_m) \sim \prod_{i=1}^N \left(\frac{e^{-\frac{\delta_i^2}{2D}}}{\sqrt{D}} \right) \sim \frac{1}{D^{\frac{N}{2}}} e^{-\frac{f}{2D}}, \quad (68)$$

где

$$f = \sum_{i=1}^N \delta_i^2 = \sum_{i=1}^N \left(y_i - \sum_{j=1}^m a_j u_j(x_i) \right)^2. \quad (69)$$

Применяя метод Байеса, нужно рассматривать коэффициенты a_i , D как случайные величины с априорной мерой вероятности

$$dp_{apr}(a_1, a_2, \dots, a_n, D) \sim da_1 da_2 \dots da_n \frac{dD}{D},$$

что приводит к распределению параметров a_j и D :

$$dp(a_1, a_2, \dots, a_n, D) = N^{-1} \frac{1}{D^{\frac{N}{2}}} \left(e^{-\frac{f}{2D}} \right) da_1 da_2 \dots da_n \frac{dD}{D}, \quad (70)$$

где функция f является квадратичной по коэффициентам регрессии a_j :

$$f = \left(\sum_{i=1}^N y_i^2 \right) - 2 \sum_{j=1}^m a_j \left(\sum_{i=1}^N y_i u_j(x_i) \right) + \sum_{j=1}^m \sum_{k=1}^m a_j a_k \left(\sum_{i=1}^N u_j(x_i) u_k(x_i) \right). \quad (71)$$

Обозначая выражения в круглых скобках как $\langle y^2 \rangle$, $\langle y u_j \rangle \equiv b_j$ и $\langle u_j u_k \rangle \equiv K_{jk}$, сразу определим, что наиболее вероятные значения статистических коэффициентов регрессии находятся в минимуме функции f , определяемом линейной системой уравнений

$$b_j = \sum_{l=1}^m K_{jl} \bar{a}_l; \quad \bar{a}_l = \sum_{j=1}^m K_{lj}^{-1} b_j. \quad (72)$$

Этот вывод является основным в теории наименьших квадратов, где основной формулой является (72).

14.1 Регрессия прямой линией

Полностью статистическую задачу распределения коэффициентов регрессии мы рассмотрим на примере определения линейной зависимости $y = a + b x$ с двумя регрессионными функциями $u_1 = 1$, $u_2 = x$.

В этом случае

$$K_{11} = \langle 1 \cdot 1 \rangle = N; \quad K_{12} = K_{21} = \sum_{i=1}^N x_i = N \bar{x}; \quad K_{22} = \sum_{i=1}^N x_i^2.$$

Чтобы матрица K_{ij} была диагональной, удобнее взять за исходные функции $u_1 = 1$, $u_2 = x - \bar{x}$, тогда $\langle u_1 \cdot u_2 \rangle = \sum (x_i - \bar{x}) = 0$ и матрица K_{ji} имеет лишь диагональные элементы:

$$K_{11} = N; \quad K_{22} = \sum_{i=1}^N (x_i - \bar{x})^2 \equiv N L^2.$$

Наилучшие коэффициенты \bar{a} и \bar{b} определяются из (72):

$$\bar{a} = \frac{\langle y \rangle}{N}; \quad \bar{b} = \frac{\langle y (x - \bar{x}) \rangle}{N L^2}.$$

Квадратичная форма при наилучших коэффициентах определяет суммарную квадратичную погрешность:

$$\Delta^2 = \sum_{i=1}^N (y_i - \bar{a} - \bar{b} (x_i - \bar{x}))^2,$$

а при произвольных a и b :

$$f = \Delta^2 + N (a - \bar{a})^2 + N L^2 (b - \bar{b})^2,$$

и формула (70) приводит к распределению параметров a , b , D :

$$dp(a, b, D) = N^{-1} \frac{1}{D^{\frac{N}{2}}} \left(e^{-\frac{\Delta^2 + N (a - \bar{a})^2 + N L^2 (b - \bar{b})^2}{2 D}} \right) da db \frac{dD}{D}. \quad (73)$$

После интегрирования по D эта формула дает *двумерное распределение Стьюдента* для коэффициентов a и b , пропорциональное

$$dp(a, b) \sim \frac{da db}{\left(1 + \frac{N ((a - \bar{a})^2 + L^2 (b - \bar{b})^2)}{\Delta^2} \right)^{\frac{N}{2}}}.$$

Если же сначала проинтегрировать по b , что просто уменьшит на $1/2$ степень по D в знаменателе, а затем по D , то результатом будет распределение коэффициента a – распределение Стьюдента с $N - 2$ степенями свободы:

$$dp(a) \sim \frac{da}{\left(1 + \frac{(N-1)(a-\bar{a})^2}{\Delta^2}\right)^{\frac{N-1}{2}}} \sim \frac{dy}{\left(1 + \frac{y^2}{m}\right)^{\frac{m+1}{2}}}. \quad (74)$$

Аналогично, интегрирование по a и D приводит к распределению Стьюдента для коэффициента b :

$$dp(b) \sim \frac{db}{\left(1 + \frac{(N-1)L^2(b-\bar{b})^2}{\Delta^2}\right)^{\frac{N-1}{2}}} \sim \frac{dy}{\left(1 + \frac{y^2}{m}\right)^{\frac{m+1}{2}}}. \quad (75)$$

Из этих выражений находим математические ожидания и дисперсии коэффициентов a и b :

$$\langle a \rangle = \bar{a} = \frac{1}{N} \sum_{i=1}^N y_i; \quad \langle b \rangle = \bar{b} = \frac{1}{N L^2} \sum_{i=1}^N (x_i - \bar{x}) y_i;$$

$$D_a = \frac{(N-2)}{(N-4)} \frac{\Delta^2}{(N-1)}; \quad D_b = \frac{(N-2)}{(N-4)} \frac{\Delta^2}{(N-1) L^2}$$

Каждый из них распределен по Стьюденту с $m = N - 2$ степенями свободы.

14.2 Регрессионный прогноз

Статистическая неопределенность коэффициентов регрессии a и b приводит к необходимости при прогнозировании проводить усреднение по этим параметрам с распределением (73):

$$dp(x) = \int_{a,b,D} dp(x/a, b, D) dp(a, b, D). \quad (76)$$

Этот интеграл все того же гауссова типа, который после интегрирования по D приводит к распределению Стьюдента. Нужно расписать квадратичную форму в показателе экспоненты:

$$f = \Delta^2 + N(a - \bar{a})^2 + N L^2 (b - \bar{b})^2 + (y - a - b(x - \bar{x}))^2.$$

Теперь нужно найти минимум этой формы по a и b :

$$\tilde{a} = \frac{\bar{a}(N L^2 + x(x - \bar{x})^2) + L^2(y - \bar{b}(x - \bar{x}))}{(N + 1)L^2 + (x - \bar{x})^2}.$$

$$\tilde{b} = \frac{\bar{b}L^2(N + 1) + (x - \bar{x})(y - \bar{a})}{(N + 1)L^2 + (x - \bar{x})^2}.$$

При этом сама квадратичная форма оказывается равной

$$f = \frac{N L^2(y - \bar{a} - \bar{b}(x - \bar{x}))^2}{(N + 1)L^2 + (x - \bar{x})^2} + \Delta^2.$$

Добавление новой точки увеличит степень D в знаменателе на $1/2$, однако интегрирование по a и b две половинки съедает, так что окончательно

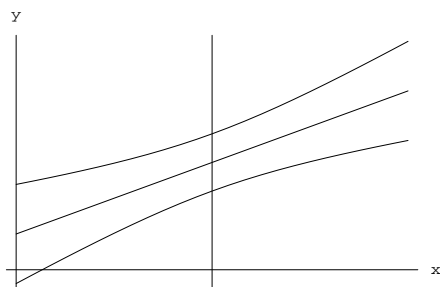
$$dp(y(x)) \sim \frac{dy}{\left(1 + \frac{N L^2 (y - \bar{a} - \bar{b}(x - \bar{x}))^2}{((N + 1)L^2 + (x - \bar{x})^2)\Delta^2}\right)^{\frac{N-1}{2}}} \sim \frac{dS}{\left(1 + \frac{S^2}{m}\right)^{\frac{m+1}{2}}}, \quad (77)$$

Это распределение Стьюдента с $m = N - 2$ степенями свободы. Сама прогнозируемая величина является суммой регулярной части – наиболее вероятной регрессионной функции – и случайной величины S , распределенной по Стьюденту:

$$y(x) = \bar{a} + \bar{b}(x - \bar{x}) + S_{N-2} \sqrt{\Delta^2 \left(1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{N L^2}\right)} \quad (78)$$

Математическое ожидание y лежит на регрессионной кривой $y = \bar{a} + \bar{b}(x - \bar{x})$, а дисперсия ее различна при различных x :

$$\begin{aligned} D_x &= \frac{(N - 2)}{(N - 4)} \Delta^2 \left(1 + \frac{1}{N} + \frac{(x - \bar{x})^2}{N L^2}\right) = \\ &= \frac{(N - 2)(N + 1)}{(N - 4)N} \Delta^2 \left(1 + \frac{(x - \bar{x})^2}{(N + 1)\Delta^2}\right) \end{aligned} \quad (79)$$



Дисперсия, а вместе с ней и доверительный интервал заданной надежности, расширяются от середины графика к концам. Минимум дисперсии в точке $x = \bar{x}$.

15 Статистические гипотезы

Никакой конечный эксперимент не может определить вид распределения, которому подчиняется измеряемая случайная величина (за исключением бинарного распределения – с двумя исходами). Поэтому исследователь, исходя из природы этой случайной величины, выдвигает *гипотезу* о виде распределения и, например, методом Байеса определяет параметры этого распределения.

Например, для непрерывной случайной величины, случайность которой определяется многими факторами, вследствие центральной предельной теоремы чаще всего принимается гипотеза о нормальном распределении.

После проведения эксперимента возникает вопрос: а согласуются ли полученные экспериментальные данные с выдвинутой гипотезой?

Критерии согласия разработаны для того, чтобы дать *вероятностный* ответ на данный вопрос.

15.1 Критерий χ^2

Применяется для проверки гипотезы о том, что эксперимент объема n с k состояниями, в результате которого в каждом i -м состоянии выпало r_i событий, определяется вероятностями p_i , оцененными

какими-либо статистическими методами. При этом

$$\sum_{i=1}^k r_i = n. \quad (80)$$

Математические ожидания чисел r_i равны $\langle r_i \rangle = m_i = n p_i$, а при достаточно больших m_i (больших 10-и) отклонения $r_i - m_i$ вследствие центральной предельной теоремы подчиняются нормальному распределению с дисперсией

$$D_i = n p_i (1 - p_i) \approx n p_i,$$

если состояний достаточно много и каждое p_i мало.

Тогда величину

$$z_i = \frac{r_i - n p_i}{\sqrt{n p_i}} = \frac{r_i - m_i}{\sqrt{m_i}}$$

можно полагать нормально распределенной с единичной дисперсией, а величина

$$u = \sum_{i=1}^k z_i^2 = \sum_{i=1}^k \frac{(r_i - m_i)^2}{m_i} \quad (81)$$

распределена примерно по закону χ^2 .

Нужно еще найти число статистических степеней свободы. Числа r_i подчиняются упомянутой связи (80), что уменьшает число статистических степеней свободы на единицу. Если данные используются также для оценки t параметров распределения, то статистически независимыми остается

$$\nu = k - 1 - t \quad (82)$$

параметров, что и определяет число степеней свободы в χ^2 -распределении.

Площадь под кривой χ^2 -распределения с ν степенями свободы от нуля до вычисленного значения u определяет вероятность *неверности гипотезы*. Если u велико, вероятность отклонения гипотезы оказывается высокой.

16 Статистическое моделирование

Статистическое моделирование (метод Монте-Карло) применяется в задачах теории вероятностей, где простые, изначальные события моделируются достаточно просто, но вероятности некоторых сложных событий, определяемых простыми, вычислить невозможно или очень сложно. Проводится многократная случайная реализация исходных событий с заданными вероятностями, а затем на основании законов физики из них выводятся необходимые сложные события.

16.1 Датчики случайных чисел

Для моделирования исходных событий нужны *датчики случайных чисел* с заданными вероятностями. В пакете Mathematica имеются следующие функции для генерации непрерывных или целых случайных чисел:

- `Random[]` – выдает случайное число от 0 до 1 с равномерным распределением.
- `Random[Integer]` – выдает 0 или 1 с вероятностью $1/2$.
- `Random[Integer,10]` – равновероятно выдает случайные целые числа от 0 до 10.
- `Random[Integer,{1,100}]` – равновероятно выдает случайные целые числа от 1 до 100.
- `Random[Real,{0,10}]` – равновероятно выдает случайные вещественные числа от 0 до 10.

16.2 Моделирование первой предельной теоремы

Модуль на Математике:

```
first[p_, n_] := Module[{s, z},
  s = 0; Do[z = Random[]; If[z < p, s++], {i, n}];
  s/n // N]
```

позволяет при заданной вероятности p и числе реализаций n получить среднее и сравнить его значение с заданной вероятностью.

16.3 Моделирование multistep-распределения

На примерах step и multistep-распределения (24) мы рассмотрим процедуру статистического моделирования выборки с заданным распределением вероятностей. Пусть заданы значения вероятностей p и q . Моделирование n шагов со случайными значениями $(-1, 0, +1)$, с вероятностями $(q, 1 - p - q, p)$ соответственно можно построить следующим модулем:

```
nstep[p_, q_, n_] := Module[{x, z, s}, s = 0;
Do[x = Random[]; If[x < p + q, If[x < q, s--, s++], {i, n}];
s]
```

Можно смоделировать серию таких сдвигов из m экземпляров:

```
nmstep[p_, q_, n_, m_] :=
Module[{s, ms, DD}, ll = {};
Do[AppendTo[ll, step[p, q, n]], {i, m}];
s = Sum[ll[[i]], {i, m}]; ms = s/m;
DD = Sum[(ll[[i]] - ms)^2, {i, m}];
{ll, {ms, DD/m}}/N
]
```

В модуле также вычисляется среднее значение (ms) и суммарный квадрат отклонения от среднего (DD).

16.4 Вычисление интегралов методом Монте-Карло

Если с помощью функции Random[] выбрасывать пары чисел и придавать им смысл координат (x, y) точки на плоскости, то при многократном повторении точки достаточно равномерно заполняют на плоскости квадрат, от 0 до 1 вдоль оси x и так же вдоль оси y (равномерность определяется *качеством* датчика случайных чисел).

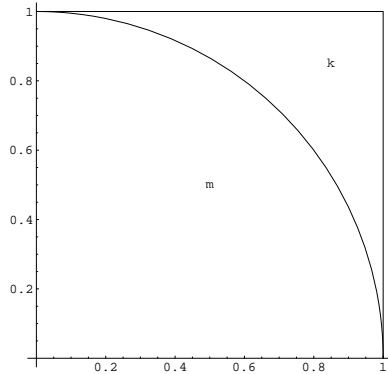
Если в этом квадрате проведена некоторая кривая, то отношение числа точек под кривой (m) к общему числу точек (n) приблизительно равняется отношению площади под кривой, к площади квадрата (1), то есть равняется

$$\int_0^1 y(x) dx \approx \frac{m}{n}.$$

Если интеграл нужно вычислить на интервале (a, b) , то координата x связывается со случайным числом ξ в интервале от нуля до единицы $x = a + (b - a)\xi$ и

$$\int_a^b y(x) dx \approx (b - a) \frac{m}{n}.$$

Рассмотрим, например, площадь под четвертью окружности в первом квадранте с уравнением $x^2 + y^2 = 1$. Радиус окружности единица и площадь должна равняться $\pi/4$. Вычисляя методом Монте-Карло площадь под этой кривой (учетверив ее) можно приблизительно вычислить число π :



Модуль такого расчета на Mathematic'e приведен ниже:

```
pi[n_] := Module[{x, y, s}, s = 0;
  Do[x = Random[ ]; y = Random[ ];
    If[x^2 + y^2 < 1, s++], {i, n}];
  4s/n // N
]
```

Разные реализации дают различные приближенные значения числа π .

Можно организовать серию вычислений с вычислением статистического среднего s и среднего квадратичного разброса DD :

```

npi[n_, m_] :=
Module[{ll, s, DD}, ll = {};
Do[AppendTo[ll, pi[n]], {i, m}];
s = Sum[ll[[i]], {i, m}]/m;
DD = Sum[(ll[[i]] - s)^2, {i, m}]/m;
{ll, {s, DD}}]

```

Работа с этим модулем показывает, что с увеличением n (100, 1000, 10000) среднеквадратичный разброс уменьшается.

Приведем еще модуль вычисления интеграла

$$\int_0^{\frac{\pi}{2}} \sin x \, dx = 1.$$

```

si[n_] := Module[{x, y, s}, s = 0;
Do[x = Random[ ]; y = Random[ ];
If[y < Sin[\[Pi] x], s++, {i, n}];
\[Pi] s/n/2 // N
]

```

Как и в предыдущем случае можно организовать серию таких испытаний с вычислением среднего значения и среднеквадратичного разброса:

```

nsi[n_, m_] :=
Module[{ll, s, DD}, ll = {};
Do[AppendTo[ll, si[n]], {i, m}];
s = Sum[ll[[i]], {i, m}]/m;
DD = Sum[(ll[[i]] - s)^2, {i, m}]/m;
{ll, {s, DD}}]

```

Функция для интегрирования может быть задана и неаналитически. Должен быть задан алгоритм определения нахождения точки под кривой.

16.5 Пакет расширения “Статистика”

Чтобы сделать доступным нужное распределение, необходимо подключить соответствующий раздел пакета:

```
< <Statistics‘DiscreteDistributions‘
```

для BinomialDistribution[n,p], PoissonDistribution[mu] и др.,

```
< <Statistics‘ContinuousDistributions‘
```

для γ -, β -распределений и других

```
< <Statistics‘NormalDistribution‘
```

для нормального, Стьюдента, χ^2 , F-распределения Фишера.

После подключения соответствующего пакета для объявленных в нем распределений становятся доступными следующие функции, где `dist` – переменная с присвоенным ей распределением, например,

`pu5=PoissonDistribution[5]` – распределение Пуассона с $\mu = 5$:

PDF[dist,x] | плотность вероятности как функция x

CDF[dist,x] | интегральная вероятность как функция x

Quantile[dist,q] | квантиль, отсекающий площадь q

Domain[dist] | область изменения переменной

Mean[dist] | математическое ожидание

Variance[dist] | дисперсия

StandardDeviation[dist] | стандартное отклонение

ExpectedValue[f, dist, x] | математическое ожидание функции x

Random[dist] | псевдослучайное число с распределением `dist`

RandomArray[dist,n] | генерирует массив из n псевдослучайных чисел с распределением `dist`

Например, для генерации 12 случайных чисел с объявленным выше распределением `pu5` нужно выполнить `RandomArray[pu5, 12]`.

Обработка массивов случайных чисел (вычисление средних, разброса и пр.) становится доступной после запуска пакета

```
< <Statistics‘DescriptiveStatistics‘
```

Help и примеры в разделе Help: Add-Ons/Standard Packages/Statistics.

Часть III

Задания

17 Задания

17.1 Распределения

Для случайной величины, имеющей распределение

- дискретное: биномиальное, Пуассона, Паскаля;
- непрерывное: нормальное, β -распределение, Γ -распределение, Максвелла, Релея, χ^2

1. выбрать значения параметров;
2. построить график;
3. на графике указать: *центр и наиболее вероятное значение* – вертикальными линиями, математическое ожидание – жирной точкой, а *стандартный интервал* $< x > \pm \sigma$; $\sigma = \sqrt{D_x}$ указать отрезками горизонтальных линий длиной σ влево и вправо от математического ожидания

17.2 Семиинварианты

Вычислить разложением в ряд производящей функции семиинвариантов математическое ожидание, дисперсию, асимметрию, эксцесс для распределений (**основного набора**):

биномиального, Пуассона, Паскаля, n step, β -распределения, Γ -распределения, Максвелла, Релея, χ^2 .

17.3 Распределение Стьюдента

- Записать явную формулу распределения Стьюдента с $m = 5$ степенями свободы.

- Построить это распределение в сравнении со стандартным нормальным (на одном графике) ($-3 < x < 3$).
- Прямым вычислением убедиться в нормировке этого распределения. Для этого создать функцию с формальным параметром m

$$B[m_] := \int_{-\infty}^{\infty} \frac{dx}{\left(1 + \frac{x^2}{m}\right)^{\frac{m+1}{2}}}.$$

Вычислить $B[m]$ при $m = 1, 3, 5, 7$.

Построить ряд отношений $B[m+2]/B[m]$.

Вычислить $B[m]$ при $m = 2, 4, 6$.

Построить ряд отношений $B[m+2]/B[m]$.

Убедиться, что эти соотношения согласованы с определением нормированной функции (47) на странице 35.

17.4 Моделирование распределений

Для *одного* из распределений *основного набора* (стр. 66) смоделировать последовательность n случайных величин, подчиняющихся этому распределению, и записать в файл.

По критерию согласия χ^2 проверить соответствие полученной последовательности заданному распределению.

17.5 Статистика параметров

Прочитать из файла (сформированного по предыдущему пункту другим студентом) последовательность n случайных чисел. Зная распределение построить статистику параметров распределения.

17.6 Критерии значимости

Сгенерировать две последовательности нормально распределенных случайных величин с одинаковой дисперсией. По критерию Стьюдента определить различие математических ожиданий, оцененных по этим последовательностям.

17.7 Регрессия

Сгенерировать последовательность точек на плоскости с линейной зависимостью $y = a + bx$ и шумом, добавляемым в виде случайной величины $k(1 - 2\text{Random}[])$, принимающей равновероятно значения от $-k$ до $+k$, где k – амплитуда шума.

По полученной последовательности точек на плоскости построить регрессионную прямую, оценить интенсивность шума и построить доверительный интервал заданной надежности.

17.8 Интегрирование методом Монте-Карло

Вычислить методом Монте-Карло заданные определенные интегралы.

Список литературы

- [1] Худсон Д. Статистика для физиков. М.: Мир, 1967
- [2] Корн Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука, 1970.
- [3] Моррис У.Т. Наука об управлении. Байесовский подход. Мю: Мир, 1971.
- [4] Рао С.Р. Линейные статистические методы и их применения. М.: Наука, 1968. [13.2](#)
- [5] Кендалл М., Стьюарт А. Статистические методы и связи. М.: Наука, 1973.
- [6] Муравьев В.А., Бурланков Д.Е. *Практическое введение в пакет Mathematica*. Нижний Новгород: Изд. ННГУ, 2000.