УДК 519.237

ГЕОМЕТРИЧЕСКАЯ ИНТЕРПРЕТАЦИЯ МНОГОМЕРНЫХ РАСПРЕДЕЛЕНИЙ

GEOMETRIC INTERPRETATION OF MULTIDIMENSIONAL DISTRIBUTIONS

А. И. Пьянзин

A. I. Pyanzin

ФГБОУ ВПО «Чувашский государственный педагогический университет им. И. Я. Яковлева», г. Чебоксары

Аннотация. В статье представлена геометрическая интерпретация одномерного и многомерного распределений. На основе логического анализа выявлен способ графической визуализации многомерного распределения с использованием цветового градиента. Если аргументы выразить через количественную насыщенность элементарных событий с помощью цветового градиента, то одномерное нормальное распределение будет представлено в виде полосы (шкалы), двумерное – в виде окружности, трехмерное – в виде сферы с разной степенью цветовой насыщенности, с максимальной насыщенностью в зоне средних значений аргументов. Приведены примеры, подтверждающие их приложение в различных областях научного знания.

Abstract. The paper presents a geometric interpretation of one-dimensional and multidimensional distributions. Based on the method of logical analysis the graphic visualization of multidimensional distributions with a color gradient has been revealed. If the arguments are expressed in terms of elementary events quantitative saturation in the color gradient, the one-dimensional normal distribution is presented as a strip, a two-dimensional distribution can be presented as a circle and three-dimensional distribution can be presented as a sphere with different degrees of color saturation with the maximum saturation in the average values zone of arguments. The author gives the examples that confirm their application in various fields of scientific knowledge.

Ключевые слова: *статистика, распределение, одномерность, многомерность, визуализация, образ, насыщенность.*

Keywords: *statistics, distribution, one-dimensionality, multidimensionality, visualization, image, saturation.*

Актуальность исследуемой проблемы. При обработке данных измерений в науке и технике обычно предполагают нормальный закон распределения. Чаще всего при решении прикладных задач используется одномерное нормальное распределение, однако не всегда при построении функции распределения одного аргумента бывает достаточно. Принципиально ничто не ограничивает нас от использования двух и более аргументов, но здесь мы сталкиваемся с проблемой геометрической интерпретации (визуализации) образа такого распределения. Ведь наше восприятие ограничено тремя пространственными измерениями, два из которых уже занимают качественные варьирующие признаки, а третье измерение графически выражает частоту их проявления. Проблема геометрической интерпретации многомерных распределений связана с тем, что необходимо изобразить три пространственных измерения и еще одно – четвертое – для иллюстрации искажения первых трех. Актуальность работы заключается в поиске ответа на вопрос, насколько возможна такая геометрическая интерпретация.

Графическая интерпретация наиболее полно представлена в специализированных пакетах Maple и Mathematica, которые позволяют получать быстрое общее графическое представление (набросок) алгебраического выражения в интерактивном режиме, пусть не совсем точное и хорошо оформленное, но отвечающее задачам, в частности исследовательского характера, когда анализ результата определяет путь дальнейшей работы и графическая составляющая играет определенную роль.

Материал и методика исследований. В ходе исследования применялись следующие методы: анализ научно-методической литературы по проблеме исследования, графическое моделирование, логический анализ. В работе использовался материал, представленный в справочной и научной литературе по математической статистике и физике, а также данные прикладных научных исследований в области спортивной подготовки.

Результаты исследований и их обсуждение. Нормальное распределение представляет собой одну из эмпирически проверенных истин относительно общей природы действительности. Оно также может рассматриваться как один из фундаментальных законов природы. Как показывают многочисленные научные данные, нормальный тип распределения широко распространен в природе, так как, в принципе, любые явления и процессы имеют вероятностный характер проявления и зависят от большого числа случайных причин без доминирующего влияния какой-либо из них, что является условием действия нормального закона. Если известен функциональный вид распределения случайной величины, то можно получить полную информацию о вероятности ее реализации в любом заданном интервале значений. При обработке данных измерений в науке и технике обычно предполагают нормальный закон распределения.

Нормально распределенная случайная величина имеет следующие свойства:

1) она может принимать непрерывный ряд значений от $-\infty$ до $+\infty$;

2) распределение, задаваемое функцией Гаусса, симметрично относительно максимума; центр распределения случайной величины одновременно является центром симметрии, т. е. одинаковые отклонения результатов измерения в противоположные стороны от центра встречаются одинаково часто;

3) малые отклонения от центра встречаются чаще больших, другими словами, реализуются с большей вероятностью.

Точная форма нормального распределения (характерная «колоколообразная кривая» – рис. 1) определяется только двумя параметрами: средним ($\boldsymbol{x}$) и стандартным отклонениями ($\boldsymbol{\sigma}$). Значение аргумента $\boldsymbol{x}$, при котором плотность вероятности максимальна, является наиболее вероятным при реализации случайной величины. Значение функции уменьшается с увеличением $\boldsymbol{\sigma}$, но полная площадь под кривой остается при этом неизменной.

Практика спортивной деятельности показала, что для большинства видов спорта лучшей формой распределения варьирующего признака является часто встречающееся в биологии и биометрии нормальное распределение [12]. В частности, М. Байбак [1], В. В. Бойко [2] показывают возможности применения закона нормального распределения в области спортивной тренировки, где ось $\boldsymbol{x}$ отражает значения качественного или варьирующего признака движений (их интенсивность, сложность). Величина $\boldsymbol{y}$ показывает частоту, иначе говоря, количество (объем) того или иного аргумента ($\boldsymbol{x}$).

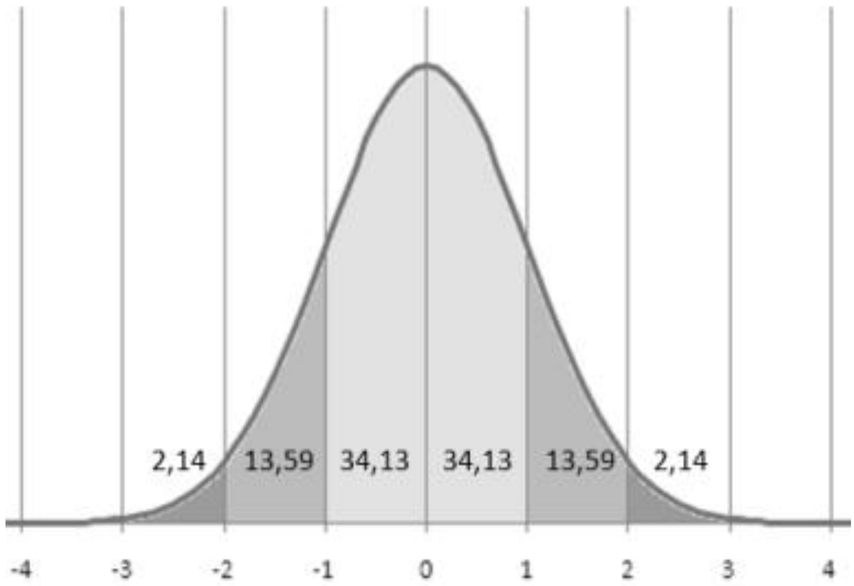


Рис. 1. ***Одномерное нормальное распределение***

Чаще всего при решении прикладных задач используется одномерное нормальное распределение, основанное на одном качественном варьирующем признаке (***x***). Однако не всегда при построении функции распределения одной переменной бывает достаточно. Тогда приходится использовать двумерное нормальное распределение (рис. 2). В этом случае частота (количественный параметр – $\boldsymbol{W(x_1, x_2)}$) принимает форму купола, в основании которого лежат два аргумента ($\boldsymbol{x_1}$ и $\boldsymbol{x_2}$).

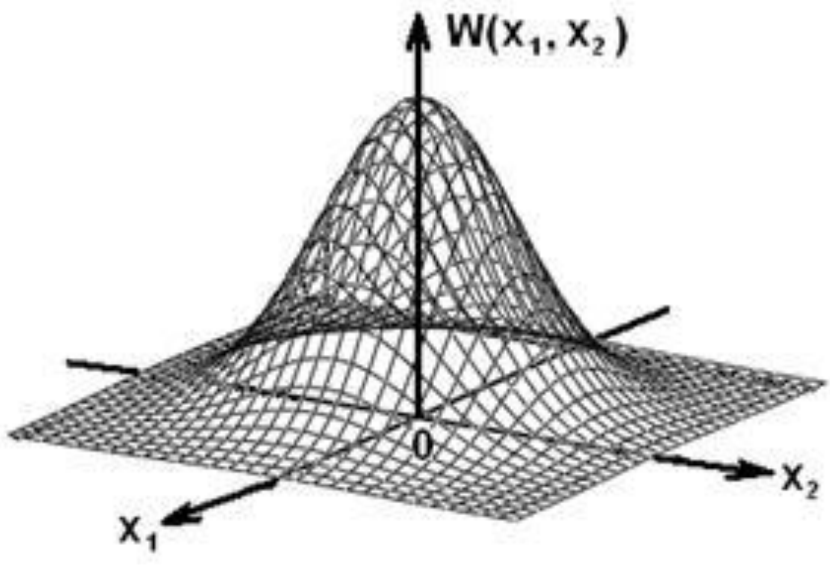


Рис. 2. ***Двумерное нормальное распределение***

Любая линия пересечения поверхности распределения с вертикальной плоскостью, параллельной одной из координатных осей, имеет форму нормального распределения, а кривая пересечения поверхности распределения с горизонтальной плоскостью представляет собой эллипс [11].

Двумерное нормальное распределение, в частности, легло в основу теоретической модели структуры тренировочной нагрузки в легкоатлетических прыжках [8], [3], [9], в которой варьирующими качественными признаками являются два ключевых биодинамических параметра взаимодействия тела спортсмена с опорой при отталкивании (вертикальная составляющая скорости вылета общего центра масс тела и среднее значение вертикальной составляющей реакции опоры), а количественным – объем нагрузки, или число повторений, приходящееся на каждое упражнение в зависимости от характерных для него значений скорости вылета и силы опорной реакции. При этом горизонтальная плоскость распределения («поле средств»), представляющая собой основание купола, раскрывает «географию проживания» отдельных упражнений с присущими им параметрами скорости и силы и позволяет увидеть их взаимное расположение, а значит и степень их качественной однородности или неоднородности при оценке воздействий на организм спортсмена.

В то же время принципиально ничто не ограничивает нас от использования более чем двух качественных варьирующих признаков, и математический аппарат позволяет это сделать. Но здесь мы сталкиваемся с проблемой геометрической интерпретации (визуализации) образа такого распределения, ведь наше восприятие ограничено тремя пространственными измерениями, два из которых уже занимают качественные варьирующие признаки двумерного нормального распределения ($\boldsymbol{x_1}$ и $\boldsymbol{x_2}$), а третье измерение графически (в виде купола) выражает частоту их проявления ($\boldsymbol{W(x_1, x_2)}$).

Подобные модели двумерного распределения используются и для иллюстрации закона всемирного тяготения как проявления искривления пространственно-временной материи вблизи массивных космических объектов (рис. 3) [5], [13]. «Согласно теории тяготения Эйнштейна, истинное гравитационное поле является ничем иным, как проявлением искривления четырехмерного пространства-времени» [7].

Падение тел в гравитационном поле И. А. Климишин [4] иллюстрирует посредством их свободного движения по геодезической (экстремальной) линии в искривленном четырехмерном пространстве-времени. Представим себе горизонтально натянутую резиновую пленку (мембрану), на которой нанесена декартова сетка координат. Если на пленку положить тяжелое тело, то под действием его веса мембрана прогнется. В этом искривленном двумерном мире геометрия уже не будет евклидовой. Шарик, двигаясь с начальной скоростью $\boldsymbol{v}$ по мембране, будет скатываться («притягиваться») к тяжелому телу.

Модель искривления пространственно-временной материи очень наглядно объясняет, почему менее массивный объект скатывается по спирали в «гравитационный мешок» пространства, «продавленный» более массивным объектом. Хотя, в силу ограниченности нашего восприятия тремя измерениями, в этой графической модели присутствуют лишь два пространственных измерения из трех, потому что место третьего измерения занято иллюстрацией искривления первых двух пространственных измерений в виде воронки или перевернутого купола.

Таким образом, гравитационная модель, изображенная на рисунке 3, отражает лишь два пространственных измерения (плоскость), искривленных массой объекта (воронка или углубление в плоскости). Назовем ее двумерной воронкой, хотя в объективной реальности данная воронка является трехмерной. Но с формированием понятного визуального образа трехмерной воронки возникают трудности, связанные с тем, что необходимо изобразить три пространственных измерения и еще одно – четвертое – для иллюстрации искажения первых трех. Насколько реальна такая геометрическая интерпретация?

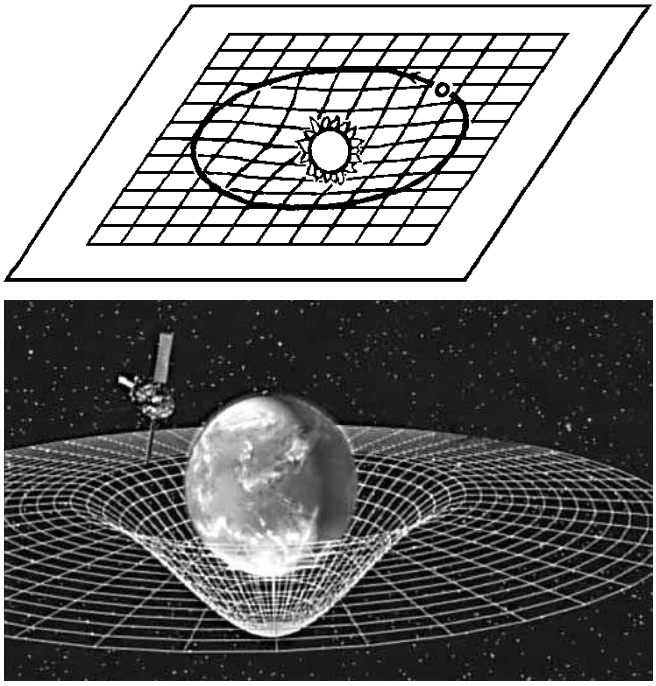
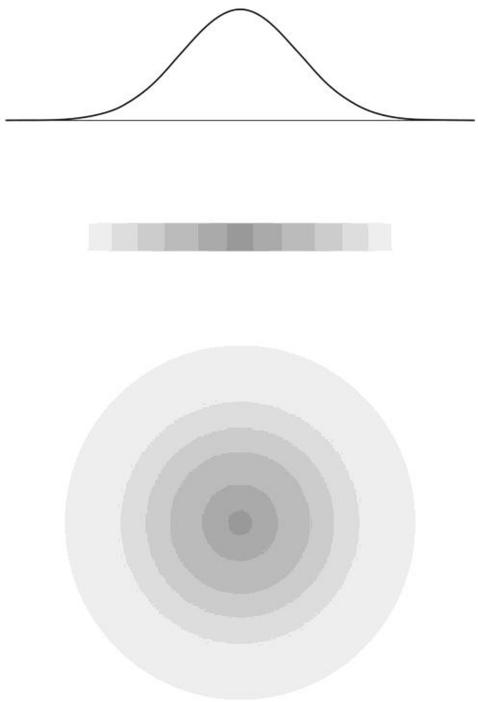


Рис. 3. ***Графическая иллюстрация притяжения тел с меньшей массой к телам с большей массой вследствие искривления пространства***

Следует отметить, что гравитационная воронка очень напоминает перевернутый (в нашем традиционном восприятии пространственных координат) купол нормального распределения (ср. рис. 2 и 3). Их схожесть не только визуальна, но и функциональна. Кривая нормального распределения отражает *концентрацию* или *количественную насыщенность* случайных событий в зависимости от степени их близости к среднему значению качественного признака, то есть к условному вероятностному *центру*. Гравитационная воронка отражает *концентрацию* или *количественную насыщенность* элементарных носителей массы вещества в зависимости от степени их близости к условному гравитационному *центру* материального объекта.

Задача геометрической интерпретации четырехмерного образа становится выполнимой с введением цветового градиента. Если количественный признак выразить через количественную насыщенность элементарных событий (или элементарных носителей массы) с помощью цветового градиента, то одномерное нормальное распределение будет представлено в виде полосы (шкалы) с разной степенью цветовой насыщенности, с максимальной насыщенностью в зоне среднего значения качественного признака, т. е. в центре шкалы (рис. 4 а).

Двумерное нормальное распределение будет представлено в виде плоской окружности с разной степенью цветовой насыщенности, с максимальной насыщенностью в зоне средних значений двух качественных признаков, т. е. в центре окружности (рис. 4 б).



а

б

Рис. 4. ***Градиентная визуализация одномерного (а) и двумерного (б) нормального распределения***

Трехмерное нормальное распределение будет представлено в виде сферы с разной степенью цветовой насыщенности, с максимальной насыщенностью в зоне средних значений трех качественных признаков, т. е. в центре сферы (рис. 5). В этом случае речь идет о многомерном нормальном (гауссовском) распределении как обобщении одномерного нормального распределения [6].

При решении задачи измерения трех нормально распределенных параметров у $\boldsymbol{n}$ индивидуумов получается ситуация, где $\boldsymbol{n}$ точек сосредоточены в трехмерном пространстве с тремя осями переменных $\boldsymbol{X}$, $\boldsymbol{Y}$, $\boldsymbol{Z}$ в облаке вокруг общего центра тяжести. Это облако точек наблюдений в общем случае имеет овальную форму и называется эллипсоидом. В частном случае, когда во всех трех направлениях дисперсия одинакова по величине, получают шар [10]. Этот шар или эллипсоид не имеют четко очерченных границ, поскольку хвосты распределения в любом пространственном измерении бесконечны.

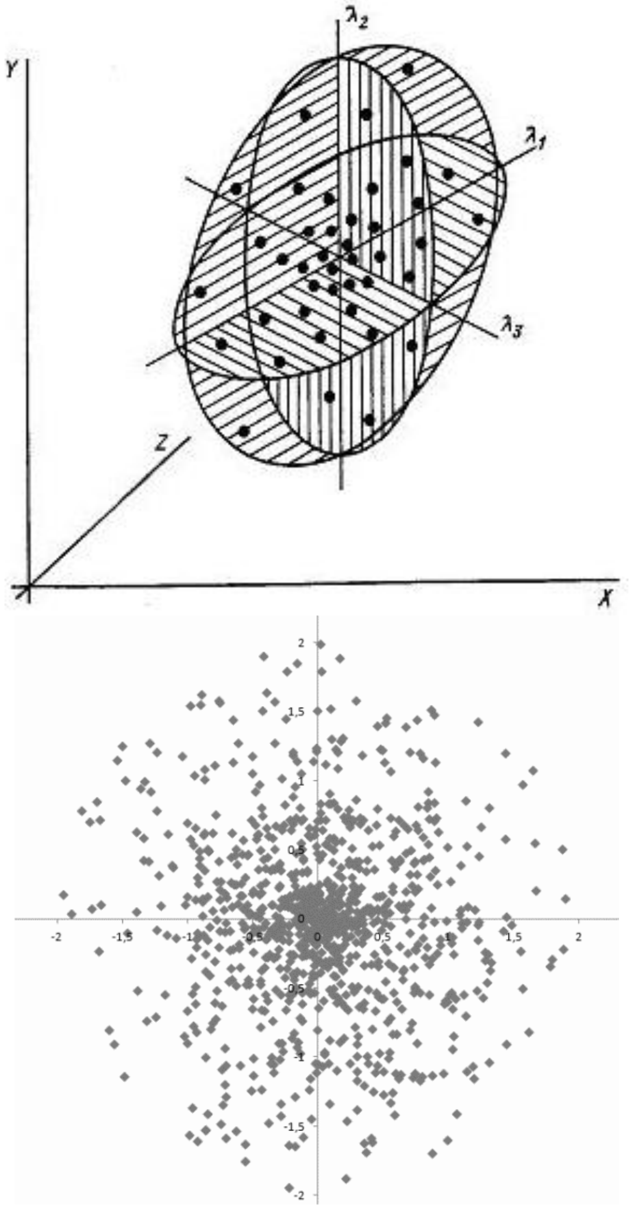


Рис. 5. ***Градиентная визуализация трехмерного распределения точек с соответствующими главными осями***

Резюме. В свете изложенного можно сделать вывод о том, что введение цветового градиента делает выполнимой задачу геометрической интерпретации многомерного распределения. Если аргументы выразить через количественную насыщенность элементарных событий с помощью цветового градиента, то одномерное нормальное распределение будет представлено в виде полосы (шкалы), двумерное – в виде окружности, трехмерное – в виде сферы с разной степенью цветовой насыщенности, с максимальной насыщенностью в зоне средних значений аргументов. Четырехмерная геометрическая интерпретация распределения представляет собой «вероятностный сгусток» – облако с различной плотностью случайных событий в случае с многомерным нормальным распределением, или «гравитационный сгусток» – облако с различной плотностью вещества как носителя массы в случае с гравитационным искажением пространственно-временной материи.

ЛИТЕРАТУРА

1. *Байбак, М.* Систематизация специальных подготовительных упражнений, применяемых при совершенствовании техники отталкиваний в тройном прыжке с разбега : автореф. дис… канд. пед. наук : 73.00.04 / М. Байбак. – М., 1988. – 22 с.
2. *Бойко, В. В.* Целенаправленное развитие двигательных способностей человека / В. В. Бойко. – М. : Физкультура и спорт, 1987. – 144 с.
3. *Взаимосвязь* специальных упражнений в легкоатлетических прыжках / А. И. Пьянзин и др. // Теория и практика физической культуры. – 2004. – № 10. – С. 33–35.
4. *Климишин, И. А.* Релятивистская астрономия / И. А. Климишин. – М. : Наука, 1983. – С. 208.
5. *Кронов, Н.* Космогония: пространство расширяется, потому что время течет [Электронный ресурс] / Н. Кронов. – Режим доступа: http://n-kronov.chat.ru/cosm2.html
6. *Многомерное* нормальное распределение [Электронный ресурс]. – Режим доступа: http://ru.wikipedia.org/wiki/Многомерное_нормальное_распределение
7. *Новиков, И. Д.* Эволюция Вселенной / И. Д. Новиков. – М. : Наука, 1983. – 192 с.
8. *Пьянзин, А. И.* Модель ранжирования специальных тренировочных средств в прыжковых видах легкой атлетики / А. И. Пьянзин // Теория и практика физической культуры. – 2001. – № 3. – С. 28–30.
9. *Пьянзин, А. И.* Спортивная подготовка легкоатлетов-прыгунов / А. И. Пьянзин. – М. : Теория и практика физической культуры, 2004. – 370 с.
10. *Факторный* анализ [Электронный ресурс]. – Режим доступа: http://www.wikiznanie.ru/ru-wz/index.php/Факторный_анализ
11. *Хальд, А.* Двумерное нормальное распределение / А. Хальд // Математическая статистика с техническими приложениями. – М. : Иностранная литература, 1956. – С. 501–504.
12. *Энгвер, Н. Н.* Построение эмпирических формул и моделей в спорте / Н. Н. Энгвер, Я. И. Савицкий, М. Г. Гибадуллин // Теория и практика физической культуры. – 1986. – № 10. – С. 35–37.
13. *Quantum* gravity [Электронный ресурс]. – Режим доступа: http://www.stfc.ac.uk/433.aspx