

## ПРИМЕНЕНИЕ МНОГОМЕРНОГО НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДЛЯ ПОСТРОЕНИЯ ВЕРОЯТНОСТНОЙ ОЦЕНКИ КОМПЕТЕНТНОСТИ

© 2014 Н.В. Колачева, О.А. Кузнецова

Тольяттинский государственный университет

Поступила в редакцию 13.01.2014

В данной статье рассмотрено многомерное нормальное распределение, определена доверительная область вектора средних, рассмотрен пример построения доверительной области оценки компетентности студентов.

*Ключевые слова:* многомерное нормальное распределение, доверительная область, компетентность.

Оценка компетентности студентов в вузе является одной из важнейших задач эффективности подготовки конкурентоспособных специалистов, в том числе экологов. Образовательные стандарты третьего поколения предписывают внедрение компетентно – ориентированного обучения в вузе.

Компетенция – это предметная область, в которой индивид хорошо осведомлен и в которой он проявляет готовность к выполнению профессиональной деятельности. Компетенция – это ситуативная категория, так как она выражается в готовности к осуществлению деятельности в конкретных профессиональных ситуациях. Компетентность – *интегральная характеристика личности, распадающаяся на спектр отдельных компетенций*, включает в себя когнитивные, мотивационно-ценностный и эмоционально-волевой компоненты.

Основные признаки или факторы компетенции можно оценить, используя таксономию Блума, согласно которой каждая компетенция может быть оценена по шести признакам: знание, понимание, применение, анализ, синтез и оценка.

Согласно ФГОС ВО (уровень бакалавриата) у выпускника должны быть сформированы общекультурные (ОК), общепрофессиональные (ОПК), профессиональные (ПК) и профессионально-прикладные (ППК) компетенции. Для каждого из направлений обучения имеются свои задачи формирования необходимых компетенций будущего специалиста, в том числе в области экологии [1-10].

Рассматривая соответствующий набор компетенций как  $n$ -мерный вектор  $X = (X_1, X_2, \dots, X_n)$ ,

получим четыре вектора компетентностей  $X_{OK}, X_{ОПК}, X_{ПК}, X_{ППК}$ . Согласно таксономии Блума каждая компонента вектора компетентности будет иметь шесть оценочных значений. Каждому признаку может быть присвоен весовой коэффициент и рассчитан обобщенный показатель оценки компетенции как компоненты соответствующего вектора компетентности.

Чтобы получить среднее значение оценки компетенции каждого вектора необходимо иметь набор многомерных статистических данных, полученных в результате опроса и тестирования специалистов данной предметной области.

Эти данные подлежат систематизации и обработке с целью выявления характера и структуры взаимосвязей между компонентами  $n$ -мерного вектора  $X = (X_1, X_2, \dots, X_n)$ . Эти задачи можно решить методами многомерного статистического анализа.

Многомерные статистические методы позволяют обоснованно выбрать ту модель, которая наилучшим образом соответствует исходным статистическим данным, характеризующим реальное поведение исследуемой совокупности объектов, оценить надежность и точность выводов, сделанных на основании ограниченного статистического материала.

Рассмотрим более подробно многомерное нормальное распределение, так как многомерные методы, основанные на нормальном распределении, нашли широкое распространение и их можно изучать более организованно и систематически. Одномерное нормальное распределение возникает очень часто, потому что изучаемый эффект является суммой действия многих независимых случайных факторов. Подобно этому многомерное нормальное распределение часто встречается в связи с тем, что множественные измерения являются суммами многих малых независимых воздействий. Кроме того, для теории нормального распределения разработаны точные математические методы.

Колачева Наталья Вениаминовна, кандидат педагогических наук, доцент кафедры высшей математики и математического моделирования. E-mail: ncolacheva@mail.ru.  
Кузнецова Ольга Александровна, кандидат педагогических наук, доцент кафедры высшей математики и математического моделирования. E-mail: oly-2009@yandex.ru

Одномерная нормальная плотность распределения может быть записана следующим образом:

$$k \cdot e^{-\frac{1}{2}\alpha(x-\beta)^2} = k \cdot e^{-\frac{1}{2}(x-\beta)\alpha(x-\beta)}, \quad (1)$$

где  $\alpha$  – положительно, а  $k$  выбирается таким образом, чтобы интеграл от (1) по всей оси  $x$  был равен единице. Плотность вероятности многомерного нормального распределения величин  $X_1, \dots, X_n$  имеет аналогичную форму. Скалярная переменная  $x$  заменяется вектором

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad (2)$$

скалярную постоянную  $\beta$  заменяют вектором

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \quad (3)$$

а положительную постоянную  $\alpha$  заменяют положительно определенной (симметрической) матрицей

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}. \quad (4)$$

Выражение  $\alpha(x-\beta)^2 = (x-\beta)\alpha(x-\beta)$  в этом случае заменяется квадратичной формой

$$(x-\mu)^T A(x-\mu) = \sum_{i,j=1}^n a_{ij}(x_i-\mu_i)(x_j-\mu_j). \quad (5)$$

Тогда плотность вероятности  $n$ -мерного нормального распределения равна

$$f(x_1, \dots, x_n) = \frac{\sqrt{|A|}}{(2\pi)^{\frac{1}{2}n}} e^{-\frac{1}{2}(x-\mu)^T A(x-\mu)} = (2\pi)^{-\frac{1}{2}n} |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)} \quad (6)$$

где  $\Sigma$  есть ковариационная матрица, равная  $A^{-1}$ .

Для двумерного нормального распределения полученная модель описывается следующим образом.

Среднее значение вектора  $X = (X_1, X_2)^T$  равно

$$M \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}; \quad (7)$$

ковариационная матрица может быть записана следующим образом:

$$\Sigma = M \begin{pmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 \end{pmatrix} =$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \rho \\ \sigma_2 \sigma_1 \rho & \sigma_2^2 \end{pmatrix}, \quad (8)$$

где  $\sigma_1^2$  является дисперсией  $X_1$ ,  $\sigma_2^2$  – дисперсией  $X_2$  и  $\rho$  – коэффициентом корреляции между  $X_1$  и  $X_2$ .

Плотность вероятности  $X_1$  и  $X_2$  равна

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{(x_1 - \mu_1)^2}{\sigma_1^2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right] \right\}. \quad (9)$$

Отметим, что плотность вероятности (6) в  $n$ -мерном евклидовом пространстве постоянна на эллипсоидах

$$(x - \mu)^T \Sigma^{-1}(x - \mu) = c \quad (10)$$

для каждого положительного значения  $c$ . Центром каждого эллипсоида является точка  $\mu$ . Форма и положение эллипсоида определяются значением  $\Sigma$ , а размеры – значением  $c$ .

Рассмотрим подробно двумерный случай плотности вероятности (9). Преобразуем координаты, переходя к стандартным нормальным случайным величинам по формуле

$$\frac{x_i - \mu_i}{\sigma_i} = y_i \quad (i=1, 2),$$

так, чтобы центры линий,

на которых плотность распределения постоянна, находились в начале координат. Эти линии определяются уравнениями

$$\frac{1}{1-\rho^2} (y_1^2 - 2\rho y_1 y_2 + y_2^2) = c. \quad (11)$$

Отрезки, отсекаемые на осях  $y_1$  и  $y_2$  равны между собой. Если  $\rho > 0$ , то большая ось эллипса наклонена под углом  $45^\circ$  к оси  $x$  и ее длина равна  $2\sqrt{c(1+\rho)}$ , а длина малой оси равна  $2\sqrt{c(1-\rho)}$ . Если  $\rho < 0$ , то большая ось эллипса наклонена под углом  $135^\circ$  к оси  $x$  и ее длина равна  $2\sqrt{c(1-\rho)}$ , а длина малой оси равна  $2\sqrt{c(1+\rho)}$  (рис. 1).

Одной из причин изучения многомерных нормальных распределений является то, что частные и условные распределения, полученные из многомерных нормальных распределений, также нормальны. Линейные комбинации нормально распределенных величин также распределены нормально.

Нормальное распределение полностью определяется вектором  $\mu$  среднего значения и кова-

риационной матрицей  $\Sigma$ . Первой статистической проблемой является вопрос о том, как оценить эти параметры по результатам выборки.

Как и в одномерном случае, методом максимального правдоподобия можно показать, что оценкой для  $\mu$  является выборочное среднее, оценкой для  $\Sigma$  - матрица, пропорциональная матрице дисперсий и ковариаций. Выборочная дисперсия равна сумме квадратов отклонений наблюдаемых значений от выборочного среднего, деленной на число наблюдений без единицы; выборочная ковариация определяется подобным образом по результатам взаимных произведений соответствующих отклонений. Выборочная ковариационная матрица является несмещенной оценкой  $\Sigma$ .

Метод наибольшего правдоподобия очень полезен для различных оценок и проблем проверки гипотез, относящихся к многомерным нормальным распределениям. Оценки, полученные по методу наибольшего правдоподобия, или их модификация обычно обладают некоторыми оптимальными свойствами – асимптотически эффективны.

Если  $x_1, \dots, x_N$  образуют выборку из многомерного нормального распределения с параметрами  $\mu$  и  $\Sigma$ , где  $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}$  ( $\rho_{ij} = 1$ ), то оценкой наибольшего правдоподобия для  $\mu$  является

$$\hat{\mu} = \bar{x} = \frac{1}{N} \sum_{\alpha=1}^N x_{\alpha},$$

оценкой наибольшего правдоподобия для  $\sigma_i^2$  является

$$\hat{\sigma}_i^2 = \frac{1}{N} \sum_{\alpha} (x_{i\alpha} - \bar{x}_i)^2,$$

где  $x_{i\alpha}$  есть  $i$ -я компонента  $x_{\alpha}$  и  $\bar{x}_i$  есть  $i$ -я компонента  $\bar{x}$ , а оценка наибольшего правдоподобия для  $\rho_{ij}$  равна

$$\hat{\rho}_{ij} = \frac{\sum_{\alpha} x_{i\alpha} x_{j\alpha} - N \bar{x}_i \bar{x}_j}{\sqrt{\sum_{\alpha} x_{i\alpha}^2 - N \bar{x}_i^2} \sqrt{\sum_{\alpha} x_{j\alpha}^2 - N \bar{x}_j^2}} = \frac{\hat{\sigma}_{ij}}{\hat{\sigma}_i \hat{\sigma}_j}. \quad (12)$$

В одномерном случае выборочное среднее распределено нормально и не зависит от выбо-

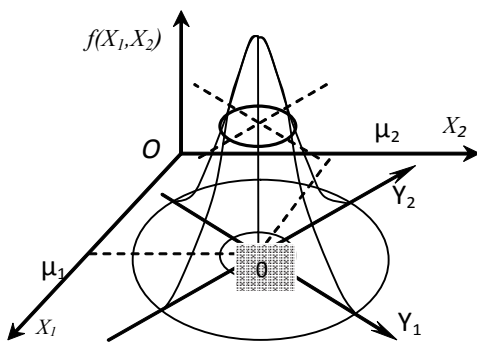


Рис. 1. Поверхность плотности нормального двумерного распределения

рочной дисперсии. Аналогично выборочное среднее  $\bar{x} = (\bar{x}_1 \dots \bar{x}_n)^T$  распределено нормально и не зависит от  $\hat{\Sigma}$ .

Среднее значение выборки объема  $N$  из многомерной нормальной генеральной совокупности  $N(\mu, \Sigma)$  распределено нормально  $N\left(\mu, \frac{1}{N} \Sigma\right)$  и не зависит от  $\hat{\Sigma}$ , являющейся оценкой наибольшего правдоподобия для  $\Sigma$ .

Одна из наиболее важных групп задач одномерной статистики связана с вопросами, касающимися оценки математического ожидания некоторого распределения, дисперсия которого неизвестна. Бывают случаи, когда по выборке хотят решить, равно ли математическое ожидание некоторому наперед заданному числу, или же указать интервал, в котором находится математическое ожидание. В одномерных случаях обычно используется статистика, являющаяся частным от деления разности между выборочным значением  $\bar{x}$  и гипотетическим математическим ожиданием генеральной совокупности на среднее квадратическое отклонение  $s$ . Если выборка произведена из совокупности  $N(\mu, \sigma^2)$ , то величина

$$t = \frac{\bar{x} - \mu}{s/\sqrt{N}} = \sqrt{N}(\bar{x} - \mu)S^{-1} \quad (13)$$

имеет  $t$ -распределение с  $N-1$  степенями свободы, где  $N$  – объем выборки. Основываясь на этом можно построить критерий для проверки гипотезы  $H_0: \mu = \mu_0$ , где  $\mu_0$  – заданное число, или построить доверительный интервал для неизвестного параметра  $\mu$ .

Многомерным аналогом квадрата величины  $t$ , определенной формулой (13), является величина  $T^2$  Хотеллинга:

$$T^2 = N(\bar{x} - \mu)^T S^{-1}(\bar{x} - \mu), \quad (14)$$

где  $\bar{x}$  – вектор среднего значения и  $S$  – ковариационная матрица выборки объема  $N$ . Покажем, как можно использовать эту статистику для проверки гипотезы о векторе среднего значения генеральной совокупности  $\mu$  и для получения доверительных областей для неизвестного параметра  $\mu$ .

Если  $\mu$  есть среднее значение многомерного нормального распределения  $N(\mu, \Sigma)$ , то вероятность получить выборку объема  $N$  со средним  $\bar{x} = (\bar{x}_1 \dots \bar{x}_n)^T$  и выборочной ковариационной матрицей  $S$  такую, что

$$T^2 = N(\bar{x} - \mu)^T S^{-1}(\bar{x} - \mu) \leq T_0^2(\alpha) \quad (15)$$

равна  $1 - \alpha$ .

Здесь

$$T_0^2 = \frac{(N-1)n}{N-n} F_{n, N-n}(\alpha), \quad (16)$$

где  $\alpha$  – уровень значимости,  $F_{n, N-n}(\alpha)$  точка  $F$ -распределения.

Таким образом, если для конкретной выборки произвести вычисления по формуле (15), то утверждение относительно  $\mu$ , выраженное формулой (15), будет справедливо с доверием  $1 - \alpha$ . Совокупность точек, координаты которых удовлетворяют неравенству

$$N(\bar{x} - m)^T S^{-1}(\bar{x} - m) \leq T_0^2(\alpha) \quad (17)$$

образует в  $n$ -мерном пространстве внутренность и границу эллипсоида с центром в точке  $\bar{x} = (\bar{x}_1 \dots \bar{x}_n)^T$ , размеры и форма которого зависят от  $S^{-1}$  и  $\alpha$  (рис. 2),  $\mu$  лежит внутри этого эллипсоида. Эллипсоид (17) является случайным, так как выборка случайна.

Описанная модель многомерного нормального распределения использовалась при оценке уровня профессиональной компетентности студентов квалификации – бакалавр по направлению подготовки 080100.62 Экономика. В качестве компонентов вектора  $X_{ПК}$  оценивались две компетенции: способность на основе описания экономических процессов и явлений строить стандартные теоретические и эконометрические модели, анализировать и содержательно интерпретировать полученные результаты (ПК 4); способность использовать для решения аналитических и исследовательских задач современные технические средства и информационные технологии (ПК 8). Таким образом, вектор  $X_{ПК}$  рассматривается как двумерный вектор компетентности, формирование и оценка уровня которого проводилась при изучении дисциплины “Эконометрика” на третьем курсе обучения. Каждая компетенция оценивалась по таксономии Блума.

В качестве диагностических средств были использованы: тесты достижения для оценки уровня знания и понимания; представление и защи-

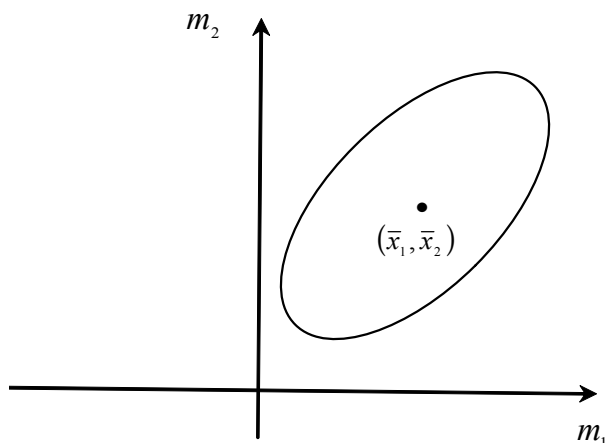


Рис. 2. Граница доверительной области

та итоговой работы по дисциплине, которая включала в себя обработку, анализ и расчет эконометрических показателей с помощью средств ЭВМ (пакет “Анализ данных” Excel), построение эконометрических моделей и их использование для построения прогнозов. При использовании данного средства диагностики оценивались уровни применения, анализа, синтеза и оценки.

Результаты диагностики компетенций в группе из 25 студентов представлены в табл. 1.

Для получения обобщенных показателей уровня компетенций использовались весовые коэффициенты:  $\alpha_1=0,05$ ,  $\alpha_2=0,05$ ,  $\alpha_3=0,1$ ,  $\alpha_4=0,2$ ,  $\alpha_5=0,2$ ,  $\alpha_6=0,4$ .

Результаты расчетов обобщенных показателей каждой компетентности приведены в табл. 2.

Найдем оценки математических ожиданий, дисперсии и коэффициента корреляции, доверительную область для вектора математических ожиданий с надежностью  $\gamma=0,95$ .

Точечные оценки для  $\mu_1$  и  $\mu_2$  – средние арифметические  $\bar{x}_1$  и  $\bar{x}_2$  равны:

$$\bar{x}_1 = \frac{1}{25} (27,3 + 30,1 + \dots + 70,3) = 46,72,$$

$$\bar{x}_2 = \frac{1}{25} (29,05 + 30,2 + \dots + 51,25) = 29,584.$$

Для определения ковариационной матрицы  $S$  выборки перейдем к центрированным случайным величинам  $u_{ij} = x_{ij} - \bar{x}_j$ :

$$U^T = \begin{pmatrix} -19,42 & -16,62 & \dots & 23,58 \\ -10,534 & -9,384 & \dots & 11,666 \end{pmatrix}.$$

$$U^T U = \begin{pmatrix} -19,42 & -16,62 & \dots & 23,58 \\ -10,534 & -9,384 & \dots & 11,666 \\ \dots & \dots & \dots & \dots \\ 23,58 & 11,666 \end{pmatrix} =$$

$$= \begin{pmatrix} 13652,78 & 8762,803 \\ 8762,803 & 6320,309 \end{pmatrix}$$

Несмещенная оценка матрицы  $\Sigma$  - ковариационная матрица  $S$  выборки равна

$$S = \frac{1}{N-1} U^T U = \frac{1}{24} \begin{pmatrix} 13652,78 & 8762,803 \\ 8762,803 & 6320,309 \end{pmatrix} =$$

$$= \begin{pmatrix} 568,87 & 356,12 \\ 365,12 & 263,35 \end{pmatrix}$$

Таким образом, несмещенные оценки дисперсий и средних квадратичных отклонений следующие:

$$s_1^2 = 568,87; \quad s_1 = 23,85; \quad s_2^2 = 263,35; \quad s_2 = 16,23.$$

Вычислим выборочный коэффициент корреляции

$$\hat{\rho}_{ij} = \frac{356,12}{23,85 \cdot 16,23} = 0,92.$$

Обратная матрица к  $S$  будет иметь вид

**Таблица 1.** Результаты формирования профессиональных компетенций в %

Номер студента в списке группы	Компетенция	Знание	Понимание	Применение	Анализ	Синтез	Оценка
1	ПК 4	61	61	58	23	20	17
	ПК 8	63	60	59	25	20	20
2	ПК 4	64	60	59	25	23	21
	ПК 8	64	64	60	26	23	20
3	ПК 4	53	50	43	18	19	16
	ПК 8	57	52	45	17	19	17
4	ПК 4	54	51	50	22	21	20
	ПК 8	57	52	52	23	20	19
5	ПК 4	72	69	69	44	40	43
	ПК 8	70	70	69	43	40	40
6	ПК 4	69	54	53	30	29	19
	ПК 8	68	55	50	28	29	20
7	ПК 4	74	65	44	23	21	18
	ПК 8	72	65	42	20	20	20
8	ПК 4	57	49	43	18	17	17
	ПК 8	56	50	46	18	18	16
9	ПК 4	74	75	67	34	23	21
	ПК 8	75	74	63	31	24	20
10	ПК 4	44	30	27	12	14	12
	ПК 8	49	32	27	11	11	13
11	ПК 4	69	70	68	34	32	18
	ПК 8	64	63	62	31	29	17
12	ПК 4	38	30	28	12	11	12
	ПК 8	40	30	30	11	10	12
13	ПК 4	49	44	35	18	19	19
	ПК 8	45	44	37	18	17	16
14	ПК 4	69	65	63	43	35	27
	ПК 8	71	67	64	42	35	29
15	ПК 4	68	68	64	23	19	20
	ПК 8	67	66	65	22	19	21
16	ПК 4	100	98	88	69	58	63
	ПК 8	72	70	68	59	53	55
17	ПК 4	100	100	97	69	62	65
	ПК 8	65	66	63	44	45	43
18	ПК 4	100	100	95	87	79	72
	ПК 8	73	70	68	56	53	58
19	ПК 4	100	100	97	78	82	68
	ПК 8	89	89	83	62	67	60
20	ПК 4	100	93	85	62	65	54
	ПК 8	85	78	70	44	50	43
21	ПК 4	100	95	95	78	72	76
	ПК 8	92	90	90	67	72	65
22	ПК 4	100	100	98	87	82	72
	ПК 8	64	62	59	32	34	50
23	ПК 4	100	80	78	55	52	65
	ПК 8	73	65	55	44	46	51
24	ПК 4	100	100	98	60	62	78
	ПК 8	95	95	91	53	51	56
25	ПК 4	100	100	99	66	64	61
	ПК 8	96	91	85	44	45	39

Таблица 2. Обобщенные показатели профессиональных компетенций в %

Номер студента в списке группы	ПК 4 ( $x_1$ )	ПК 8 ( $x_2$ )
1	27,3	29,05
2	30,1	30,2
3	23,25	23,95
4	26,85	26,85
5	47,95	46,5
6	30,85	30,55
7	27,35	27,05
8	23,4	23,5
9	33,95	32,75
10	16,4	16,35
11	34,15	31,35
12	15,6	15,5
13	23,15	21,55
14	39,4	40,3
15	29,6	29,75
16	69,3	58,3
17	71,9	47,85
18	81,5	58,95
19	78,9	67
20	65,15	51,15
21	79,65	71,9
22	82,4	45,4
23	64,2	50,8
24	75,4	61,8
25	70,3	51,25

$$S^{-1} = \begin{pmatrix} 0,01596 & -0,02213 \\ -0,02213 & 0,03448 \end{pmatrix}.$$

По формуле (14) находим значение статистики Хотеллинга

$$T^2 = 25 \cdot (46,72 - \mu_1 \quad 29,584 - \mu_2) \begin{pmatrix} 0,01596 & -0,02213 \\ -0,02213 & 0,03448 \end{pmatrix} \begin{pmatrix} 46,72 - \mu_1 \\ 29,584 - \mu_2 \end{pmatrix} = \\ = 0,399(46,72 - \mu_1)^2 - 1,1065(46,72 - \mu_1)(29,584 - \mu_2) + 0,862(29,584 - \mu_2)^2.$$

При уровне значимости  $\alpha = 1 - \gamma = 0,05$  и числе степеней свободы  $\nu_1 = 2$  и  $\nu_2 = 23$  находим по таблице точку  $F$ -распределения -  $F(0,05; 2; 23) = 3,42$ .

Тогда по формуле (16)

$$T_0^2 = \frac{(25 - 1) \cdot 2}{25 - 2} 3,42 = 7,137.$$

Из неравенства (17) получим уравнение эллипса (включая внутреннюю область)

$$0,399(46,72 - \mu_1)^2 - 1,1065(46,72 - \mu_1)(29,584 - \mu_2) + 0,862(29,584 - \mu_2)^2 \leq 7,137,$$

которое определяет доверительную область для вектора средних значений.

Пусть вектор  $X$  подчиняется  $n$ -мерному нормальному распределению  $N_n(\mu, \Sigma)$ , а матрица  $C_{n \times k}$  - матрица ранга  $k$  ( $k \leq n$ ). Тогда вектор

$C^T X$  подчиняется  $k$ -мерному нормальному распределению  $N_k(C^T \mu, C^T \Sigma C)$ , так как линейные комбинации нормально-распределенных величин также распределены нормально.

В этом случае статистика  $T^2$  Хотеллинга имеет вид

$$T^2 = N(C^T \bar{x} - C^T \mu)^T (C^T S C)^{-1} (C^T \bar{x} - C^T \mu) \quad (18)$$

со значением, отвечающим вероятности  $\alpha = 1 - \gamma$

$$T_0^2 = \frac{(N - 1)k}{N - k} F_{k, N - k}(\alpha). \quad (19)$$

Поскольку вектор  $C^T \mu$  содержит  $k$  генеральных средних, то в отличие от уравнения (16) теперь число степеней свободы равно  $k$  и  $N - k$ . В частном случае, когда  $C$  - единичная матрица порядка  $n$ , то уравнение (18) сводится к уравнению (16) и число степеней свободы становится прежним.

Использование линейных комбинаций компонент вектора  $\mu$  позволяет расширить область применения статистики  $T^2$  Хотеллинга при интервальном оценивании в задачах сравнения.

С помощью линейных комбинаций можно найти совместные доверительные интервалы или проверить гипотезу относительно первых ( $k \leq n$ ) средних значений генеральной совокупности. Для этого достаточно принять

$$C^T = \left( \begin{array}{cccccc} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{array} \right) \Bigg\} k. \quad (20)$$

Поскольку  $C^T$  имеет размерность  $k \times n$  и ранг  $k$ , то вектор  $C^T \mu$  имеет размерность  $k \times 1$  и содержит  $k$  генеральных средних.

Чтобы построить доверительный интервал для генерального среднего  $\mu_j$  ( $j = 1, 2, \dots, n$ ), достаточно принять, что  $C_j$  есть  $j$ -й столбец единичной матрицы размерности  $n$ . Тогда  $C_j^T \mu = \mu_j$  и статистика Хотеллинга

$$T^2 = \frac{N [C_j^T (\bar{x} - \mu)]^2}{C_j^T S C_j} \quad (21)$$

имеют распределение, зависящее от числа степеней свободы 1 и  $N-1$ .

Таким образом, согласно уравнению (18) доверительная область для  $C^T \mu$  с надежностью  $\gamma$  будет ограничена поверхностью

$$N(C^T \bar{x} - C^T \mu)^T (C^T S C)^{-1} (C^T \bar{x} - C^T \mu) = \frac{k(N-1)}{N-k} F_{k, N-k}(\alpha). \quad (22)$$

В частности, с надежностью  $\gamma$  доверитель-

ные границы для линейной комбинации  $C_j^T \mu$  определяются как

$$C^T \bar{x} \pm \sqrt{\frac{1}{N} C_j^T S C_j F_{1, N-1}(\alpha)}. \quad (23)$$

По данным табл. 1 с помощью линейных комбинаций найдем с надежностью  $\gamma = 0,95$  интервальные оценки генеральных средних  $\mu_1$  и  $\mu_2$ .

В нашем случае

$$\alpha = 1 - \gamma = 0,05 \quad \nu_1 = 1 \quad \nu_2 = N - 1 = 24.$$

Согласно таблице  $F$ -распределения точка  $F(0,05; 1; 24) = 4,26$ .

Для построения интервальной оценки средней  $\mu_1$  примем  $C_1 = (1, 0)^T$ , так что  $C_1^T \bar{x} = 46,72$ ;  $C_1^T S C_1 = S_1^2 = 568,87$ . Тогда границы доверительного интервала для  $\mu_1$  имеют вид

$$46,72 \pm \sqrt{\frac{1}{25} \cdot 568,87 \cdot 4,26} = 46,72 \pm 9,85.$$

Откуда

$$36,87 \leq \mu_1 \leq 56,57.$$

Для построения интервальной оценки генерального среднего  $\mu_2$  принимаем  $C_2 = (0, 1)^T$ , откуда  $C_2^T \bar{x} = 29,584$ ;  $C_2^T S C_2 = S_2^2 = 263,35$ .

Тогда границы доверительного интервала для  $\mu_2$  имеют вид

$$29,584 \pm \sqrt{\frac{1}{25} \cdot 263,35 \cdot 4,26} = 29,584 \pm 2,06.$$

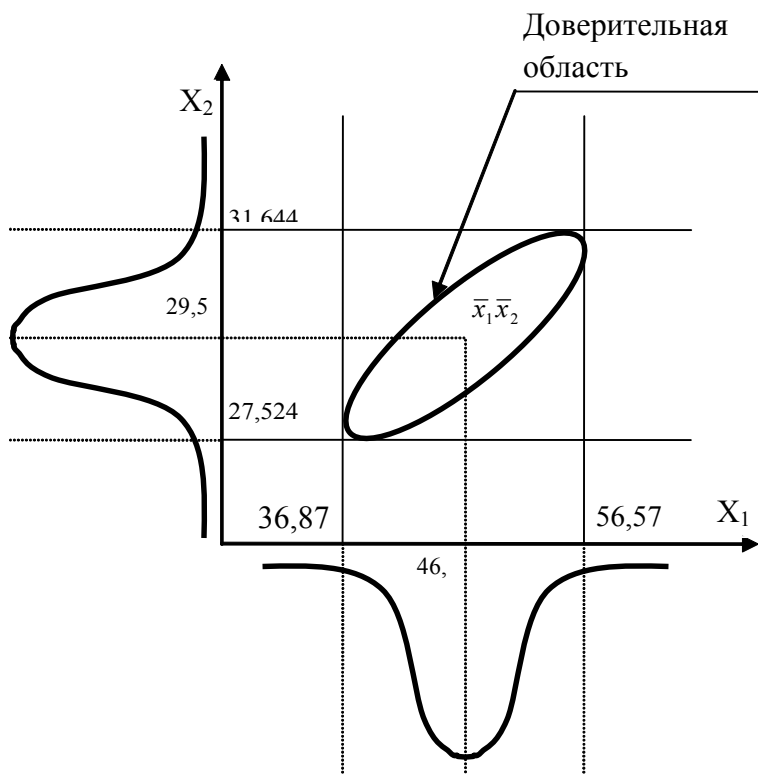


Рис. 3. Доверительная область для двумерного вектора  $\mu$  средних

Откуда

$$27,524 \leq \mu_2 \leq 31,644.$$

Доверительная область для двумерного вектора  $\mu$  средних представлена на рис. 3.

### СПИСОК ЛИТЕРАТУРЫ

1. *Андерсон Т.* Введение в многомерный статистический анализ [пер. с англ.]. М. : ГИФМЛ, 1963. 500 с.
2. *Васильев А.В.* Инженер-эколог: профессия будущего // Известия Самарского научного центра Российской академии наук. 2005. № S2. С. 320-323.
3. *Васильев А.В.* Опыт обучения студентов вузов по проблемам защиты окружающей среды. Безопасность в техносфере. 2010. № 4. С. 55-58.
4. *Васильев А.В.* Обеспечение экологической безопасности в условиях городского округа Тольятти: учебное пособие. Самара: Изд-во Самарского научного центра РАН, 2012. 201 с., ил.
5. *Васильев А.В., Заболотских В.В., Терещенко И.О., Терещенко Ю.П.* Разработка программного обеспечения для комплексной оценки рисков здоровью населения на урбанизированных территориях // В сборнике: ELPIT-2013. Экология и безопасность жизнедеятельности промышленно-транспортных комплексов: сборник трудов IV Международного экологического конгресса (VI Международной научно-технической конференции). Научный редактор: А.В. Васильев. 2013. Т. 4. С. 26-35.
6. *Васильев А.В., Заболотских В.В., Терещенко И.О., Терещенко Ю.П.* Информационно-аналитическая система оценки рисков здоровью населения в условиях урбанизированных территорий // Экология и промышленность России. 2013. № 12. С. 29-31.
7. *Павлова Е.С.* Моделирование систем управления // Актуальные проблемы гуманитарных и естественных наук. 2013. № 2. С. 13-19.
8. *Пивнева С.В.* Реализация методов прогнозирования // Известия Самарского научного центра Российской академии наук. 2008. № S7. С. 171-177.
9. *Ярыгин А.Н., Палферова С.И.* Формирование базовых компетенций студентов при изучении математики в техническом вузе // Вектор науки тольяттинского государственного университета. Серия: Педагогика, психология. 2013. № 1 (12). С. 294-298.
10. *Vasilyev A.V., Zabolotskikh V.V., Vasilyev V.A.* Development of methods for the estimation of impact of physical factors on the health of population // Safety of Technogenic Environment. 2013. № 4. С. 42-45.

### THE APPLICATION OF MULTIVARIATE NORMAL DISTRIBUTION FOR BUILDING A PROBABILISTIC ASSESSMENT OF COMPETENCE

© 2014 N.V. Kalacheva, O.A. Kuznetsova

Togliatti State University

This article describes the multidimensional normal distribution, defined trust the area of vector medium, consider an example of building trust the assessment of competence of students.

*Key words:* multivariate normal distribution, trust region, competence.

---

*Natalya Kolacheva, Candidate of Pedagogical Sciences, Associate Professor of Higher Mathematics and Mathematical Modeling Department. E-mail: nkolacheva@mail.ru*  
*Olga Kuznetsova, Candidate of Pedagogical Sciences, Associate Professor of Higher Mathematics and Mathematical Modelling Department. E-mail: oly-2009@yandex.ru*