Theorem 1.11. *Случайный вектор* (X_1, X_2) *имеет двумерное нормальное распределение с одинаковыми дисперсиями компонент тогда и только тогда, когда случайные величины* $X_1 + X_2$ *и* $X_1 - X_2$ *независимы и распределены нормально.*

Примем эту теорему без доказательства.

Theorem 1.12. *Если случайный вектор* $\vec{X} \sim N_p(\vec{\mu}, \Sigma)$, *то случайная величина*
$Y^2 = (\vec{X} - \vec{\mu})^T \Sigma^{-1} (\vec{X} - \vec{\mu})$ *имеет хи-квадрат распределение с* p *степенями свободы.*

Доказательство. Для любой положительно определенной матрицы Σ^{-1} найдется такая невырожденная матрица C, что

$$C^T \Sigma^{-1} C = E$$

Рассмотрим вектор $Z = C^{-1}(X - \mu)$. В силу теоремы о линейном преобразовании гауссовского случайного вектора $\vec{Z}$ имеет нормальное распределение с вектором средних $\mu = (0, 0, ..., 0))$ и ковариационной матрицей $Cov(Z) = E$. Так как вектор Z представляет собой линейное преобразование вектора X, то по теореме о линейном преобразовании он имеет многомерное гауссовское распределение $N_p(\vec{0}, E)$. А значит из следствия 1 его координаты стандартные нормальные одномерные с.в.

Если теперь рассмотреть квадратичную форму следующего вида:

$$Y^2 = (\vec{X} - \vec{\mu})^T \Sigma^{-1} (\vec{X} - \vec{\mu}) = (CZ)^T \Sigma^{-1} (CZ) = Z^T C^T \Sigma^{-1} C Z$$

Так как $C^T \Sigma^{-1} C = E \Rightarrow Z^T Z = \sum_{j=1}^{p} Z_j$. В силу того что $Z_j \sim N(0; 1)$, то $\sum_{j=1}^{p} Z_j \sim \chi^2_p$. □

2 Оценки параметров многомерного нормального распределения и их некоторые свойства.

2.1 Оценивание вектора параметров средних μ и ковариационной матрицы Σ

. Пусть проводится n независимых экспериментов и получены реализации за p-мерным случайным вектором $X = (X_1, \ldots, X_p)^T \sim N_p(\mu, \Sigma)$. Тогда результаты наблюдений можно записать в виде матрицы :

$$x = \begin{pmatrix} \vec{x_1} \\ \vec{x_2} \\ \vdots \\ \vec{x_p} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{p1} & x_{p2} & \ldots & x_{pn} \end{pmatrix}$$

Каждый столбец матрицы $x = (x_{j\,i})$ gпредставляет собой вектор значений наблюдаемых показателей в i- ом эксперименте, $i = \overline{1, n}$. Например, набор значений показателей анализов по некоторому пациенту. А j- ая строка – это реализация j- го показателя во всех n экспериментах, например, уровень гемоглобина у всех пациентов.

Для оценивания по методу моментов приравняем теоретические моменты к соотвествующим им эмпирическим. Тогда оценка вектора средних μ с.в. X имеет вид:

$$\widehat{\mu} = \overline{x} = \begin{pmatrix} \overline{x_1} \\ \overline{x_2} \\ \vdots \\ \overline{x_p} \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum\limits_{i=1}^{n} x_{1\,i} \\ \frac{1}{n}\sum\limits_{i=1}^{n} x_{2\,i} \\ \vdots \\ \frac{1}{n}\sum\limits_{i=1}^{n} x_{p\,i} \end{pmatrix}$$

Аналоично, чтобы оценить элементы ковариационной матрицы $\sigma_{j\,k}$, воспользуемся оценками для математических ожиданий. Так как $cov(X_j, X_k) = E(X_j\, X_k) - E(X_j)\, E(X_k)$, то

$$\widehat{\sigma}_{j\,k} = \frac{1}{n}\sum_{i=1}^{n} x_{j\,i}\, x_{k\,i} - \overline{x_j}\; \overline{x_k}$$

Отсюда сразу же получаем оценки для дисперсий компонент вектора X:

$$s_j^2 = \frac{1}{n}\sum_{i=1}^{n}(x_{ji})^2 - (\overline{x_j})$$

Тогда в матричной форме оценку ковариационной матрицы можно записать в следующем виде:

$$\widehat{\Sigma} = \frac{1}{n}\, x\, x^T - \bar{x}\, \overline{x}^T.$$

Соответственно оценки для парных коэффициентов корреляции примут вид:

$$\widehat{\rho}_{jk} = \frac{\widehat{\sigma}_{j\,k}}{s_j\, s_k}.$$

Как связаны оценки максимального правдоподобия с оценками, полученными методом моментов?

Theorem 2.1. *Если вектор наблюдаемых признаков имеет многомерное нормальное распределение с вектором средних μ и ковариационной матрицей Σ, то*

- *ОМП совпадают с ОММ*
- *Максимум функции правдоподобия равен*

$$\max_{\mu, \Sigma} L(x) = (2e\pi)^{\frac{-pn}{2}}\, |\widehat{\Sigma}|^{-\frac{n}{2}}$$

Доказательство. Так как функция плотности p - мерного гауссовского распределения имеет вид

$$f(\vec{x}) = \frac{1}{\sqrt{(2\pi)^p||\Sigma||}} \cdot exp\{\frac{-1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})\},$$

то функция правдоподобия выглядит следующим образом:

$$\ln L = \ln \prod_{i=1}^{n} f(\vec{x_i}) = \sum_{i=1}^{n} \ln f(\vec{x_i}) = \tag{7}$$

$$= -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}(\vec{x_i}-\vec{\mu})^T\Sigma^{-1}(\vec{x_i}-\vec{\mu}) \tag{8}$$

$$= -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}\sum_{j,k=1}^{p} q_{j\,k}(\vec{x_{ji}}-\mu_j)(x_{ki}-\mu_k), \tag{9}$$

где q_{jk} – элементы обратной матрицы Σ^{-1}.

1)a) Найдем максимум по координатам вектора μ:

$$\frac{\partial \ln L}{\partial \mu_j} = \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{p} q_{j\,k}(x_{ki}-\mu_k) = \frac{1}{2}\sum_{k=1}^{p} q_{j\,k}\sum_{i=1}^{n}(x_{ki}-\mu_k) = \frac{n}{2}\sum_{k=1}^{p} q_{j\,k}(\overline{x_k}-\mu_k)$$

Эту же систему можно записать в матричной форме:

$$\Sigma^{-1}(\bar{x}-\mu) = \vec{0}. \tag{10}$$

Так как матрица Σ^{-1} невырождена, то уравнение (10) имеет место лишь при $\mu = \bar{x}$.

b) Найдем оценку ковариационной матрицы. Так как матрицы Σ и Σ^{-1} друг друга однозначно определяют, будем искать максимум по элентам $q_{j\,k}$ матрицы Σ^{-1}. найдем производную функции правдоподобия по элементу $q_{j\,k}$. Для этого разложим определитель матрицы Σ^{-1} по j-ой строке:

$$|\Sigma^{-1}| = \sum_{m=1}^{p} q_{jm}(-1)^{j+m}|\left(\Sigma^{-1}\right)_{jm}|,$$

где $|\left(\Sigma^{-1}\right)_{jm}|$ - определитель, который получается после вычеркивания j-ой строки и m- го столбца. Тогда

$$\frac{\partial|\Sigma^{-1}|}{\partial q_{jk}} = (-1)^{j+k}|\left(\Sigma^{-1}\right)_{jk}| = \frac{\sigma_{jk}}{|\Sigma|},$$

так как это почти готовое выражение для элемента матрицы, обратной к Σ^{-1}.

Вернемся теперь к функции правдоподобия и продифференцируем ее по q_{jk} с учетом найденных оценок для μ_j:

$$\frac{\partial|\Sigma^{-1}|}{\partial q_{jk}} = \frac{n}{2}\frac{1}{|\Sigma^{-1}|}\frac{\partial|\Sigma^{-1}|}{\partial q_{jk}} - \frac{1}{2}\sum_{i=1}^{n}\sum_{j,k=1}^{p}(x_{ji}-\overline{x_j})(x_{ki}-\overline{x_k}) = \frac{n}{2}(\sigma_{kj}-\widehat{\sigma}_{kj}) = 0;$$

Отсюда получаем, что оценками для элементов коварианнной матрицы являются как раз оценки, полученные методом моментов.

2) Подставляя, теперь полученные оценки в выражение (7) для функции правдоподобия:

$$\ln L(\widehat{\mu}, \widehat{\Sigma}) = \max_{\mu,\Sigma} \ln L(x) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln|\widehat{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}(\vec{x_i} - \widehat{\mu})^T \widehat{\Sigma}^{-1}(\vec{x_i} - \widehat{\mu})$$

Так как произведение $(\vec{x_i} - \widehat{\mu})^T \widehat{\Sigma}^{-1}(\vec{x_i} - \widehat{\mu})$ является скаляром, то она совпадает со своим следом. Так как оценку для матрицы коварияций можно записать в виде:

$$\frac{1}{n}\sum_{i=1}^{n}(\overline{x}_i - \overline{x})(\overline{x}_i - \overline{x})^T$$

Внутри функции следа матрицы можно переставлтяь местами, поэтому:

$$\sum_{i=1}^{n}(\vec{x_i} - \overline{x})^T \widehat{\Sigma}^{-1}(\vec{x_i} - \overline{x}) = tr\left[(\vec{x_i} - \overline{x})^T \widehat{\Sigma}^{-1}(\vec{x_i} - \overline{x})\right] =$$

$$= tr\left[\sum_{i=1}^{n}\widehat{\Sigma}^{-1}(\vec{x_i} - \overline{x})^T(\vec{x_i} - \overline{x})\right] = tr\left[\widehat{\Sigma}^{-1}\sum_{i=1}^{n}(\vec{x_i} - \overline{x})^T(\vec{x_i} - \overline{x})\right] = tr(nE) = np.$$

Тогда максимум логарифмической функции правдоподобия равен:

$$\max_{\mu,\Sigma} \ln L(x) = -\frac{np}{2}\ln(2\pi) + \frac{n}{2}\ln|\widehat{\Sigma}| - \frac{1}{2}np.$$

Тогда максимум самой функции правдоподобия:

$$\max_{\mu,\Sigma} L(x) = (2e\pi)^{\frac{-pn}{2}}|\widehat{\Sigma}|^{-\frac{n}{2}}.$$

□

Свойства полученных оценок.

1) Оценки для вектора средних и матрицы ковариаций нормальной генеральной совокупности являются состоятельными;

2) Оценка вектора средних является несмещенной с математическим ожиданием $E\widehat{\mu} = \mu$;

3) Оценка ковариационной матрицы асимптотически несмещенная с математическим ожиданием $E\widehat{\sigma}_{jk} = \frac{n}{n-1}\sigma_{jk}$;

4) Оценка для вектора средних является нормальной случайной величиной;

5) Если r_n - оценка коэффициента корреляции по выборке объема n, то величина $\sqrt{n}(r_n - \rho)$ имеет асимптотически нормальное одномерное распределение $N_1(0, \tau)$.