Корреляционный анализ

Парная корреляция

Корреляционный анализ.

- Он используется для установления статистических связей между параметрами оптимизации.
- Для множества объектов матрицу парных корреляций **R** получают в ходе следующих преобразований матриц:

$$X \to Z \to Z^T Z \to \frac{1}{n} Z^T Z \to R$$

- где Z – матрица стандартных значений, а ее элементы получают:

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j}$$

-

Элементы матрицы коэффициентов

- получают по данным матрицы частных корреляций.

$$r_{ij} = -\frac{A_{ij}}{\left(A_{ii}A_{jj}\right)^{\frac{1}{2}}}$$

- Коэффициент множественной корреляции $\mathbf{R_o}$ представляет собой численную характеристику силы связей отклика со всеми факторами.

$$R_o = \left(1 - \left|\frac{R}{R_j}\right|\right)^{\frac{1}{2}}$$

Парная корреляция

- Корреляционный анализ – метод установления статистических связей между выходными параметрами сложной системы.
- Коэффициент парной корреляции является мерой тесноты линейной связи между двумя случайными величинами. В общем случае его величина меняется от 0 до 1. Если коэффициент   = 0, то связь отсутствует, а если 1, то связь линейная.

Определение коэффициента парной корреляции

$$r(x,y)=\sum_{i=1}^{n}\frac{(x_i-x)\cdot(y_i-y)}{S(y)S(x)}$$

$$r_{xy}=\frac{\sum_{i=1}^{N}(x_i-\bar{x})\cdot(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2(y_i-\bar{y})^2}}$$

Упрощение расчетов

$$\sum_{i=1}^{N}(x_i-\bar{x})\cdot(y_i-\bar{y})=\sum_{i=1}^{N}y_i x_i-\frac{\sum_{i=1}^{N}x_i\sum_{i=1}^{N}y_i}{N}$$

$$\sum_{i=1}^{N}(x_i-\bar{x})^2=\sum_{i=1}^{N}x_i^2-\frac{\left(\sum_{i=1}^{N}x_i\right)^2}{N}$$

$$\sum_{i=1}^{N}(y_i-\bar{y})^2=\sum_{i=1}^{N}y_i^2-\frac{\left(\sum_{i=1}^{N}y_i\right)^2}{N}$$

Заполняем таблицу

№ п.п.	Значение выходных параметров		Произведение параметров	Квадрат параметров	
	1-го	2 -го		1 -го	2-го
1	x_1	y_1	x_1y_1	x_1^2	y_1^2
i	x_i	y_i	x_iy_i	x_i^2	y_i^2
N	x_N	y_N	x_Ny_N	x_N^2	y_N^2
Сумма	Σx_N	Σy_N	Σx_Ny_N	Σx_N^2	Σy_N^2

Статистическая значимость коэффициента

- Для этого по выбранному уровню доверительной вероятности α (для обычных технических расчетов α принимается равной 0,95 или 0,99) и числу степеней свободы f=N-2 определяется критическое значение коэффициента парной корреляции ($r_{кр}$).
- Выбор значений $r_{кр}$ производится по таблице , имеющейся в приложении. В случае, если абсолютная величина коэффициента парной корреляции не меньше критического, то линейная связь между параметрами считается статистически значимой.
- В противном случае линейная связь статистически не значима и, следовательно, необходимо переходить более сложным математическим зависимостям.

Построение уравнения регрессии

- Линейное уравнение регрессии имеет вид:

$$y = b_o + b_1 x$$

- Коэффициенты уравнения регрессии можно рассчитать по следующим формулам (за x и y можно принять ту или другую величину):

$$b_o = \frac{\sum_{i=1}^{N} y_i \cdot \sum_{i=1}^{N} x_i^2 - \sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} x_i y_i}{N \cdot \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2} \qquad b_1 = \frac{N \cdot \sum_{i=1}^{N} y_i x_i - \sum_{i=1}^{N} x_i \cdot \sum_{i=1}^{N} y_i}{N \cdot \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i \right)^2}$$

Анализ полученных результатов

- После установления статистически значимой линейной связи необходимо определить параметр, который будет определяться экспериментально, и по которому будет осуществляться оптимизация технологического процесса.
- Оценку линейных связей параметров необходимо осуществлять с учетом абсолютного значения коэффициента парной корреляции.
- При прочих равных условиях предпочтение отдается тем параметрам, для которых метод определения более прост или позволяет проводить измерения с высокой точностью.
- Для упрощения анализ полученных результатов регрессионное уравнение может быть представлено в графическом виде.

Коэффициент парной корреляции

Число степеней свободы	Доверительная вероятность			Число степеней свободы	Доверительная вероятность		
	0,95	0,99	0,999		0,95	0,99	0,999
1.	0,997	1,000	1,000	17.	0,456	0,575	0,693
2.	0,950	0,990	0,999	18.	0,444	0,561	0,679
3.	0,878	0,959	0,992	19.	0,433	0,549	0,665
4.	0,811	0,917	0,974	20.	0,423	0,537	0,652
5.	0,754	0,874	0,951	25.	0,381	0,487	0,597
6.	0,707	0,834	0,925	30.	0,349	0,449	0,554
7.	0,666	0,798	0,898	35.	0,325	0,418	0,519
8.	0,632	0,765	0,872	40.	0,304	0,393	0,490
9.	0,602	0,735	0,847	45.	0,287	0,372	0,465
10.	0,576	0,708	0,823	50.	0,273	0,354	0,443
11.	0,533	0,684	0,801	60.	0,250	0,325	0,408
12.	0,514	0,661	0,780	70.	0,232	0,302	0,380
13.	0,504	0,641	0,760	80.	0,217	0,283	0,357
14.	0,497	0,623	0,742	90.	0,205	0,267	0,337
15.	0,482	0,606	0,725	100.	0,195	0,254	0,321
16.	0,468	0,590	0,708				

Множественная корреляция

Множественная корреляция

- На практике, весьма часто, приходиться анализировать связь между зависимой переменной **у** и группой факторов x_1; x_2; .... x_l. Для оценки используют:
- а) коэффициент множественной корреляции.
- б) коэффициент парциальной корреляции

Коэффициент множественной корреляции

- выражает степень связи между **y** и всей группой независимых переменных

$$r_{y(x_1;x_2;\dots x_l)} = \sqrt{1 - \frac{R}{R_{11}}}$$

- R – матрица парных корреляций
- R_{11} – алгебраическое дополнение определителя **R** к элементу $\mathbf{r_{yy}}$.

- Для ***l*** – независимых переменных и **n** измеренных значений **y:**

$$R = \begin{vmatrix} 1 & r_{yx_1} & r_{yx_2} & \cdots & r_{yx_l} \\ r_{yx_1} & 1 & r_{x_1x_2} & \cdots & r_{x_1x_l} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{yx_l} & r_{x_lx_1} & r_{x_lx_2} & \cdots & 1 \end{vmatrix}$$

Для случая двух независимых переменных

$$r_{yx_1x_2} = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}}$$

$$r_{x_1x_2} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_{1i} \cdot x_{2i} - \bar{x}_1\bar{x}_2}{\sigma_x\sigma_y}$$

Коэффициент парциальной корреляции

- позволяет оценить влияние на **у** каждой из независимых переменных последовательно

$$r_{y\cdots x_1} = -R_{12}\sqrt{R_{11}R_{22}} \qquad R_{11},\ R_{12},\ R_{22}$$

- алгебраические дополнения к элементам $r_{yy}, r_{yx_1}, r_{x_1x_2}$
- Для частных случаев можно воспользоваться формулами

$$r_{yx_1x_2} = \frac{r_{yx_1} - r_{yx_2}\cdot r_{x_1x_2}}{\sqrt{\left(1-r^2_{yx_2}\right)\cdot\left(1-r^2_{x_1x_2}\right)}}$$

$$r_{yx_2x_1} = \frac{r_{yx_2} - r_{yx_1}\cdot r_{x_1x_2}}{\sqrt{\left(1-r^2_{yx_1}\right)\cdot\left(1-r^2_{x_1x_2}\right)}}$$

$$r_{yx_1x_2x_3} = \frac{r_{yx_1x_2} - r_{yx_3x_2}\cdot r_{x_1x_3x_2}}{\sqrt{\left(1-r^2_{yx_3x_2}\right)\cdot\left(1-r^2_{x_1x_3x_2}\right)}}$$

- Другие коэффициенты получают циклической перестановкой индексов

Оценка статистической значимости гипотезы

- Если ($x_1.....x_l$) – факторы, а ($y_1........y_n$) – опыты на точках, то:

$$t = |r| \frac{\sqrt{l \cdot n - 2}}{\sqrt{1 - r^2}}$$

- где $l \cdot n - 2$ это число степеней свободы.
- При наличии линейной связи проводят проверку по критерию Фишера:

$$F_M = \frac{(H - l - 1) \cdot \left(r_{yx_1x_2\cdots\cdots x_l}\right)}{l \cdot \left(1 - r^2_{yx_1x_2\cdots\cdots}\right)}$$

- Можно пользоваться корреляционным отношением:

$$\rho = \frac{\frac{1}{l}\sum_{i=1}^{l} m(y_i - \bar{y})^2}{\frac{1}{l}\sum_{i=1}^{l}\sum_{j=1}^{m}\left(y_{ij} - \bar{y}\right)^2}$$

- где m – количество измерений на одну точку

Пример:

- При анализе связи $\sigma_в(y)$ размеры частиц η – фазы (x_1) и межчастичным расстоянием (x_2) после 5 режимов обработки при испытании трех образцов получено: n=15
- v_1 = числу свободных переменных = *l*.
- v_2= 15-2-1=12

$$F_{табл}^{0,05}(2;12) = 3,89$$

$$t_{табл}^{0,05}(12) = 2,18$$

$$r_{yx_1x_2} = 0,95 \qquad F_м = 54$$

- Парциальные корреляции:

$$r_{yx_1x_2} = 0,42 \qquad t_Г = 1,88$$

$$r_{yx_2x_1} = 0,90 \qquad t_Г = 7,62$$

Вывод: по коэффициенту множественной корреляции оба параметра оказывают влияние на прочностные свойства. По парциальной корреляции влияет только межчастичное расстояние.

КАНОНИЧЕСКАЯ КОРРЕЛЯЦИЯ

Сущность и теоретические основы метода

- Метод канонических корреляций относится к статистическим методам анализа связей между массовыми явлениями и процессами. Если рассматривается зависимость между одним результативным показателем *Y* и одним фактором *X*, то речь идет о парной корреляции. Когда имеется несколько переменных *X* и одна переменная У, проводится множественный корреляционный анализ для установления и измерения степени связи между переменными. Каноническая корреляция — это распространение парной корреляции на случай, когда имеется несколько результативных показателей *Y* и несколько факторов *X.* Основная цель применения этого метода состоит прежде всего в поиске максимальных корреляционных связей между группами исходных переменных: показателями-факторами и результативными качественными показателями. Кроме того, метод канонических корреляций дает возможность сократить объем исходных данных за счет отсева малозначимых факторов.

Матрица значений исходных переменных

Номер наблюдения	X_1	X_2	....	X_g	Y_1	Y_2...	Y_p
1	x_{11}	x_{12}	....	x_{1g}	y_{11}	y_{12}...	y_{1p}
2	x_{21}	x_{22}	....	x_{2g}	y_{21}	y_{22}...	y_{2p}
3	x_{31}	x_{32}	....	x_{3g}	y_{31}	y_{32}...	y_{3p}
...	...	...	...	...	...	...	...
n	x_{n1}	x_{n2}	....	x_{ng}	y_{n1}	y_{n2}...	y_{np}

- x_1, x_2, x_g — переменные факторы;
- У$_1$, Y_2, Y_p — результативные показатели.
- Так как на практике количество факторов значительно пре восходит количество результативных показателей, то будем предполагать, что $p < g$.
- Каноническая корреляция — это корреляция между новыми компонентами (каноническими переменными) U и V:

$$U = a_1 x_1 + a_2 x_2 + \dots + a_q x_q$$

$$V = b_1 y_1 + b_2 y_2 + \dots + b_p y_p$$

Подготовка информации и вычисления канонических корреляций

- По аналогии с парной корреляцией теснота связи между каноническими переменными будет определятся каноническими коэффициентами:

$$r = \frac{\operatorname{cov}(U,V)}{\sqrt{\operatorname{var}(U)\operatorname{var}(V)}}$$

$$\operatorname{cov}(x,y) = \sum_{j=1} \sum_{i=1} \left[x_i - E(x_i)\right] \cdot \left[y_i - E(y_i)\right] \cdot P_{ij}$$

- cov - некоторое число
- E – математическое ожидание величины.
- P_{ij}– совместная вероятность **x** и **y**. $\left(\sum P_{ij} = 1\right)$
- var – дисперсия случайной величины (вспомним 2 случая среднеквадратических отклонения). var(x) = 0.

Вычисление канонических коэффициентов корреляции

$$S=\begin{pmatrix}
\sigma_{x_1x_1} & \sigma_{x_1x_2} & \cdots & \sigma_{x_1x_q} & \sigma_{x_1y_1} & \cdots & \sigma_{x_1y_p} \\
\sigma_{x_2x_1} & \sigma_{x_2x_2} & \cdots & \sigma_{x_2x_q} & \sigma_{x_2y_1} & \cdots & \sigma_{x_2y_p} \\
\sigma_{x_3x_1} & \sigma_{x_3x_2} & \cdots & \sigma_{x_3x_q} & \sigma_{x_3y_1} & \cdots & \sigma_{x_3y_p} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
\sigma_{y_px_1} & \sigma_{y_px_2} & \cdots & \sigma_{y_px_q} & \sigma_{y_py_1} & \cdots & \sigma_{y_py_p}
\end{pmatrix}$$

$$S=\left(\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right)$$

$$\sigma_{x_1y_1}=\operatorname{cov}(x_1,y_1)=\frac{\sum_{i=1}^{n}(x_{1i}-\bar{x}_1)\cdot(y_{1i}-\bar{y}_1)}{n}$$

- S_{12} , S_{21} – матрица взаимодействия **x** и **y** (размерность).
- (S_{12} – **g** x **p и** S_{21} – **p** x **g**)
- S_{21} – результат транспонирования S12 .
- S_{11} – ковариационная матрица исходных переменных, ее размер **g** x **g**.
- S_{22} – ковариантная матрица **y**, **p** x **p**.

Решение задачи

- необходимо решить уравнения: $U = X \cdot A \quad V = Y \cdot B$
- U, V – векторы канонических переменных.
- X, Y – матрицы исходных значений.
- A, B – векторы коэффициентов. $r = \dfrac{\operatorname{cov}(U,V)}{\sqrt{\operatorname{var}(U)\cdot\operatorname{var}(V)}} = \dfrac{\operatorname{cov}(XA,YB)}{\sqrt{\operatorname{var}(XA)\cdot\operatorname{var}(YB)}}$
- Если предположить, что средние

  значения канонических переменных *U и V* равны нулю, а их дисперсии равны единице, σ≈r
- Для упрощения расчетов, считаем, что каждая из переменных имеет единичную дисперсию и нулевое математическое ожидание, следовательно знаменатель этого выражения = 1.

$$r = \frac{A^T S_{12} B}{\sqrt{A^T S_{11} A B^T S_{22} B}}$$

$$\operatorname{var}(XA) = A^T S_{11} A = 1$$

$$\operatorname{var}(YB) = B^T S_{22} B = 1$$

$$r = A^T S_{12} B$$

Находим максимальный коэффициент корреляции

- воспользуемся способом множителей Лагранжа для нахождения условного экстремума (**λ** – множитель Лагранжа), продифференцируем функцию Лагранжа по компонентам векторов А и В и приравняем их к нулю, получим систему:

$$\begin{cases} S_{11}B - \lambda S_{11}A = 0 \\ S_{21}A - \lambda S_{22}B = 0 \end{cases}$$

- Домножим полученные выражения на **λ** и обратную матрицу соответсвенно, получим:

$$\begin{cases} S_{11}\lambda B - \lambda^2 S_{11}A = 0 \\ S_{12}^{-1}S_{21}A = \lambda_{22}B \end{cases} \Rightarrow \quad S_{12}S_{22}^{-1}S_{21}A = \lambda^2 S_{11}$$

$$S_{11}^{-1}S_{12}S_{22}^{-1}S_{22}A = \lambda^2 A$$

$$S_{11}^{-1}S_{12}S_{22}^{-1}S_{22} - \lambda^2 E = 0$$

- Умножив обе части на S_{11}^{-1}, получим
- Рассуждая аналогично

$$(S_{11}^{-1}S_{12}S_{22}^{-1}S_{22} - \lambda^2 E)B = 0$$

Решение последнего уравнения

- Чтобы решить уравнение, необходимо найти характеристические корни и характеристические векторы. Из предположения, что *p* < *g*, вытекает, что размерность вектора *B* меньше размерности вектора *A*. Можно определить вектор А из 1 уравнения системы:

$$A = \frac{S_{11}^{-1} S_{12} B}{\lambda}$$

- Для того чтобы найти компоненты вектора *A*, необходимо определить векторы *B* и *λ*.
- Значения λ^2 находятся как собственные значения матрицы С:
- $$C = S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}$$

1 Можно показать, что $\lambda = r$

$$\begin{cases} S_{11}B - \lambda S_{11} A = 0 \\ S_{21} A - \lambda S_{22} B = 0 \end{cases} \xrightarrow[B^{-1}]{A^{-1}} \begin{cases} A^{-1} S_{11} B = \lambda A^{-1} S_{11} A \\ B^{-1} S_{21} A = \lambda B^{-1} S_{22} B \end{cases} \Longrightarrow \lambda = r \Longrightarrow R = \left(\begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{21} & R_{22} \end{array} \right)$$

1

РАСЧЕТ КАНОНИЧЕСКИХ КОРРЕЛЯЦИЙ 3 ФАКТОРА 2 ПАРАМЕТРА ОПТИМИЗАЦИИ

Пример

Матрица исходных данных

	X_1	X_2	X_3	Y_1	Y_2
1	0,45	170	1860	10,1	23,0
2	0,21	185	1455	8,6	13,2
3	0,18	160	1290	9,5	11,0
4	0,38	175	1710	9,0	9,4
5	0,35	140	1850	7,6	9,2
6	0,50	105	1650	11,5	10,0
7	0,32	90	1935	12,0	19,5
8	0,54	134	1795	6,8	9,0
9	0,47	98	2800	8,5	12,0
10	0,38	100	1635	9,4	10,6

Матрица ковариаций

$$S = \begin{pmatrix} \sigma_{x1x1} & \sigma_{x1x2} & \cdots & \sigma_{x1xq} & \sigma_{x1y1} & \cdots & \sigma_{x1yp} \\ \sigma_{x2x1} & \sigma_{x2x2} & \cdots & \sigma_{x2xq} & \sigma_{x2y1} & \cdots & \sigma_{x2yp} \\ \sigma_{x3x1} & \sigma_{x3x2} & \cdots & \sigma_{x3xq} & \sigma_{x3y1} & \cdots & \sigma_{x3yp} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{ypx1} & \sigma_{ypx2} & \cdots & \sigma_{ypxq} & \sigma_{ypy1} & \cdots & \sigma_{ypyp} \end{pmatrix}$$

$$S = \left(\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array}\right)$$

Матрица парных коэффициентов корреляции

$$R = \left(\begin{array}{ccc|cc} 1 & -0{,}409 & 0{,}520 & -0{,}108 & -0{,}060 \\ -0{,}409 & 1 & 0{,}488 & 0{,}351 & 0{,}047 \\ 0{,}520 & 0{,}488 & 1 & -0{,}117 & 0{,}143 \\ \hline -0{,}108 & -0{,}351 & -0{,}117 & 1 & 0{,}530 \\ -0{,}060 & 0{,}143 & 0{,}143 & 0{,}530 & 1 \end{array}\right)$$

$$r = \frac{\operatorname{cov}(U, V)}{\sqrt{\operatorname{var}(U)\operatorname{var}(V)}}$$

Вспомогательные матрицы

$$R_{11}^{-1}=\begin{pmatrix}1,4375 & 0,2999 & -0,6045\\ 0,2929 & 1,3766 & 0,5195\\ -0,6045 & 0,5195 & 1,5712\end{pmatrix}\qquad R_{22}^{-1}=\begin{pmatrix}1,3890 & -0,7360\\ -0,7360 & 1,3890\end{pmatrix}$$

Для определения собственных значений найдем матицу С

$$C=R_{22}^{-1}R_{21}R_{11}^{-1}R_{12}=\begin{pmatrix}0,4009 & -0,1299\\ -0,2712 & 0,1212\end{pmatrix}$$

Т.к. эта матрица имеет размер 2 х 2, то она будет иметь два собственных значения, и два собственных вектора.

Вспомогательные матрицы

$$R_{11}^{-1} = \begin{pmatrix} 1,4375 & 0,2999 & -0,6045 \\ 0,2929 & 1,3766 & 0,5195 \\ -0,6045 & 0,5195 & 1,5712 \end{pmatrix} \quad R_{22}^{-1} = \begin{pmatrix} 1,3890 & -0,7360 \\ -0,7360 & 1,3890 \end{pmatrix}$$

Для определения собственных значений найдем матицу С

$$C = R_{22}^{-1} R_{21} R_{11}^{-1} R_{12} = \begin{pmatrix} 0,4009 & -0,1299 \\ -0,2712 & 0,1212 \end{pmatrix}$$

Т.к. эта матрица имеет размер 2 х 2, то она будет иметь два собственных значения, и два собственных вектора.

$$\left| C - \lambda^2 \cdot E \right| = 0$$

Собственный вектор: Коэффициент корреляции:

$$\lambda_1^2 = 0{,}491 \quad (\quad -1{,}1470 \qquad 0{,}8411 \quad) \quad r_1 = 0{,}701$$

$$\lambda_2^2 = 0{,}031 \quad (\quad 0{,}2746 \qquad 0{,}8269 \quad) \quad r_2 = 0{,}176$$

$$A = \frac{S_{11}^{-1} S_{12} B}{\lambda}$$

аналогично

$$A_2 = \begin{pmatrix} -1{,}0370 \\ -0{,}3291 \\ 0{,}8674 \end{pmatrix}$$

$$A_1 = \frac{1}{\lambda_1} \begin{pmatrix} 1{,}4375 & 0{,}2929 & -0{,}6045 \\ 0{,}2999 & 1{,}3766 & 0{,}5195 \\ -0{,}6045 & 0{,}5195 & 1{,}5712 \end{pmatrix} \cdot \begin{pmatrix} -0{,}108 & -0{,}060 \\ -0{,}351 & 0{,}047 \\ -0{,}117 & 0{,}143 \end{pmatrix} \times \begin{pmatrix} -1{,}147 \\ 0{,}8411 \end{pmatrix} = \begin{pmatrix} 0{,}1132 \\ 1{,}085 \\ 0{,}8345 \end{pmatrix}$$

Канонические переменные

- И так максимальный коэффициент канонической корреляции 0,71.

$$U_1 = 0{,}1132x_1 + 1{,}085x'_2 + 0{,}8345x'_3$$

$$V_1 = -1{,}147y'_1 + 0{,}8411y'_2$$

$$U_2 = -1{,}037x_1 - 0{,}329x'_2 + 0{,}8674x'_3$$

$$V_2 = 0{,}2746y'_1 + 0{,}827y'_2$$

Проверка статистической значимости

- Проверку статистической значимости коэффициентов проводят по критерию Бартлета:

$$\chi^2 = \left[n - 1 - \frac{1}{2} \cdot \left(p + q + 1 \right) \right] \cdot \ln W_0$$

- И для данного числа степеней свободы сравнивают с табличными:

$$W_0 = \left(1 - \lambda_1\right) \cdot \left(1 - \lambda_2\right) = 0,493221$$

- $$\chi^2 = -\left[10 - 1 - \frac{1}{2} \left(2 + 3 + 1 \right) \right] \ln W_0 = 4.24$$
- для числа степеней свободы (p-1)(g-1)=2, и уровня значимости 0,95.

$$\chi^2_{табл.} = 1,635$$

Получение реальных коэффициентов

- Для того чтобы получить коэффициенты, относящиеся к исходным данным, необходимо помнить, что мы все дисперсии приравняли к 1 =>

$$a_1' \approx \frac{a_1}{\sigma_{x1}} \qquad a_1 = \frac{0,1132}{0,118} = 0,9553$$

Т.о., все остальные коэффициенты будут равны:

a2=0.03004 a3=0.0021

b1=-0.7147 b2=0.1764

Уравнение канонических корреляций будет выглядеть следующим образом:

U1=0,9553x1+0,0304x2+0,0021x3

V1=-0,7147y1+0,176y2

В том случае, если нельзя ограничиться одним выходным параметром, то необходимо перейти к обобщенному параметру оптимизации.

Выводы

- Максимальный коэффициент корреляции 0,701, что означает наличии тесной связи между факторами.
- Сами факторы Y тесно связаны между собой (их корреляция 0,53), также высокую сязь имеют факторы X1 и X3 (0,52)
- Второй коэффициент корреляции не велик и говорит о том, что другие линейные комбинации маловероятны.
- В обеих линейных комбинациях наиболее значима величина X3, коэффициенты при других величинах существенно меняются по величине и меняют знак, т.е. достоверно только влияние фактора X3.
- Для уточнения результатов следует повторить расчеты для других сочетаний факторных и результативных переменных, можно отбрасывать одну из переменных, и рассчитывать новые коэффициенты.
- В случае определения канонических корреляций нет необходимости добиваться независимости исходных переменных.