

Лосева Ольга Владиславовна

кандидат педагогических наук,
доцент кафедры статистики
Всероссийского заочного
финансово-экономического института
(филиал в г. Пензе)
lov191171@yandex.ru

СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ ПЕРСОНАЛА

Аннотация:

В статье описана методика статистического анализа результатов тестирования персонала, направленная на выявление разного рода различий, имеющих управленческое значение. Методика основана на проверке гипотез с помощью соответствующих критериев.

Ключевые слова:

результаты тестирования, проверка гипотез, критерии различия.

Loseva Olga Vladislavovna

Candidate of Pedagogics,
associate professor of statistics,
All-Russian Correspondence
Financial and Economic Institute
(affiliate in Penza)
lov191171@yandex.ru

STATISTICAL ANALYSIS OF RESULTS OF TESTING PERSONNEL

The summary:

The article deals with methodic of statistical analysis of the results of testing personnel, aimed at identifying different kinds of differences that have managerial significance. The methodic is based on the verification of hypothesis testing using the appropriate criteria.

Keywords:

test results, test hypotheses, criteria for distinction.

Тестирование персонала является одной из важных составляющих системы кадрового менеджмента организации. Этому в немалой степени способствует появление на отечественном рынке информационных технологий разнообразных компьютерных программ для психологической диагностики и аттестации работников организации.

Существует ряд должностей, предъявляющих особые требования к личности человека, его свойствам и поведению. В этих случаях диагностика сотрудника является совершенно необходимой для профессионального отбора и подготовки человека к работе. Тестирование персонала позволяет руководителю:

- определить личностные свойства и качества сотрудников для построения более эффективных систем управления;
- своевременно выявить назревание конфликтных ситуаций и определить пути их разрешения;
- планировать вопросы карьерного роста и ротации кадров;
- обосновать кадровые перемещения и т. д.

Основными направлениями тестирования являются социально-психологическая или профессиональная диагностика персонала. Первая направлена, как правило, на определение темперамента, уровня тревожности, мотивации, депрессивных состояний, стрессоустойчивости, стратегий психологической защиты, структуры ценностей, склонностей к алкоголизму и асоциальному поведению и других характеристик личности. Вторая может исследовать оценку успешности работника в различных сферах деятельности, уровень ситуативной надежности личности, профессиональной пригодности и компетентности сотрудника, его лояльности и т. п.

Однако помимо основной цели – оценки сотрудника по ряду параметров – полученные в результате тестирования данные могут использоваться менеджерами по персоналу для различных сопоставлений, в частности:

- сравнительного сопоставления групп работников по какому-то признаку;

– сопоставления результатов исследования работников с первоначальными данными или результатами контрольной группы;

– сопоставления двух и более признаков или двух и более профилей признаков (иерархий), измеренных на одной и той же выборке тестируемых, с целью определения согласованности их изменений;

– сопоставления индивидуальных значений полученных характеристик работника с измененными характеристиками под влиянием контролируемых факторов.

Каждый вид сопоставления представляет собой определенную задачу статистического анализа, решение которой имеет то или иное управленческое значение. Рассмотрим каждую из них подробнее.

Выявление групповых различий работников в уровне исследуемого признака.

Эта задача является актуальной при определении различий по какому-то признаку между двумя группами (командами) работников, выполняющих одну и ту же задачу, например занимающихся разработкой инновационных проектов. В качестве признака в этом случае может выступать уровень компетенции, сплоченности, креативности и т. д. Другим вариантом этой же задачи является упорядочивание работников группы (команды) по определенному параметру (например, уровню толерантности, лидерским качествам), которая сформирована по какому-то другому признаку (например, уровню образования, возрасту, типу темперамента и т. п.).

В зависимости от количества обследуемых групп и их объема для проверки соответствующих гипотез используются различные непараметрические критерии. Основным условием применения критериев является возможность упорядочивания (ранжирования) работников по полученным результатам, которую представляют большинство тестов. В таблице 1 приведены основные задачи данного вида сопоставления, рекомендуемые критерии для независимых групп тестирующихся и ограничения их применения.

Таблица 1 – Выбор критерия для выявления групповых различий, полученных при тестировании по исследуемому признаку

№	Задача	Критерий	Ограничения
1	Определить, существуют ли различия между двумя независимыми группами (командами) работников по какому-то признаку?	U-критерий Манна-Уитни	Границы объемов групп: ($n_1, n_2 \geq 3$) или ($n_1 = 2, n_2 \geq 5$) ($n_1, n_2 \leq 60$)
2	Определить, существуют ли различия между тремя и более независимыми группами (командами) работников по какому-то признаку?	H-критерий Крускала-Уоллеса	Кол-во групп $c = 3$, объемы групп: ($n_1, n_2, n_3 \leq 5$)
		Критерий χ^2 Пирсона	Кол-во групп $c = 4$ и более, объем хотя бы одной группы больше 5. Число степеней свободы $v = c - 1$
3	Упорядочить группы, сформированные по качественному признаку (полу, профессии и т. п.), по какому-то количественному признаку	S-критерий Джонкира (мера выраженности признака)	Кол-во групп $3 \leq c \leq 6$, объемы групп должны совпадать, причем быть не менее 2 и не более 10

Продолжение таблицы 1

4	Упорядочить группы, сформированные по количественному признаку (возрасту, уровню интеллекта), по какому-то другому количественному признаку	S-критерий Джонкира (мера связи количественных признаков)	Кол-во групп $3 \leq s \leq 6$, объемы групп должны совпадать, причем быть не менее 2 и не более 10
---	---	---	--

В основе применения приведенных критериев лежит следующий алгоритм.

1. Формулируются альтернативные гипотезы:

H_0 – выбранные группы не имеют значимых различий по исследуемому признаку;

H_1 – выбранные группы значимо различаются по исследуемому признаку.

2. Определяются эмпирические значения критериев по расчетным формулам [1].

3. По таблице находятся теоретические значения критериев, зависящие от выбранного уровня значимости ($\alpha = 0,05$ или $\alpha = 0,01$), а также от количества и объема групп.

4. Если эмпирическое значение равно или превышает теоретическое значение критерия, то гипотезу H_0 отвергают и принимают гипотезу H_1 .

5. Если гипотезу H_0 не удалось отвергнуть, то это не означает, что различий действительно не существует. Возможно, что применение другого более мощного критерия поможет их выявить, например угловое преобразование Фишера ϕ^* , которое предназначено для сопоставления двух выборок по частоте встречаемости интересующего исследователя эффекта (например, в какой группе работников креативные способности выше заданного значения). Верхнего предела по объему групп у данного критерия не существует, а нижний – не менее пяти наблюдений.

Для расчета критериев рекомендуется применять специализированные статистические пакеты обработки данных – *SPSS*, *STATISTICA*.

Оценка достоверности сдвига в значениях исследуемого признака.

В этом случае важно проследить – произошли ли существенные изменения (сдвиги) в значениях измеряемых показателей в результате действия каких-либо факторов.

В частности, рассмотрим следующие наиболее часто встречающиеся виды сдвигов:

1) временной сдвиг – изучение психологических, профессиональных параметров в зависимости от возраста или стажа работы;

2) сдвиг под влиянием контролируемых воздействий – изменение интересующих параметров работника под влиянием тренинга, реализации программ саморазвития, обучения и т.п.;

3) ситуационный сдвиг – изучение показателей в разных условиях измерения (компьютерный и бумажный варианты тестирования, тестирование в спокойной или напряженной обстановке и т. п.).

В первом случае возможно использование независимых выборок – одновременное тестирование двух и более экспериментальных разновозрастных групп (когда отсутствует возможность длительного проведения исследований).

Во втором случае полученные результаты экспериментальной группы необходимо сравнить с результатами контрольной группы, которая не подвергалась никакому воздействию. Иначе мы не можем с уверенностью утверждать, что достигнутые изменения объясняются именно нашими контролируемыми воздействиями, а не появились в силу действия других неучтенных факторов. При этом экспериментальная и контрольная группы испытуемых являются независимыми выборками.

Для независимых выборок можно применять уже рассмотренный U-критерий Манна-Уитни, а также угловое преобразование Фишера ϕ^* (для зависимых и независимых выборок).

Если измерения проводились на одной и той же группе испытуемых, то необходимо использовать критерии для связанных (зависимых) выборок. Выбор соответствующего критерия определяется видом сдвигов и количеством измерений признака. Критерии оценки статистической достоверности полученных сдвигов приведены в таблице 2.

**Таблица 2 – Критерии оценки достоверности сдвигов
в значениях исследуемого признака**

№	Вид сдвига	Условия применения и ограничения	Критерий
1	Временной, ситуационный	Кол-во измерений – 2, кол-во групп – 1, объем группы $5 \leq n \leq 50$	Т-критерий Вилкоксона (определяет направление и интенсивность сдвига, причем нулевые сдвиги из рассмотрения исключаются)
		Кол-во измерений от 3 до 6, кол-во групп – 1, объем группы $n \leq 12$	L-критерий Пейджа (позволяет определить направление изменений)
		Кол-во измерений – 2, кол-во групп – 2	U-критерий Манна-Уитни, угловое преобразование Фишера ϕ^*
2	Под влиянием контролируемых воздействий	1) при отсутствии контрольной группы	
		Кол-во измерений – 2	Т-критерий Вилкоксона
		Кол-во измерений от 3 до 6, объем группы $n \leq 12$	L-критерий Пейджа
		2) при наличии контрольной группы	
		сопоставление значений «до» и «после» отдельно по двум группам	Т-критерий Вилкоксона (если кол-во измерений – 2); L-критерий Пейджа (если кол-во измерений от 3 до 6)
		Сопоставление сдвигов в двух группах	U-критерий Манна-Уитни, угловое преобразование Фишера ϕ^*

Алгоритм применения критериев для связанных выборок аналогичен ранее рассмотренному алгоритму применения критериев для независимых выборок. Отличие заключается в формулировках альтернативных гипотез. Для Т-критерия Вилкоксона они имеют вид:

H_0 – интенсивность сдвигов значений исследуемого признака в типичном направлении не превосходит интенсивность сдвигов в нетипичном направлении;

H_1 – интенсивность сдвигов значений исследуемого признака в типичном направлении превышает интенсивность сдвигов в нетипичном направлении.

Типичным направлением сдвига считается более часто встречающееся направление (положительное или отрицательное), нетипичным – реже встречающееся.

Если эмпирическое значение критерия равно или превышает теоретическое значение, найденное по таблице, то сдвиг в типичную сторону достоверно преобладает.

Для L-критерия тенденций Пейджа формулировки альтернативных гипотез следующие:

H_0 – увеличение индивидуальных значений при смене условий случайно;

H_1 – увеличение индивидуальных значений при смене условий неслучайно.

Соответственно, если эмпирическое значение критерия больше или равно теоретического, то выявленная тенденция к росту параметра достоверна.

Для ϕ^* -критерия Фишера в общем виде альтернативные гипотезы формулируются так:

H_0 – доля лиц, обладающая интересующим исследователя эффектом, в первой группе не превышает долю таких же лиц во второй группе;

H_1 – доля лиц, обладающая интересующим исследователя эффектом, в первой группе значимо отличается от доли таких же лиц во второй группе.

Если эмпирическое значение критерия равно или превышает теоретическое (табличное) значение критерия, то гипотеза H_0 отвергается.

Определение степени согласованности признаков или объектов.

Здесь предполагается наличие двух или нескольких рядов данных для определения степени согласованности:

- 1) двух или нескольких признаков, измеренных по одной и той же совокупности испытуемых (например, уровень мотивации и уровень средней зарплаты работника);
- 2) двух или нескольких испытуемых по одной и той же совокупности признаков (например, схожи ли работники по типу темперамента или уровню лояльности);
- 3) индивидуальных значений признака со среднегрупповыми;
- 4) двух или нескольких групп одного объема по одной и той же совокупности признаков.

Во всех четырех случаях речь идет о корреляционной связи, которая свидетельствует о том, что изменение одной величины или совокупности величин сопровождается изменениями другой величины или других величин. Содержательное ограничение корреляционного анализа состоит в том, что он позволяет обнаружить наличие связи, но не дает оснований для установления причинно-следственных связей. Иными словами, наличие тесной корреляционной связи между величинами X и Y еще не говорит о том, что Y зависит от X (или наоборот), а может свидетельствовать о наличии третьей латентной величины Z , с которой они обе связаны. Это необходимо учитывать при интерпретации результатов. В качестве меры корреляционной связи в нашем исследовании использовались следующие коэффициенты:

- параметрический коэффициент линейной корреляции r :

$$r = \frac{\sum yx - \frac{\sum y \sum x}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n} \right] \left[\sum y^2 - \frac{(\sum y)^2}{n} \right]}} ; \quad (1)$$

- непараметрический коэффициент множественной конкордации w и его частный случай – коэффициент парной конкордации w' [2]:

$$w = 1 - \frac{\sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n |x_{ik} - x_{jk}|}{nm(m-1)(k-1)} ; \quad (2)$$

$$w' = 1 - \frac{\sum_{j=1}^n |x_{1j} - x_{2j}|}{n(k-1)} . \quad (3)$$

Их характеристика, условия применения приведены в таблице 3.

**Таблица 3 – Сравнительная характеристика
показателей согласованности изменений величин**

№	Характеристика	Коэффициент линейной корреляции r	Коэффициент конкордации (согласованности) w
1	Расчетная формула	1) для двух величин – коэффициент парной корреляции Пирсона (формула (1)); 2) для нескольких величин – базируется на вычислении парных коэффициентов корреляции	1) для двух величин – коэффициент парной конкордации w' (формула (3)); 2) для нескольких величин – коэффициент множественной конкордации (формула (2))
2	Пределы изменения	$-1 \leq r \leq 1$ $R > 0 \rightarrow$ связь прямая $R < 0 \rightarrow$ связь обратная	$0 \leq w \leq 1$
3	Степень тесноты	Оценивается по таблице Чэддока: $0 < r < 0,1 \rightarrow$ отсутствует; $0,1 < r < 0,3 \rightarrow$ слабая; $0,3 < r < 0,5 \rightarrow$ умеренная; $0,5 < r < 0,7 \rightarrow$ заметная; $0,7 < r < 0,9 \rightarrow$ тесная; $0,9 < r < 0,99 \rightarrow$ очень тесная; $r = 1 \rightarrow$ функциональная	$w = 0 \rightarrow$ отсутствует $w < 0,65 \rightarrow$ слабая $0,65 \leq w < 0,75 \rightarrow$ умеренная $0,75 \leq w < 0,85 \rightarrow$ хорошая $0,85 \leq w < 1 \rightarrow$ высокая $w = 1 \rightarrow$ полная согласованность
4	Условия применения	1) величины должны быть измерены в шкале интервалов или шкале отношений; 2) значения всех величин имеют нормальный закон распределения; 3) сопоставляемые ряды данных одинаковы по объему; 4) изменения значений величин носят линейный характер; 5) для определения значимости применяется t -критерий Стьюдента	1) величины должны быть измерены в категориях качества, задаваемых исследователем; 2) подчинение нормальному закону распределения не требуется; 3) сопоставляемые ряды данных одинаковы по объему; 4) изменения значений величин носят произвольный характер; 5) не требует определения значимости, так как находит процент согласованности

Коэффициент конкордации, на наш взгляд, является более универсальным: он не требует нормализации исходных «сырых» тестовых баллов испытуемых, доказательства своей значимости. Его недостатком является трудоемкость расчетов (для множественной конкордации), поскольку он не автоматизирован в специальных прикладных пакетах обработки данных. Но этот недостаток можно устранить, создав специальную программу для определения значения данного коэффициента.

Вместе с тем применение линейного коэффициента корреляции является оправданным, так как процедура стандартизации тестовых баллов, описанная во многих учебных пособиях по психометрике [3], приводит к соблюдению всех условий его применения и расчет данного коэффициента автоматизирован как для парной, так и для множествен-

ной корреляции. Гипотезы, которые мы проверяем с помощью данного коэффициента, формулируются в общем виде следующим образом:

H_0 – корреляция между двумя или более признаками (испытуемыми, группами испытуемых) не отличается от нуля;

H_1 – корреляция между двумя или более признаками (испытуемыми, группами испытуемых) статистически значимо отличается от нуля.

Таким образом, если коэффициент конкордации свидетельствует о достаточно высокой степени согласованности, то необходимо определить и линейный коэффициент корреляции. Это позволит, во-первых, установить направление связи. Во-вторых, значимое отличие коэффициента корреляции от нуля говорит о том, что целесообразно построить линейную модель регрессии, на основе которой получить дополнительные выводы о зависимости между исследуемыми величинами.

Анализ изменчивости признака под влиянием контролируемых факторов.

Для изучения изменчивости признака под влиянием одного или нескольких факторов проводят однофакторный или многофакторный (как правило, двухфакторный) дисперсионный анализ. При дисперсионном анализе устанавливаются причинно-следственные связи, при которых одни переменные рассматриваются как факторы, а другие являются результативными признаками, зависящими от факторов. В этом заключается отличие дисперсионного анализа от корреляционного, в котором просто изучается согласованность изменений величин.

Задача дисперсионного анализа состоит в том, чтобы из общей вариативности признака вычленили:

- 1) вариативность, обусловленную действием одного или двух и более независимых факторов;
- 2) вариативность, обусловленную взаимодействием факторов (при двухфакторном дисперсионном анализе);
- 3) случайную вариативность, обусловленную всеми другими неизвестными факторами.

Другими словами, сущность дисперсионного анализа заключается в расчленении общей дисперсии изучаемого результативного признака на отдельные компоненты, обусловленные влиянием конкретных факторов. Сравнивая компоненты дисперсии друг с другом с помощью F -критерия Фишера, можно определить, какая доля общей вариативности результативного признака обусловлена действием факторов.

Однофакторный дисперсионный анализ базируется на расчете дисперсий трех видов:

- общей дисперсии, вычисленной по всей совокупности экспериментальных данных;
- межгрупповой дисперсии, характеризующей вариативность групповых средних под влиянием признака-фактора;
- внутригрупповой дисперсии, характеризующей случайную (остаточную) вариацию результативного признака под влиянием неучтенных факторов.

Данный метод может рассматриваться в двух вариантах:

- 1) для дисперсионных таблиц, в которых влиянию разных градаций фактора подвергнуты разные выборки тестируемых (анализ для несвязанных выборок);
- 2) для дисперсионных таблиц, представляющих данные одной и той же выборки, подвергнутых влиянию разных градаций фактора (анализ для связанных выборок).

Дисперсионная таблица должна быть равномерной, т. е. каждой градации фактора должно соответствовать одинаковое число наблюдений, причем не менее двух. Это поз-

воляет обойти требование равенства дисперсий для каждой градации фактора. Если все же количество наблюдений разное, то «лишние» наблюдения необходимо отсечь путем случайного отбора.

Число градаций фактора должно быть не менее трех. В противном случае проще использовать непараметрические критерии для анализа исследуемой зависимости, рассмотренные выше. Для первого варианта подойдет Н-критерий Крускала-Уоллеса, для второго – L-критерий Пейджа. Однако, если объемы выборок неограниченны, то преимущество остается за дисперсионным анализом.

Алгоритм проведения однофакторного дисперсионного анализа для несвязанных выборок представляет собой следующую последовательность этапов.

1. Составляется дисперсионная таблица, в столбцах которой находятся группы испытуемых, получивших ту или иную градацию факторного признака, а в ячейках – соответствующие значения результативного признака.

2. Обосновывается нормальность распределения результативного признака.

Рассмотрим достаточно простую методику проверки нормальности распределения результативного признака, предложенную Е.И. Пустыльниковым [4]:

а) рассчитываются эмпирические значения коэффициентов асимметрии и эксцесса по формулам:

$$A_{\text{эмп}} = \frac{\sum (y - \bar{y})^3}{n \sigma^3}; \quad (4)$$

$$E_{\text{эмп}} = \frac{\sum (y - \bar{y})^4}{n \sigma^4} - 3; \quad (5)$$

б) определяются теоретические значения коэффициентов асимметрии и эксцесса по формулам:

$$A = 3 \sqrt{\frac{6(n-1)}{(n+1)(n+3)}}; \quad (6)$$

$$E = 5 \sqrt{\frac{24n(n-2)(n-3)}{(n+1)^2(n+3)(n+5)}}; \quad (7)$$

в) если $|A_{\text{эмп}}| < A_{\text{кр}}$ и $|E_{\text{эмп}}| < E_{\text{кр}}$, то распределение результативного признака не отличается от нормального.

Данный метод эффективен, если объем результативного признака не превышает 40. В противном случае лучше использовать λ-критерий Колмогорова, расчет которого автоматизирован.

3. Выдвигаются альтернативные гипотезы:

H_0 – различия в градациях фактора являются не более выраженными, чем случайные различия внутри каждой группы;

H_1 – различия в градациях фактора являются более выраженными, чем случайные различия внутри каждой группы.

4. Осуществляется дисперсионный анализ (например, с помощью пакета «Анализ данных» в Excel).

При проведении однофакторного дисперсионного анализа для связанных выборок исследуется влияние разных градаций фактора на одну и ту же группу испытуемых. Если в предыдущем случае различия между градациями фактора отражали индивидуальные

различия (испытываемые с разным уровнем степени сосредоточения), то теперь различия между градациями фактора могут проявиться только вопреки индивидуальным различиям. Сам фактор индивидуальных различий может оказаться более значимым, чем изменения экспериментальных условий.

Двухфакторный дисперсионный анализ позволяет оценить не только влияние каждого фактора в отдельности, но и их взаимодействие. Может оказаться, например, что фактор X_1 значимо действует на результирующий признак только при небольших значениях другого фактора X_2 . Или один из факторов может усилить действие другого.

Сама дисперсионная таблица должна быть симметрична: каждой градации фактора X_1 должно соответствовать одинаковое число градаций фактора X_2 , причем количество наблюдений должно быть одинаковым для каждой группы.

Если мы измеряем одни и те же показатели у одних и тех же испытуемых несколько раз (в разных условиях, после обучения, с помощью параллельных форм диагностирующих методик), то нам необходимо провести множественное сравнение показателей при переходе от одного условия к другому. Непараметрические критерии не позволяют это сделать, так как необходимо определить тенденцию изменения признака под влиянием двух факторов одновременно. Только двухфакторный дисперсионный анализ для связанных выборок способен справиться с этой задачей. При этом последовательность действий аналогична анализу для несвязанных выборок.

Таким образом, предложенная методика статистического анализа полученных результатов психологического или профессионального тестирования персонала является инструментом для решения дополнительных задач кадрового менеджмента, в частности:

- выявления статистически значимых различий между отделами, группами по диагностируемым параметрам;
- выявления коррелируемости (согласованности) изменений различных характеристик сотрудников;
- определение изменения заданного параметра при контролируемом влиянии другого фактора или совокупности факторов, например при оценке эффективности тренинга или других управляющих воздействий.

Ссылки:

1. Носс И.Н. Психодиагностика. Тест, психометрия, эксперимент. М., 1999. 320 с.
2. Статистический анализ многомерных объектов произвольной природы / В.И. Васильев, В.В. Красильников, С.И. Плакий и др. М., 2004. 382 с.
3. Бурлачук Л. Ф. Психодиагностика : учебник для вузов. СПб., 2006. 351 с.
4. Ермолаев О.Ю. Математическая статистика для психологов. М., 2003. 336 с.

References (transliterated):

1. Noss I.N. Psihodiagnostika. Test, psihometriya, eksperiment. M., 1999. 320 p.
2. Statisticheskiy analiz mnogomernykh objektov proizvol'noy prirody / V.I. Vasil'ev, V.V. Krasil'nikov, S.I. Plaksiy, et al. M., 2004. 382 p.
3. Burlachuk L. F. Psihodiagnostika : university textbook. SPb., 2006. 351 p.
4. Ermolaev O.Y. Matematicheskaya statistika dlya psihologov. M., 2003. 336 p.