

Метод расчета оптимальной надежности испытаний при выпуске качественной продукции

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ»

Рассмотрены вопросы практического применения метода расчета количественных характеристик контроля качества испытаний на наукоемком высокотехнологическом предприятии. Предложены различные процедуры определения коэффициента надежности и проведена сравнительная оценка различных методов определения количественной оценки теста. Предложен оптимальный способ расчета коэффициента надежности проведенного контроля качества продукции. Выбранный метод отличается от существующих тем, что позволяет рассчитать количественную оценку полноты контроля качества выпускаемой продукции. Это дает достоверную оценку проверке и при этом обеспечивает показатель надежности в заданном интервале.

Ключевые слова: методы, тестовые результаты, гомогенность испытаний, коэффициент надежности, формула Кюдера – Ричардсона, средняя корреляция.

1. Постановка проблемы, актуальность и современное состояние проблемы

Для современных предприятий особо острым стал вопрос получения сертификата качества на выпускаемую продукцию, что необходимо для дальнейшего выхода на международный рынок. Политика в области качества может быть сформулирована в виде принципа деятельности или долгосрочной цели и включать в себя улучшение экономического положения предприятия; расширение или завоевание новых рынков сбыта; достижение технического уровня продукции, превышающего уровень ведущих фирм; ориентацию на удовлетворение требований потребителей определенных отраслей или определенных регионов; освоение изделий, функциональные возможности которых реализуются на новых принципах; улучшение важнейших показателей качества продукции; снижение уровня дефектности изготавливаемой продукции; увеличение сроков гарантии на продукцию; развитие сервиса. Существующие методы проверки качества изделий зачастую не охватывают полный спектр исследуемых параметров или вообще не являются адекватными.

2. Анализ последних исследований

В настоящее время не существует однозначного ответа на вопрос об объеме обследуемой выборки. Исследование большого числа реальных задач многомерного анализа данных показывает, что основная часть процесса проверки качества содержит от 30 до 200 испытаний. В зависимости от объема выборки используют различные статистические критерии [1].

Параметры диагностических моделей (отбора информативных пунктов испытаний, нахождение весовых коэффициентов) определяют с помощью методов многомерного статистического анализа. Однако, как было отмечено выше, специфика измерений (высокая размерность, нормальный и качественный характер исходных признаков) накладывает свой отпечаток на применение этих методов. Часто нецелесообразно, а то и невозможно использовать классические версии того или иного метода. Нередко проверяющему имеет смысл ограничиться

упрощенными моделями указанных методов и остановиться на их реализации в форме отбора признаков в самой приблизительной оценке весомых коэффициентов для пунктов испытаний, вошедших в правило проверки результирующего показателя. В табл. 1 приведены отдельные методы определения параметров диагностирующих моделей.

Таблица 1

Методы определения параметров диагностических моделей

Наименование метода или модели	Смысловая нацеленность критерия информативности данных	Используемые предложения о структуре данных
<i>Методы, основанные на критерии автоинформативности</i>		
Метод главных компонент	Нахождение в пространстве исходных признаков новой координатной оси с максимальной дисперсией	Основная часть исходных признаков согласованно отражает требуемый диагностический конструкторов
Факторный анализ	Максимизация точности воспроизведения корреляционных связей между исходными признаками с помощью новых вспомогательных переменных	Одна или несколько групп взаимосвязанных признаков отражают один или несколько диагностических конструкторов
Метод контрастных групп	Исключение признаков из «чернового» варианта диагностической модели, уменьшающих вытянутость гиперэллипсоида рассеивания	Большая часть признаков, вошедших в «черновой» вариант диагностической модели, подобрана правильно
<i>Методы, использующие внешний критерий</i>		
Регрессионный анализ	Минимизация ошибки восстановления значений критериального показателя по значениям исходных признаков	Значение критериального показателя, выражены количественной переменной. Они линейно связаны с диагностическими признаками
Дискриминантный анализ	Минимизация вероятности ошибочного отнесения объектов к заданным классам	Критериальный показатель является номинальной переменной.
Типологический подход	Минимизация ошибки восстановления критериального показателя для отдельных подгрупп объектов (типов)	Существует нелинейная сложная связь значений критериального показателя с исходными признаками для целевой выборки объектов

При контроле качества независимо от совершенства применяемых для этого методик предлагается прежде всего отделение хороших изделий от плохих. Естественно, что качество изделия не повышается путем отделения некачественных. Поэтому современные предприятия сосредотачивают внимание не на выявлении брака, а на его предупреждении, на тщательном контроле производственного процесса и осуществляют свою деятельность в соответствии с концепцией «регулирование качества».

Большую роль в обеспечении качества продукции играют статистические методы.

Целью методов статистического контроля является исключение случайных изменений качества продукции. Такие изменения вызываются конкретными принципами, которые нужно установить и устранить. Статистические методы контроля качества подразделяют на следующие:

- статистический приемочный контроль по альтернативному признаку;
- выборочный приемочный контроль по варьирующим характеристикам качества;
- стандарты статистического приемочного контроля;
- система экономических планов;
- планы непрерывного выборочного контроля;
- методы статистического регулирования технологических процессов.

3. Цель статьи

Целью предлагаемой работы являются исследование существующих процедур проверки подсистемы испытаний выпускаемой продукции и выделение оптимального расчета количественной оценки, которая позволит адекватно выполнить такие функции, как выбор тестовых заданий, накопление статистической информации о результатах контроля и коррекцию базы тестовых заданий (путем предоставления рекомендаций, указаний и т.д.). Благодаря процедуре контроля качества испытаний определяют пригодность полученных результатов проверки для той цели, ради чего проводилась исследование выпускаемой продукции, а также рассчитывают коэффициент надежности проведенных исследований.

4. Изложение основных материалов исследования

Система проверки качества испытания означает, что в испытании собраны такие тесты (задания), которые обладают системообразующими свойствами. Любое испытание состоит из тестовых заданий, которые представляют собой не совокупность произвольно объединенных заданий, а именно систему. Испытание как система обладает составом, целостностью и структурой.

Качество испытания сводится к определению меры его надежности и вопросов валидности полученных результатов. Как и объективным, так и качественным можно назвать только тот метод измерения, который обоснован научно и способен дать требуемые результаты [2].

Таблица 2

Показатели, влияющие на качество испытания

<p>Валидность Степень пригодности результатов испытания для той цели, ради чего его и проводили</p>	<p>Надежность Степень устойчивости результатов испытания пройденных каждым проверяемым продуктом к воздействию случайных факторов</p>	<p>Эффективность (оптимальность) Наилучшее из возможных вариантов измерения проверяемой продукции с меньшим количеством заданий</p>
<ul style="list-style-type: none"> ▪ от качества заданий; ▪ их числа; ▪ степени полноты и глубины измерения всех параметров; ▪ баланса и распределения заданий по трудности; ▪ метода отбора заданий в испытании из общей базы заданий, интерпретации результатов испытаний; ▪ организации сбора данных; подбора выборочной совокупности испытуемых и ряда других факторов 	<ul style="list-style-type: none"> ▪ количество заданий; ▪ содержание заданий (чем разнообразнее измеряемые параметры, тем выше надежность); ▪ гетерогенность группы проверяемых (чем больше в группе тех, кто соответствует требованиям, тем выше надежность); ▪ корреляция между результатами выполнения заданий (чем больше позитивных корреляций, тем выше надежность); ▪ характеристика заданий (если испытание имеет большое количество заданий с высоким коэффициентом дискриминативности и среднем уровнем сложности, то его надежность высока) 	<ul style="list-style-type: none"> ▪ количество заданий; ▪ качество выпускаемой продукции; ▪ принцип подбора заданий

Для иллюстрации рассмотренных методов (табл. 1) воспользуемся исходными данными, приведенными в табл. 3.

Таблица 3

Распределение тестовых баллов

282

Номер проверяемой продукции, n	Количество тестов в испытании, k																						Число пройденных тестов, $\sum_{j=1}^k x_{ij}$	Достижение, $\frac{\sum_{j=1}^k x_{ij}}{k}$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22		
1	1	1	1	1	0	1	0	1	1	1	0	0	1	1	1	1	0	1	0	0	0	0	13	0,59
2	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	4	0,18
3	1	1	1	1	0	1	1	1	1	0	0	0	1	1	1	1	0	1	0	1	0	0	14	0,64
4	1	1	1	1	0	1	1	1	1	0	0	0	0	1	1	1	0	1	0	1	0	0	13	0,59
5	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	0	0	14	0,63
6	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	18	0,82
7	0	1	1	1	0	1	1	1	1	1	0	0	1	0	1	1	0	1	1	0	0	0	13	0,59
8	0	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	0	0	0	1	0	1	12	0,55
9	1	1	1	0	0	1	0	1	1	1	1	0	1	0	0	0	0	1	0	1	0	0	11	0,5
10	1	1	1	0	0	1	1	1	1	0	0	0	1	1	0	1	0	1	0	1	0	0	12	0,55
11	1	1	1	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	1	0	0	8	0,36
12	1	0	1	1	0	0	0	1	1	1	0	0	1	1	0	0	0	0	0	1	0	0	9	0,41
$\sum_{i=1}^n x_{ij}$	9	11	11	7	0	8	7	11	12	7	3	2	10	8	6	9	0	8	1	9	0	2	$\sum_{i=1}^n \sum_{j=1}^k x_{ij} = 141$	$\sum_{i=1}^n \frac{\sum_{j=1}^k x_{ij}}{k} = 6,41$
$\frac{\sum_{i=1}^n x_{ij}}{n}$	0,8	0,9	0,9	0,6	0,0	0,7	0,6	0,9	1,0	0,6	0,3	0,2	0,8	0,7	0,5	0,8	0,0	0,7	0,1	0,8	0,0	0,2	$\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n} = 11,75$	
σ_j^2	0,19	0,08	0,08	0,24	0,00	0,22	0,24	0,08	0,00	0,24	0,19	0,14	0,14	0,22	0,25	0,19	0,00	0,22	0,08	0,19	0,00	0,14	$\sum_{j=1}^k \sigma_j^2 = 3,118$	

4.1. Метод определения валидности тестовых результатов

Существуют следующие виды валидности:

- содержательная (очевидная);
- критериальная (эмпирическая или прогностическая);
- конструктивной (концептуальная);
- дискриминативная (конкурентная).

Содержательная валидность характеризуется корреляцией каждого тестового задания с суммой тестовых баллов испытуемой продукции. Коэффициент корреляции показывает, насколько валидным является данное задание теста. При $r_{xy} > 0,3$ задание тестирования считается валидным. Например, рассчитаем коэффициент корреляции по формуле Пирсона (1) для первого тестового задания, используя вспомогательную табл. 4.

Таблица 4

Исходные данные для расчета коэффициента корреляции первого задания

Номер проверяемой продукции, n	Баллы в первом задании x_{i1}	Сумма баллов по всем заданиям каждой продукцией, которая прошла проверку, $\sum_{j=1}^k x_{ij}$	$x_{i1} \cdot \sum_{j=1}^k x_{ij}$	x_{i1}^2	$\left(\sum_{j=1}^k x_{ij}\right)^2$
1	1	13	13	1	169
2	0	4	0	0	16
3	1	14	14	1	196
4	1	13	13	1	169
5	1	14	14	1	196
6	1	18	18	1	324
7	0	13	0	0	169
8	0	12	0	0	144
9	1	11	11	1	121
10	1	12	12	1	144
11	1	8	8	1	64
12	1	9	9	1	81
Сумма	$\sum_{i=1}^n x_{i1} = 9$	$\sum_{i=1}^n \sum_{j=1}^k x_{ij} = 141$	$\sum_{i=1}^n x_{i1} \cdot \sum_{j=1}^k x_{ij} = 112$	$\sum_{i=1}^n x_{i1}^2 = 9$	$\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij}\right)^2 = 1793$
среднее	$\frac{\sum_{i=1}^n x_{i1}}{n} = 0,75$	$\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n} = 11,75$	$\frac{\sum_{i=1}^n \left(x_{i1} \cdot \sum_{j=1}^k x_{ij}\right)}{n} = 9,333333$	$\frac{\sum_{i=1}^n x_{i1}^2}{n} = 0,75$	$\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij}\right)^2}{n} = 149,416667$

$$r_1 = \frac{\sum_{i=1}^n \left(x_{i1} \cdot \sum_{j=1}^k x_{ij}\right) - \frac{\sum_{i=1}^n x_{i1} \cdot \sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n}}{\sqrt{\left(\frac{\sum_{i=1}^n x_{i1}^2}{n} - \left(\frac{\sum_{i=1}^n x_{i1}}{n}\right)^2\right) \cdot \left(\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij}\right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n}\right)^2\right)}} = \frac{9,333333 - 0,75 \cdot 11,75}{\sqrt{0,75 - 0,75^2} \cdot \sqrt{149,42 - 11,75^2}} = 0,357. \tag{1}$$

Коэффициенты корреляции по каждому заданию занесены в табл. 5

Таблица 5

Коэффициенты корреляции по каждому тестовому заданию

Номер теста в испытании, k	Коэффициент корреляции r_j	Средний коэффициент корреляции $\overline{r_j}$	$\overline{r_j^2}$	Смещенная оценка дисперсии σ_j^2	$\overline{r_j^2} \cdot \sigma_j^2$
1	0,357	0,01623	0,0002633	0,19	4,937E-05
2	0,246	0,01118	0,000125	0,08	9,551E-06
3	0,693	0,0315	0,0009923	0,08	7,58E-05
4	0,489	0,02223	0,0004941	0,24	0,0001201
5	0	0	0	0	0
6	0,734	0,03336	0,0011131	0,22	0,0002474
7	0,690	0,03136	0,0009837	0,24	0,0002391
8	0,693	0,0315	0,0009923	0,08	7,58E-05
9	0	0	0	0	0
10	0,389	0,01768	0,0003126	0,24	7,599E-05
11	0,443	0,02014	0,0004055	0,19	7,603E-05
12	0,564	0,02564	0,0006572	0,14	9,128E-05
13	0,431	0,01959	0,0003838	0,14	5,331E-05
14	0,262	0,01191	0,0001418	0,22	3,152E-05
15	0,717	0,03259	0,0010622	0,25	0,0002655
16	0,414	0,01882	0,0003541	0,19	6,64E-05
17	0	0	0	0	0
18	0,734	0,03336	0,0011131	0,22	0,0002474
19	0,112	0,00509	2,592E-05	0,08	1,98E-06
20	-0,271	-0,0123	0,0001517	0,19	2,845E-05
21	0	0	0	0	0
22	0,431	0,01959	0,0003838	0,14	5,331E-05
Сумма	$\sum_{j=1}^k r_j =$ =8,128	$\sum_{j=1}^k \overline{r_j} =$ =0,36945	$\sum_{j=1}^k \overline{r_j^2} =$ =0,0099556	$\sum_{j=1}^k \sigma_j^2 =$ =3,12	$\sum_{j=1}^k \overline{r_j^2} \cdot \sigma_j^2 =$ =0,0018082

Коэффициент корреляции в данном случае показывает силу (интенсивность) линейной связи. В случае положительной корреляции можно говорить о линейной зависимости между тестами испытания. Отрицательная – свидетельствует об обратной линейной связи. В случае нулевой корреляции такого рода зависимость отсутствует.

Выделенные коэффициенты корреляции показывают, что данные тестовые задания не являются валидными и требуют коррекции.

Критериальную валидность определяет дифференцирующая способность испытания, которую рассчитывают по формуле (2), используя табл. 4.

$$ДС = ДГ_B - ДГ_H = \frac{\sum_{j=1}^k x_{ij}}{k_{max}} - \frac{\sum_{j=1}^k x_{ij}}{k_{min}} = 0,81818182 - 0,18181818 = 0,63636364, \quad (2)$$

где $ДГ_B$ ($ДГ_H$) – максимальный (минимальный) результат, который достигла продукция, прошедшая проверку.

Дифференцирующая способность теста считается удовлетворительной, если показатель $ДС > 0,15$.

Дискриминативная валидность определяет способность отдельных заданий и испытания в целом дифференцировать проверяемую продукцию относительно минимального и максимального результатов. Дискриминативность теста (δ) оценивают количеством совпадений результатов тестирования в группе и рассчитывают по формуле (3) с помощью табл. 4.

$$\delta = \frac{n^2 - \sum_{i=1}^n f_i^2}{n^2 - \frac{n^2}{k+1}} = \frac{12^2 - (3^2 + 1^2 + 2^2 + 1^2 + 2^2 + 1^2 + 1^2 + 1^2)}{12^2 - \frac{12^2}{22+1}} = \frac{122}{137,74} = 0,8857, \quad (3)$$

где f_i – число продукции, прошедшей проверку и показавшей при тестировании i -й результат (одна проверяемая продукция прошла 18 тестов в испытании, две – 14 испытаний, три – 13 испытаний, две – 12 испытаний, одна – 11 испытаний, одна – 9 испытаний, одна – 8 испытаний и одна – 4 испытания).

Приемлемое пороговое значение $\delta = 0,5$

4.2. Метод определения надежности тестовых результатов

Существует несколько методов определения количественной оценки прохождения тестовых проверок, определяющейся коэффициентом надежности, область допустимых значений которого находится в интервале [0;1].

1. Метод половинного деления

Приведем пример отдельного коррелирования результатов выполнения централизованного теста (табл. 6).

Таблица 6

Исходные данные для расчета надежности методом половинного деления

Номер проверяемой продукции, n	Сумма баллов в нечетных заданиях, $\sum_{j=1}^k x_{i(2 \cdot j-1)}$	Сумма баллов в четных заданиях, $\sum_{j=1}^k x_{i2 \cdot j}$	$\sum_{j=1}^k x_{i(2 \cdot j-1)} \cdot \sum_{j=1}^k x_{i2 \cdot j}$	$\left(\sum_{j=1}^k x_{i(2 \cdot j-1)} \right)^2$	$\left(\sum_{j=1}^k x_{i2 \cdot j} \right)^2$
1	5	8	40	25	64
2	1	3	3	1	9
3	6	8	48	36	64
4	5	8	40	25	64
5	7	7	49	49	49
6	7	11	77	49	121
7	6	7	42	36	49
8	4	8	32	16	64
9	5	6	30	25	36
10	5	7	35	25	49
11	4	4	16	16	16
12	4	5	20	16	25
Сумма	$\sum_{i=1}^n \sum_{j=1}^k x_{i(2 \cdot j-1)} =$ =59	$\sum_{i=1}^n \sum_{j=1}^k x_{i2 \cdot j} =$ =82	$\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2 \cdot j-1)} \cdot \sum_{j=1}^k x_{i2 \cdot j} \right) =$ =432	$\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2 \cdot j-1)} \right)^2 =$ =319	$\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2 \cdot j} \right)^2 =$ =610
среднее	$\frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2 \cdot j-1)}}{n} =$ n =4,9167	$\frac{\sum_{i=1}^n \sum_{j=1}^k x_{i2 \cdot j}}{n} =$ n =6,8333	$\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2 \cdot j-1)} \cdot \sum_{j=1}^k x_{i2 \cdot j} \right)}{n} =$ n =36	$\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2 \cdot j-1)} \right)^2}{n} =$ n =26,583	$\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2 \cdot j} \right)^2}{n} =$ n =50,833

Рассчитаем коэффициент корреляции по формуле Пирсона:

$$r_{x_{i(2,j-1)}x_{i2,j}} = \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2,j-1)} \cdot \sum_{j=1}^k x_{i2,j} \right)}{\sqrt{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2,j-1)} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2,j-1)}}{n} \right)^2}} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2,j-1)} \cdot \sum_{i=1}^n \sum_{j=1}^k x_{i2,j}}{\sqrt{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2,j} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{i2,j}}{n} \right)^2}} =$$

$$= \frac{36 - 4,9167 \cdot 6,833}{\sqrt{26,583 - 4,9167^2} \cdot \sqrt{50,833 - 6,833^2}} = 0,761.$$

Коэффициент $r_{x_{i(2,j-1)}x_{i2,j}}$ показывает надежность теста при определении коэффициента корреляции между двумя половинами.

Коэффициент надежности теста в целом может быть найден по формуле Спирмана-Брауна:

$$r_{H1} = \frac{2 \cdot r_{x_{i(2,j-1)}x_{i2,j}}}{1 + r_{x_{i(2,j-1)}x_{i2,j}}} = \frac{2 \cdot 0,761}{1 + 0,761} = 0,864. \quad (8)$$

2. Метод подсчета средней корреляции тестов испытания

Если сложить все значения r_j в корреляционной таблице тестов испытаний (табл. 6) и разделить эту сумму на общее число тестов $k=22$, то получим среднюю корреляцию:

$$\bar{R} = \frac{\sum_{j=1}^k r_j}{k} = \frac{8,128}{22} = 0,369450. \quad (9)$$

Надежность теста может быть рассчитана по формуле:

$$r_{H2} = \frac{k \cdot \bar{R}}{1 + (k-1) \cdot \bar{R}} = \frac{8,128}{1 + 21 \cdot 0,36945} = 0,928. \quad (10)$$

3. Метод расчета коэффициента надежности теста по формуле Кюдера-Ричардсона (11):

$$r_{H3} = \frac{k}{k-1} \left(1 - \frac{\sum_{j=1}^k \sigma_j^2}{s_x^2} \right) = \frac{22}{21} \cdot \left(1 - \frac{3,118}{12,386} \right) = 0,784, \quad (11)$$

где s_x^2 – общая дисперсия баллов, которую рассчитывают по формуле:

$$s_x^2 = \frac{SS_x}{n-1} = \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n} \right)^2}{n-1} = \frac{136,25}{11} = 12,386. \quad (12)$$

В литературе самым распространенным способом определения надежности является расчет по формуле Кюдера-Ричардсона, но ее применение требует учета определенных ограничений, в связи с чем она уместна только в редких случаях равенства дисперсий большинства заданий испытания. Такого равенства практически не бывает, да и не должно быть при небольшом числе заданий гомогенного испытания. Фактически это означало бы подбор заданий одного уровня трудности.

4. Метод оценки высоких и низких достижений группы

Достаточно точно рассчитывают коэффициент надежности по формуле:

$$r_{H4} = \frac{DT_B + DT_H}{2} = \frac{2 \cdot \sum_{i=1}^n \left(\frac{\sum_{j=1}^k x_{ij}}{k} \right)_{max} + 2 \cdot \sum_{i=1}^n \left(\frac{\sum_{j=1}^k x_{ij}}{k} \right)_{min}}{2} = \frac{\sum_{i=1}^n \left(\frac{\sum_{j=1}^k x_{ij}}{k} \right)_{max} + \sum_{i=1}^n \left(\frac{\sum_{j=1}^k x_{ij}}{k} \right)_{min}}{n}, \quad (13)$$

где DT_B (DT_H) – сумма максимальных (минимальных) результатов, которых достигла продукция, прошедшая проверку.

В литературе предлагается при достаточно большом количестве проведения испытаний формировать каждую исследуемую подгруппу в соотношении 27% от всех пройденных испытаний.

В данном случае испытаний проведено недостаточно большое количество, в связи с чем подгруппы составляют 50% от общего количества проведенных испытаний. Это целесообразно, так как в противном случае дальнейший расчет коэффициента надежности будет занижен. Наиболее высокие результаты прохождения рассматриваемого тестового задания (В): 18;14;14;13;13;13; наиболее низкие (Н): 4;8;9;11;12;12.

Для расчета коэффициента надежности используют данные из табл. 4:

$$DT_B = \frac{0,591 + 0,636 + 0,591 + 0,636 + 0,818 + 0,591}{6} = 0,6439,$$

$$DT_H = \frac{0,182 + 0,545 + 0,5 + 0,545 + 0,364 + 0,409}{6} = 0,424,$$

$$r_{H4} = \frac{0,771 + 0,424}{2} = 0,5341.$$

5. Метод расчета коэффициента надежности теста по формуле Кюдера-Ричардсона при гомогенности испытаний

Если предположение о гомогенности испытания находит подтверждение, то надежность более обоснованно считать по формуле:

$$r_{H5} = \frac{s_x^2 - \sum_{j=1}^n \sigma_j^2}{2s_x^2} + \left[\left(\frac{s_x^2 - \sum_{j=1}^n \sigma_j^2}{2s_x^2} \right)^2 + \frac{\sum_{j=1}^n (r_j^2 \cdot \sigma_j^2)}{2 \cdot s_x^2} \right]^{0.5}. \quad (14)$$

Эту формулу практически не применяли из-за некоторой громоздкости.

Для сравнения с полученными результатами рассчитаем коэффициент надежности и по формуле (14). Воспользуемся данными табл. 5.

$$r_{H5} = \frac{12,386 - 3,118056}{2 \cdot 12,386} + \left[\left(\frac{12,386 - 3,118056}{2 \cdot 12,386} \right)^2 + \left(\frac{0,0018}{2 \cdot 12,386} \right) \right]^{0,5} =$$

$$= 0,7413 + (0,13997 + 0,00007)^{0,5} = 0,37413 + 0,37422 = 0,74835.$$

Выводы

Как видно из результатов, приведенных в табл. 7, показатели надежности, рассчитанные различными методами, существенно отличаются.

Таблица 7

Значения коэффициента надежности теста, полученные разными способами

Метод половинного деления или расщепления	Метод подсчета средней корреляции	Формула Кюдера-Ричардсона при равенстве дисперсий большинства заданий испытания	Метод оценки достижений группы	Формула Кюдера – Ричардсона при гомогенности испытаний
0,864 (очень хорошая)	0,928 (очень хорошая)	0,784 (удовлетворительная)	0,534 (неудовлетворительная)	0,74836 (удовлетворительная)

Все расчетные формулы надежности были приведены к такому виду, чтобы можно было рассмотреть зависимость надежности r_H от $\sum_{i=1}^n \sum_{j=1}^k x_{ij}$. Расчетные формулы имеют следующий вид:

$$r_{H1} = \frac{2 \cdot r_{x_{i(2-j-1)}x_{i2-j}}}{1 + r_{x_{i(2-j-1)}x_{i2-j}}} = \frac{2 \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2-j-1)} \cdot \sum_{j=1}^k x_{i2-j} \right)}{n} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2-j-1)}}{n} \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} - \sum_{j=1}^k x_{i(2-j-1)} \right)}{n}}{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2-j-1)} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2-j-1)}}{n} \right)^2} \cdot \frac{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2-j} \right)}{n} \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} - \sum_{j=1}^k x_{i(2-j-1)} \right)}{n}}{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2-j} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} - \sum_{j=1}^k x_{i(2-j-1)} \right)}{n} \right)^2}}{1 + \frac{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2-j-1)} \right) \cdot \sum_{j=1}^k x_{i2-j}}{n} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2-j-1)}}{n} \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} - \sum_{j=1}^k x_{i(2-j-1)} \right)}{n}}{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i(2-j-1)} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{i(2-j-1)}}{n} \right)^2} \cdot \frac{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2-j} \right) \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} - \sum_{j=1}^k x_{i(2-j-1)} \right)}{n}}{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{i2-j} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} - \sum_{j=1}^k x_{i(2-j-1)} \right)}{n} \right)^2}}{n}}{n}} \right)^2} \quad (15)$$

$$r_{H2} = \frac{k \cdot \bar{R}}{1 + (k-1) \cdot \bar{R}} = \frac{\sum_{j=1}^k \left(\frac{\sum_{i=1}^n \left(x_{ij} \cdot \sum_{j=1}^k x_{ij} \right) - \sum_{i=1}^n x_{ij} \cdot \sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n \cdot n \cdot n} \right)}{\sqrt{\frac{\sum_{i=1}^n x_{ij}^2}{n} - \left(\frac{\sum_{i=1}^n x_{ij}}{n} \right)^2} \cdot \sqrt{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n} \right)^2}} \quad (16)$$

$$1 + (k-1) \cdot \frac{\sum_{j=1}^k \left(\frac{\sum_{i=1}^n x_{ij} \cdot \sum_{j=1}^k x_{ij}}{n} - \sum_{i=1}^n x_{ij} \cdot \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n} \right)}{\sqrt{\frac{\sum_{i=1}^n x_{ij}^2}{n} - \left(\frac{\sum_{i=1}^n x_{ij}}{n} \right)^2} \cdot \sqrt{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2}{n} - \left(\frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{n} \right)^2}} \quad k$$

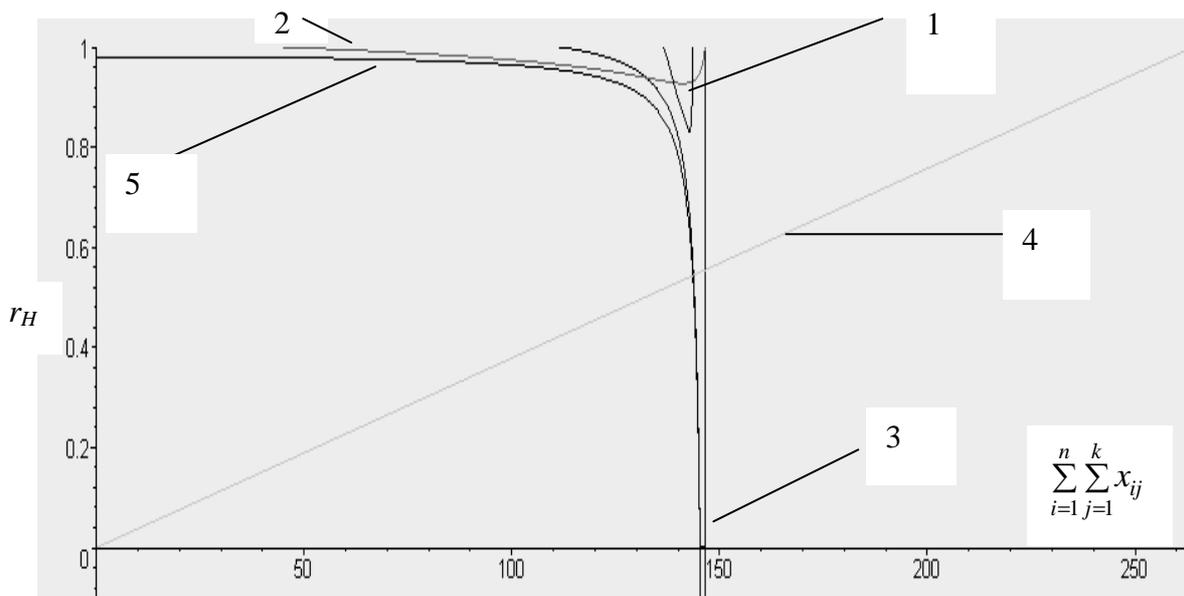
$$r_{H3} = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^N \sigma_j^2}{s_x^2} \right) = \frac{k}{k-1} \left(1 - \frac{\sum_{i=1}^N \sigma_j^2}{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n-1}} \right) \quad (17)$$

$$r_{H4} = \frac{\sum_{i=1}^n \sum_{j=1}^k x_{ij}}{k \cdot n} \quad (18)$$

$$r_{H5} = \frac{s_x^2 - \sum_{i=1}^n \sigma_j^2}{2s_x^2} + \left[\left(\frac{s_x^2 - \sum_{i=1}^n \sigma_j^2}{2s_x^2} \right)^2 + \frac{\sum_{i=1}^n \left(\bar{r}_j^2 \cdot \sigma_j^2 \right)}{2 \cdot s_x^2} \right]^{0.5} =$$

$$= \frac{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n-1} - \sum_{i=1}^n \sigma_j^2}{2 \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n-1}} + \left[\left(\frac{\frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n-1} - \sum_{i=1}^n \sigma_j^2}{2 \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n-1}} \right)^2 + \frac{\sum_{i=1}^n \left(\bar{r}_j^2 \cdot \sigma_j^2 \right)}{2 \cdot \frac{\sum_{i=1}^n \left(\sum_{j=1}^k x_{ij} \right)^2 - \left(\sum_{i=1}^n \sum_{j=1}^k x_{ij} \right)^2}{n-1}} \right]^{0.5} \quad (19)$$

Результаты расчетов показаны на рисунке.



Зависимости надежности от $\sum_{i=1}^n \sum_{j=1}^k x_{ij}$, рассчитанные различными способами

где 1 – метод половинного деления или расщепления; 2 – метод подсчета средней корреляции; 3 – метод расчета по формуле Кюдера-Ричардсона при равенстве дисперсий большинства тестов испытания; 4 – метод оценки достижений групп; 5 – метод расчета по формуле Кюдера-Ричардсона при гомогенности испытаний.

Проанализировав полученные результаты, можно сделать вывод о том, что для *методов* половинного деления и подсчета средней корреляции оценка явно завышена и ниже 0,8 не принимает значения.

Метод оценки достижения группы несовершенен, так как на расчет коэффициента надежности существенным образом влияет уровень достижения продукцией, прошедшей проверку. При большом количестве изделий точное значение для каждой подгруппы проверяемых изделий невозможно рассчитать.

При использовании метода расчета по формуле Кюдера-Ричардсона возникает ситуация, когда коэффициент надежности принимает отрицательное значение, что невозможно.

Таким образом, можно выделить оптимальный способ расчета количественной оценки полноты контроля качества – *метод подсчета по формуле Кюдера-Ричардсона при гомогенных испытаниях*, с помощью которого можно рассчитать достоверную оценку проверки и при этом обеспечить показатель надежности в заданном интервале.

Список литературы

1. Шашкина М.Б. Критерии качества педагогического теста по математике /М.Б. Шашкина// Современное образование. – 2001. – №3. – С. 97-101.
2. Вартамян В.М. Количественная оценка в методах контроля качества наукоемкого высокотехнологического производства /В.М. Вартамян В.М., А.В. Артемова // Вестник науки и техники. – Х. – 2006. –№1-2 (24-25). – С. 80-86.

Рецензент: д-р техн. наук, проф. И.Б. Туркин, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков

Поступила в редакцию 11.03.10

Метод розрахування оптимальної надійності випробувань при випуску якісної продукції

Розглянуто питання практичного застосування методу розрахунку кількісних характеристик контролю якості випробувань на наукоємному високотехнологічному підприємстві. Запропоновано різні процедури визначення коефіцієнта надійності і проведено порівняльне оцінювання різних методів визначення кількісної оцінки тесту. Запропоновано оптимальний засіб розрахунку коефіцієнта надійності проведеного контролю якості продукції. Вибраний метод відрізняється від існуючих тим, що дозволяє розрахувати кількісну оцінку повноти контролю якості продукції, що випускається. Це дає достовірну оцінку перевірки і при цьому забезпечує показник надійності в заданому інтервалі.

Ключові слова: методи, тестові результати, гомогенність випробувань, коефіцієнт надійності, формула Кюдера – Річардсона, середня кореляція

Method of calculation of test optimal reliability during the process of manufacturing of high quality products

The questions of practical application of method of computation of quantitative descriptions of control of quality of tests on naukoemkom are considered highly technological enterprise. Different procedures of determination of coefficient of reliability are offered and comparative estimation of different methods of determination of quantitative estimation of test is conducted. The optimum method of computation of coefficient of reliability of the conducted control of quality of products is offered. The chosen method differs from existing to those, that allows to expect quantitative estimation of plenitude of control of quality of the products let out, that gives reliable estimation to verification and here provides a reliability index in the set interval.

Keywords: methods, test results, homogeneity of tests, coefficient of reliability, the Kyudera-Richardsona formula, middle correlation