ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 53161— 2008 (ИСО 5495:2005)

ОРГАНОЛЕПТИЧЕСКИЙ АНАЛИЗ. МЕТОДОЛОГИЯ

Метод парного сравнения

ISO 5495:2005
Sensory analysis — Methodology — Paired comparison test (MOD)

Издание официальное

БЗ 11—2008/436

Москва
Стандартинформ
2009

gost.e-ecolog.ru

Предисловие

Цели и принципы стандартизации в Российской Федерации установлены Федеральным законом от 27 декабря 2002 г. № 184-ФЗ «О техническом регулировании», а правила применения национальных стандартов Российской Федерации — ГОСТ Р 1.0 — 2004 «Стандартизация в Российской Федерации. Основные положения»

Сведения о стандарте

1 ПОДГОТОВЛЕН Государственным научным учреждением «Всероссийский научно-исследовательский институт консервной и овощесушильной промышленности» (ГНУ ВНИИКОП) на основе аутентичного перевода стандарта, указанного в пункте 4

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 335 «Методы испытаний агропромышленной продукции на безопасность»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 18 декабря 2008 г. № 599-ст

4 Настоящий стандарт является модифицированным по отношению к международному стандарту ИСО 5495:2005 «Органолептический анализ. Методология. Метод парного сравнения» (ISO 5495:2005 «Organoleptic analysis — Methodology — Paired comparison test»). При этом дополнительные слова, фразы, включенные в текст стандарта для учета потребностей национальной экономики Российской Федерации и особенностей российской национальной стандартизации, выделены курсивом.

5 ВВЕДЕН ВПЕРВЫЕ

Информация об изменениях к настоящему стандарту публикуется в ежегодно издаваемом информационном указателе «Национальные стандарты», а текст изменений и поправок — в ежемесячно издаваемых информационных указателях «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ежемесячно издаваемом информационном указателе «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии

© Стандартинформ, 2009

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область *применения* 1
2 Нормативные ссылки 3
3 Термины и определения 3
4 Принцип 4
5 Общие условия проведения испытаний 4
6 Испытатели 5
6.1 Квалификация 5
6.2 Число испытателей 5
7 *Проведение испытаний* 6
8 Анализ и интерпретация результатов 6
8.1 Испытания на различие 6
8.2 Испытания на подобие 7
9 Отчет по испытаниям 7
10 Прецизионность метода и смещение (*систематическая погрешность*) 8
Приложение А (обязательное) *Статистические* таблицы 9
Приложение В (справочное) Примеры 15
Библиография 20

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ОРГАНОЛЕПТИЧЕСКИЙ АНАЛИЗ. МЕТОДОЛОГИЯ

Метод парного сравнения

Organoleptic analysis — Methodology — Paired comparison test

Дата введения — 2010—01—01

1 Область *применения*

Настоящий стандарт устанавливает *метод*, позволяющий определить, существует ли заметное *органолептическое* различие или подобие между образцами двух продуктов по интенсивности рассматриваемой характеристики. Данный *метод* также иногда называют тестом установления направленного различия или тестом 2-АПВ (тестом альтернативного принудительного выбора). Действительно, метод парного сравнения представляет собой обязательный выбор между двумя альтернативными вариантами.

Примечание — Метод парного сравнения — наиболее простой из существующих классификационных тестов, поскольку он предусматривает сопоставление только двух образцов продукта.

Метод применим независимо от того, существует ли различие между продуктами по одной или по нескольким характеристикам, и дает возможность сказать, действительно ли имеет место заметное различие по данной характеристике, и указать направление различия, но не позволяет ответить на вопрос о том, какова степень этого различия. Обнаружение отсутствия разницы между продуктами по изучаемой *органолептической* характеристике не означает, что не существует различия между продуктами по любой другой *органолептической* характеристике.

Метод применим только тогда, когда продукты сравнительно однородны.

Метод эффективен:

а) когда необходимо установить,

что имеется заметное различие между продуктами (метод парного сравнения при проверке различия) или

что не имеется заметного различия между продуктами (метод парного сравнения при проверке подобия) после изменения, например, ингредиентного состава или технологии производства продукта, или его упаковки, или условий хранения и обращения;

б) для целей отбора, обучения и проверки испытателей.

Прежде чем начать испытания, необходимо определить, является ли тест односторонним (руководителю испытаний априори известно направление различия, и альтернативная гипотеза состоит в том, что между образцами продукта существует различие в предполагаемом направлении) или двусторонним (руководитель испытаний не имеет какой-либо *исходной* информации относительно направления различия, и альтернативная гипотеза состоит в том, что между образцами существует различие в одном или в другом направлении).

Метод парного сравнения может быть также использован для того, чтобы сравнивать два продукта с точки зрения предпочтения одного другому. Рисунок 1 иллюстрирует различные случаи использования метода парных сравнений.

Издание официальное

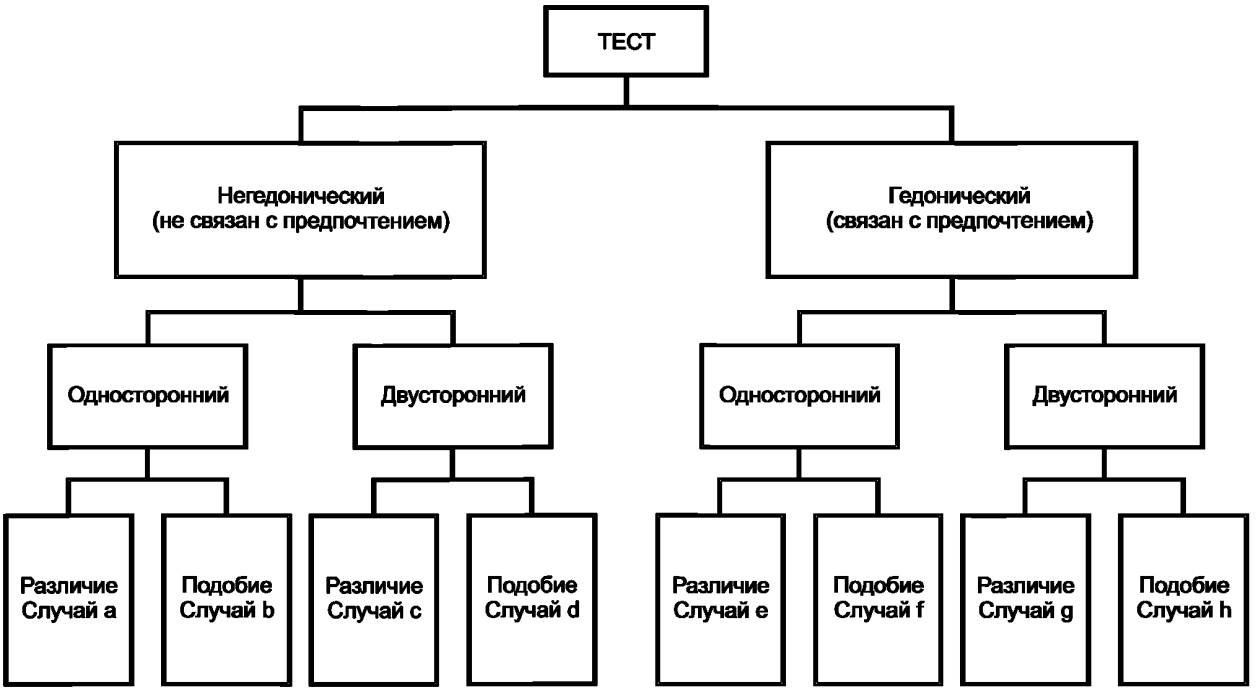

Рисунок 1 — Различные возможные случаи использования метода парного сравнения

Примечание — К настоящему стандарту относится только случай негедонических испытаний.

Пример 1 (Случай a) — Производство бисквита изменено с целью получения более хрустящего продукта. Необходимо проверить, действительно ли нововведение привело к увеличению привлекательности продукта. Следовательно, необходимо попытаться выявить разницу — удостовериться, действительно ли новый продукт ощущается как более хрустящий по сравнению с традиционным (контрольным) продуктом.

Пример 2 (Случай b) — Изготовителю известно, что продукт может содержать небольшие количества ингредиента, который придает ему посторонний флейвор. Поэтому он надеется определить то максимально допустимое количество этого ингредиента, когда разница в флейворе по сравнению с продуктом, не содержащим этого ингредиента, едва-едва заметна и, следовательно, не будет чревата нежелательными последствиями.

Пример 3 (Случай c) — Перед началом производства нового вида супа требуется сравнить два возможных ингредиента, добавляемых для придания продукту соленого вкуса. Исходя из экономических соображений, предпочтение было отдано одному из двух ингредиентов, который в одной и той же концентрации должен обеспечивать более соленый вкус. Следовательно, необходимо подтвердить, что различие в интенсивности вкуса при использовании разных ингредиентов действительно имеет место. Но априори неизвестно, какой из двух видов ингредиента будет обеспечивать более соленый вкус.

Пример 4 (Случай d) — Изготовитель пластмасс, используемых, в частности, в производстве автомашин для изготовления щитков, из экономических соображений хочет заменить обычный смазочный ингредиент новым. При этом он не имеет целью изменить в большую или меньшую сторону скользкость поверхности изделий по сравнению со скользкостью изделий из обычно применяемой пластмассы. Задача состоит в том, чтобы определить, действительно ли в одной и той же концентрации новый смазочный материал обеспечивает такую же скользкость поверхности, как ранее использовавшийся. Необходимо показать, что оба смазочных материала сходны с точки зрения обеспечения скользкости поверхности, но априори неизвестно, какой из них дает наилучший эффект.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р ИСО 5492—2005 Органолептический анализ. Словарь (ИСО 5492:1992, IDT)

ГОСТ Р ИСО 8589—2005 Органолептический анализ. Руководство по проектированию помещений для исследований (ИСО 8589:1988, IDT)

ГОСТ Р 53180—2008 Органолептический анализ. Общее руководство по отбору, обучению и контролю испытателей. Часть 1. Отобранные испытатели (ИСО 8586-1:1993, MOD)

ГОСТ Р 53181—2008 Органолептический анализ. Общее руководство по отбору, обучению и контролю испытателей. Часть 2. Испытатели (ИСО 8586-2:1994, MOD).

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодно издаваемому информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по соответствующим ежемесячно издаваемым информационным указателям, опубликованным в текущем году. Если ссылочный стандарт заменен (изменен), то при пользовании настоящим стандартом следует руководствоваться заменяющим (измененным) стандартом. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, применяется в части, не затрагивающей эту ссылку.

3 Термины и определения

В настоящем стандарте применены термины по ГОСТ Р ИСО 5492, а также следующие термины с соответствующими определениями:

3.1 α **(альфа)-риск**: Вероятность сделать заключение о наличии заметного различия между образцами тогда, когда это различие не существует.

Примечание — В равной мере применимы термины: вероятность ошибки I рода, уровень значимости или частота получения ложноположительных заключений.

3.2 β **(бета)-риск**: Вероятность сделать заключение об отсутствии сколь-нибудь заметного различия тогда, когда это различие существует.

Примечание — В равной мере применимы термины: вероятность ошибки II рода или частота получения ложноотрицательных заключений.

3.3 **различие**: Ситуация, при которой образцы продукта могут быть отличены друг от друга на основе их *органолептических* характеристик.

Примечание — Долю оценок, когда было отмечено заметное различие между двумя образцами продукта по изучаемой *органолептической* характеристике, обозначают символом p_d.

3.4 **тест с использованием одностороннего критерия**: Тест, при проведении которого руководителю испытаний априори известно направление различия.

Примечание — Нулевая гипотеза H_0: образцы продукта не имеют различия; доля *зафиксированных* правильных ответов $p = 1/2$. Альтернативная гипотеза H_1: $p > 1/2$.

3.5 **тест с использованием двустороннего критерия**: Тест, при проведении которого руководителю испытаний априори не известно направление различия.

Примечание — Нулевая гипотеза H_0: образцы продукта не имеют различия; доля *зафиксированных* правильных ответов в отношении одного из образцов $p = 1/2$. Альтернативная гипотеза H_1: $p \neq 1/2$.

3.6 **правильные ответы (предполагаемые ответы)**: В тесте с использованием одностороннего критерия: число испытателей, выбравших тот образец продукта, относительно которого руководитель испытаний надеется, что он обладает бóльшей интенсивностью изучаемой *органолептической* характеристики.

3.7 **согласующиеся ответы**: В тесте с использованием двустороннего критерия: наибольшее из значений числа испытателей, указавших на образец А (или на образец В) как на продукт, обладающий бóльшей интенсивностью исследуемой *органолептической* характеристики.

Примечание — Подсчеты делают так, как сказано выше, поскольку в данном случае ответы не могут расцениваться как правильные или как неправильные.

3.8 **продукт**: Оцениваемый материал.

3.9 **образец**: Единица продукции, подготовленная, предоставленная испытателям и оцениваемая ими в процессе проведения теста.

3.10 **чувствительность**: Обобщенный термин, используемый для того, чтобы суммарно обозначить условия проведения теста.

Примечание — С точки зрения математической статистики чувствительность теста определяется величинами α, β и p_d.

3.11 **подобие**: Ситуация, при которой любые *возможные* различия между образцами продукта настолько малы, что продукты можно рассматривать как взаимозаменяемые.

4 Принцип

Число испытателей выбирают, исходя из желаемой чувствительности теста (см. 6.2 и сноски к таблицам А.4 и А.5, *приложение А*).

Испытатели получают комплект из двух образцов продукта (т.е. образцы, представляющие собой пару). Задача испытателя — выбрать из них тот из образцов, который имеет более интенсивно выраженную *органолептическую* характеристику, подлежащую оценке, даже в том случае, если выбор будет основан только на догадках.

Примечание — Один из образцов может быть контрольным.

Затем подсчитывают суммарное количество голосов в пользу одного и другого образца всеми *испытателями* и определяют значимость результата испытаний со ссылкой на статистическую таблицу, выбранную исходя из того, было ли априори известно, какой из образцов должен быть признан обладающим более интенсивно выраженной *органолептической* характеристикой (односторонний критерий), или это не было известно (двусторонний критерий).

5 Общие условия проведения испытаний

5.1 Четко оговаривают цель испытаний и устанавливают, будет ли проводиться тест с использованием одностороннего или двустороннего критерия и идет ли речь о проверке различия между продуктами или о проверке подобия продуктов, а также определяют, какова желательная чувствительность теста.

5.2 Испытания проводят в условиях, которые исключают возможность обмена информацией между испытателями, пока все оценки не будут ими выполнены; для этого используют оборудование и помещения для испытаний, соответствующие требованиям *ГОСТ Р ИСО 8589*.

5.3 Образцы продукции для испытаний готовят в отсутствие испытателей одинаковым способом: применяют одинаковое оборудование, одинаковую посуду.

5.4 Испытатели не должны получать никакой информации об интенсивности контролируемой *органолептической* характеристики продукта, исходя из способа подготовки и предоставления им образцов продукта. Например, если исследуется только тактильная характеристика продукта, то следует избегать каких-либо различий во внешнем виде образцов. Следует замаскировывать цветовые различия образцов (поскольку не изучается именно эта характеристика), используя при освещении светофильтры или обеспечивая приглушенное освещение. Можно также, если образцы несколько отличны по внешнему виду, подавать их испытателю последовательно, а не одновременно.

5.5 Образцы или посуду, в которой их подают, кодируют одинаковым способом; предпочтительно использовать в качестве кодов трехзначные числа, в каждом тесте выбирая их случайным образом. Каждая пара состоит из двух образцов, и каждый образец имеет свой собственный код. Желательно, чтобы в течение всей серии испытаний для всех испытателей использовались разные коды образцов. Однако могут быть использованы и два одних и тех же кода образцов для всех испытателей, если при этом каждый код используется для данного испытателя лишь единожды в течение данной серии испытаний (например, когда в одной серии испытаний метод парного сравнения используется для нескольких различных продуктов).

5.6 Количество (объем) продукта, которое подают испытателю, должно быть одинаковым в обоих образцах в паре, как должно быть одинаковым и во всей серии испытаний данного вида продукта. Количество (объем) проб продукта, которое испытатель должен использовать при опробовании, можно установить заранее. Если это не сделано, то следует, тем не менее, указать испытателям, что

они должны использовать одинаковые количества продукта при оценивании пробы каждого из сравниваемых образцов.

5.7 Температура образцов продукта в каждой паре должна быть одинакова; это касается также всех других образцов в серии испытаний данного вида продукта. Желательно предоставлять испытателям образцы продукта при той температуре, при которой данный продукт обычно потребляют.

5.8 Испытателей нужно уведомить, должны ли они следовать определенным правилам при оценивании продуктов (например, в тестах на сопоставление вкуса продуктов, должны ли они проглатывать пробу, или в тестах на сопоставление тактильных ощущений, должны ли делать какие-либо специфические движения), или могут поступать по своему усмотрению. В этом последнем случае испытатели должны быть предупреждены, что следует поступать одинаковым образом в отношении всех испытуемых образцов продукта.

5.9 В течение всей серии испытаний до того момента, пока все испытания не будут завершены, необходимо избегать возможности получения испытателями какой-либо информации, которая может помочь в идентификации пробы продукта, в определении возможных эффектов от той или иной его обработки, изготовления.

6 Испытатели

6.1 Квалификация

Все испытатели должны обладать одинаковым уровнем квалификации, и этот уровень должен быть выбран исходя из цели испытаний (*см. ГОСТ Р 53180 и ГОСТ Р 53181)*. Практический опыт и предварительное знакомство с продуктом могут положительно сказаться на качестве работы испытателя и тем самым увеличить вероятность обнаружения значимых различий при сопоставлении образцов продукта. Проводимая время от времени проверка *квалификации* испытателей может быть полезна для повышения их чувствительности.

Все испытатели должны быть хорошо знакомы с сущностью метода парного сравнения (с формами протоколов, задачей испытаний, процедурой оценки). Кроме того, испытатели должны понимать, в чем состоит та *органолептическая* характеристика продукта, которую им предстоит оценивать. Эта характеристика должна быть определена вербально, посредством ссылок на известные вещества или путем предварительного предоставления испытателям нескольких образцов продукта, в которых интенсивность исследуемой характеристики имеет разные уровни.

6.2 Число испытателей

Число испытателей выбирают исходя из того, какую чувствительность теста желательно обеспечить (см. таблицу А.4, *приложение А,* — для случая одностороннего критерия и таблицу А.5, *приложение А,* — для случая двустороннего критерия). Использование большого числа испытателей является предпосылкой успеха в выявлении незначительных различий между двумя продуктами. Однако на практике выбор числа испытателей часто определяется обстоятельствами материального плана, такими как ограниченное время, отпущенное на проведение теста, имеющееся в наличии число подходящих испытателей, выделенное для теста количество продукта и пр. При проведении теста на различие число испытателей составляет примерно 24—30 человек. При проведении теста на подобие для обеспечения эквивалентной чувствительности теста число испытателей должно быть примерно удвоено (т.е. приблизительно должно составлять 60 человек). При проведении теста на подобие оценки не должны осуществляться повторно одними и теми же испытателями. При проведении теста на различие допускается повторное использование испытателя, но по возможности этого следует избегать. Однако если параллельные оценки все-таки предполагается использовать, чтобы обеспечить достаточное общее число оценок, то должны быть предприняты все возможные усилия к тому, чтобы все испытатели провели равное число параллельных оценок. Например, если имеются только 10 испытателей, то каждый из них должен сделать три параллельных оценки так, чтобы общее число оценок равнялось 30.

Примечание — Если проводят тест на подобие и используют данные таблицы А.3, *приложение А*, то рассмотрение полученных 10 испытателями в трех повторностях 30 оценок как независимых результатов не является правомочным. Однако если проводят тест на различие с применением данных таблиц А.1 и А.2, *приложение А*, то использование параллельных оценок допускается [1], [2]. В некоторых публикациях [3], [4] при использовании параллельных оценок при различительных испытаниях предлагаются альтернативные подходы к статистической обработке результатов теста.

7 Проведение испытаний

7.1 Перед проведением теста заготавливают необходимое число бланков рабочих листов и протоколов (см. рисунки В.1, В.2. и В.3, *приложение В*), исходя из того, что будет использовано равное число проб продуктов А и В — подаваемых в двух различных положениях.

7.2 Испытателю подают (последовательно или одновременно) по два образца продукта, составляющих пару (см. 5.4). Если два образца подают одновременно, то при подаче пространственно их упорядочивают так, чтобы парные пробы располагались перед каждым испытателем либо в ряд (слева направо), либо в столбик (снизу вверх). Испытатели должны опробовать два образца, составляющие пару, в том порядке, как это указано в рабочем листе, но обычно испытателям разрешают, если те пожелают, провести повторные опробования пары образцов (в том случае, конечно, когда природа продукта позволяет повторить опробование).

7.3 Рабочий лист оформляют для каждой отдельной парной пробы образцов. Если испытатель будет выполнять испытания более чем одной пары образцов в течение серии испытаний, то следует получить от него заполненный бланк рабочего листа и убрать со стола остатки продуктов, прежде чем предлагать ему следующую пару образцов. Испытатель никогда не должен возвращаться к предыдущей паре образцов и никогда не должен изменять свой вердикт относительно результата ранее проведенного испытания.

7.4 Не следует задавать испытателю какие-либо вопросы относительно предпочтительности, приемлемости или степени различия после того, как он сделал свой выбор образца, имеющего более выраженную исследуемую характеристику продукта. Выбор, который был им только что сделан, может оказать влияние на ответы на последующие дополнительные вопросы. Ответы на такие вопросы можно получить в отдельных испытаниях на предпочтение, приемлемость, степень различия и т.п. [5]. Для того чтобы испытатель мог пояснить, почему он сделал тот или иной выбор, в рабочий лист может быть введен раздел «Примечания».

7.5 Метод парного сравнения является тестом, в котором предусмотрена процедура принудительного выбора: испытателям не разрешается использовать такой вариант ответа, как «нет различия». Испытателя нужно проинформировать, что если он не сумеет обнаружить различия между образцами продукта, то должен просто выбрать наугад один из образцов и в разделе рабочего листа «Примечания» отметить, что сделанный выбор является не более чем догадкой.

8 Анализ и интерпретация результатов

8.1 Испытания на различие

8.1.1 Случай одностороннего критерия

Для анализа данных, полученных методом парного сравнения, используют таблицу А.1, *приложение А*. Если число правильных ответов больше или равно числу, указанному в таблице А.1, *приложение А* (для соответствующего числа испытателей и уровня α-риска, выбранного при проведении теста), то делают вывод, что между образцами продукта существует заметное различие (см. раздел В.1, *приложение В*).

При желании можно вычислить доверительный интервал для доли испытателей, способных различить образцы продукта. Методика расчета приведена в разделе В.5, *приложение В*.

8.1.2 Случай двустороннего критерия

Для анализа данных, полученных методом парного сравнения, используют таблицу А.2, *приложение А*. Если число согласующихся ответов больше или равно числу, указанному в таблице А.2, *приложение А* (для соответствующего числа испытателей и уровня α-риска, выбранного при проведении теста), то делают вывод, что между образцами продукта существует заметное различие (см. В.3, *приложение В*).

При желании можно вычислить доверительный интервал для доли тех, кто способен различить образцы продукта. Методика расчета приведена в разделе В.5, *приложение В*.

8.2 Испытания на подобие[1)]

8.2.1 Случай одностороннего критерия

Для анализа данных, полученных методом парного сравнения, используют таблицу А.3, *приложение А*. Если число правильных ответов меньше или равно числу, указанному в таблице А.3, *приложение А* (для соответствующего числа испытателей, уровня β-риска и величины p_d, выбранных при проведении теста), то делают вывод, что между образцами продукта нет существенного различия (см. В.2, *приложение В*). Если результаты, получаемые в разных тестах, сопоставимы, то значение p_d может быть вычислено по данным всех этих тестов.

При желании можно вычислить доверительный интервал для доли испытателей, способных различить образцы продукта. Методика расчета приведена в разделе В.5, *приложение В*.

Нельзя сделать какие-либо заключения, если максимальное число правильных ответов меньше, чем $n/2$.

8.2.2 Случай двустороннего критерия

Для анализа данных, полученных методом парного сравнения, используют таблицу А.3, *приложение А*. Если число согласующихся ответов меньше или равно числу, указанному в таблице А.3, *приложение А* (для соответствующего числа испытателей, уровня β-риска и величины p_d, выбранных при проведении теста), то делают вывод, что между образцами продукта нет существенного различия (см. В.4, *приложение В*). Если результаты, получаемые в разных тестах, сопоставимы, то значение p_d может быть вычислено по данным всех этих тестов.

При желании можно вычислить доверительный интервал для доли испытателей, способных различить образцы продукта. Методика расчета приведена в разделе В.5, *приложение В*.

9 Отчет по испытаниям

В отчете по испытаниям сообщают цель испытаний, результаты испытаний и заключения.

Рекомендуется внести в отчет следующую дополнительную информацию:

- задача испытаний и характер изучаемого способа обработки продукта;
- подробные сведения для идентификации испытуемых образцов продукта: их происхождение, способ подготовки, количество, форма, хранение перед испытаниями, количество продукта, предоставленного испытателям, температура продукта (эта информация, касающаяся образцов, должна свидетельствовать, что все проведенные операции их хранения, обработки и подготовки выполнены так, что можно утверждать, что различие между образцами, если оно обнаружено, может быть вызвано только теми изменениями в их характеристиках, которые связаны с изучаемой причиной);
- число испытателей, число правильных или согласующихся ответов и результаты статистической обработки (включая значения величин α, β и p_d, которые использовались при испытаниях);
- сведения об испытателях (их практический опыт в проведении органолептических испытаний, знакомство с продуктом и с образцами продукта, подвергавшегося испытаниям), возраст и пол (*ГОСТ Р 53180 и ГОСТ Р 53181)*;
- любые специальные сведения и рекомендации, предоставленные испытателям в связи с проводившимися испытаниями, особенно касающиеся точных вербальных определений и эталонных образцов продукта, с помощью которых разъяснялось, какое именно свойство продукта подлежит рассмотрению, и/или касающиеся доведения до сведения испытателей специальных правил проведения испытаний;
- условия при проведении испытаний: использовавшиеся средства испытаний, осуществлялась ли подача образцов испытателям одновременно или последовательно, была ли после проведения испытаний предоставлена информация, позволяющая идентифицировать образцы, и если была предоставлена, то каким образом;
- место и дату проведения испытаний, фамилию руководителя испытаний.

[1)] В настоящем стандарте «подобный» не означает «идентичный». Этот термин скорее означает, что два продукта достаточно схожи, чтобы можно было использовать один вместо другого. Невозможно доказать, что два продукта являются идентичными. Однако можно продемонстрировать, что любое различие, какое существует между продуктами, настолько невелико, что не имеет практического значения.

10 Прецизионность метода и смещение (*систематическая погрешность*)

Поскольку результаты *органолептических* различительных испытаний зависят от индивидуальной чувствительности испытателей, не представляется возможным сделать общее заключение относительно воспроизводимости результатов, которое было бы применимо по отношению к любой группе испытателей. Прецизионность метода, связанная с конкретной группой испытателей, возрастает по мере увеличения числа участников испытания и по мере возрастания их квалификации, а также зависит от подвергаемого испытаниям продукта.

Поскольку используется процедура принудительного выбора, получаемые данным методом результаты не содержат систематической погрешности (смещения), если полностью соблюдаются требования раздела 7.

Приложение А
(обязательное)

Статистические таблицы

А.1 Определение заметного различия или подобия на основе таблиц А.1 — А.3

Таблица А.1 — Минимальное число правильных ответов, необходимое для заключения о существовании заметного различия между образцами на основании результатов теста парного сравнения, односторонний критерий[1), 2)]

n	*Минимальное число правильных ответов при уровне α-риска*					*n*	*Минимальное число правильных ответов при уровне α-риска*				
	0,20	0,10	0,05	0,01	0,001		0,20	0,10	0,05	0,01	0,001
10	7	8	9	10	10	36	22	23	24	26	28
11	8	9	9	10	11	37	22	23	24	27	29
12	8	9	10	11	12	38	23	24	25	27	29
13	9	10	10	12	13	39	23	24	26	28	30
14	10	10	11	12	13	40	24	25	26	28	31
15	10	11	12	13	14						
16	11	12	12	14	15	44	26	27	28	31	33
17	11	12	13	14	16	48	28	29	31	33	36
18	12	13	13	15	16	52	30	32	33	35	38
19	12	13	14	15	17	56	32	34	35	38	40
20	13	14	15	16	18	60	34	36	37	40	43
21	13	14	15	17	18	64	36	38	40	42	45
22	14	15	16	17	19	68	38	40	42	45	48
						72	41	42	44	47	50
23	15	16	16	18	20	76	43	45	46	49	52
24	15	16	17	19	20	80	45	47	48	51	55
25	16	17	18	19	21						
26	16	17	18	20	22	84	47	49	51	54	57
27	17	18	19	20	22	88	49	51	53	56	59
28	17	18	19	21	23	92	51	53	55	58	62
29	18	19	20	22	24	96	53	55	57	60	64
30	18	20	20	22	24	100	55	57	59	63	66
31	19	20	21	23	25	104	57	60	61	65	69
32	19	21	22	24	26	108	59	62	64	67	71
33	20	21	22	24	26	112	61	64	66	69	73
34	20	22	23	25	27	116	64	66	68	71	76
35	21	22	23	25	27	120	66	68	70	74	78

Примечания

1 Числа, указанные в таблице, определены по формуле биномиального распределения и являются точными. Для значений *n*, не включенных в таблицу, могут быть вычислены приближенные значения для опущенных входов в таблицу, что делается следующим образом: минимальное число ответов *x* равняется ближайшему целому числу, бóльшему, чем

$$x = (n + 1)/2 + z\sqrt{0{,}25n},$$

где *z* — функция уровня значимости: $z = 0{,}84$ для $\alpha = 0{,}20$; $z = 1{,}28$ для $\alpha = 0{,}10$; $z = 1{,}64$ для $\alpha = 0{,}05$: $z = 2{,}33$ для $\alpha = 0{,}01$; $z = 3{,}09$ для $\alpha = 0{,}001$.

2 Значения $n < 18$ обычно не рекомендуется использовать при испытаниях на различие, проводимых по методу парного сравнения.

---

1) Значения, приведенные в таблице, рассчитаны по точной формуле биномиального распределения для параметра $p = 0{,}5$ и для *n* повторностей благодаря сведениям SAS, развитым в [6].

2) Приведенные в таблице значения соответствуют минимальному числу правильных ответов, необходимому для обеспечения статистической значимости при заданном уровне α-риска (см. соответствующую графу) и заданном числе испытателей *n* (см. соответствующую строку). Утверждение «нет различия» отвергается, если число правильных ответов больше или равно числу, указанному в таблице.

Таблица А.2 — Минимальное число согласующихся ответов, необходимое для заключения о существовании заметного различия между образцами на основании результатов теста парного сравнения, двусторонний критерий[1), 2)]

n	*Минимальное число согласующихся ответов при уровне α-риска*					*n*	*Минимальное число согласующихся ответов при уровне α-риска*				
	0,20	0,10	0,05	0,01	0,001		0,20	0,10	0,05	0,01	0,001
10	8	9	9	10		36	23	24	25	27	29
11	9	9	10	11	11	37	23	24	25	27	29
12	9	10	10	11	12	38	24	25	26	28	30
13	10	10	11	12	13	39	24	26	27	28	31
14	10	11	12	13	14	40	25	26	27	29	31
15	11	12	12	13	14						
16	12	12	13	14	15	44	27	28	29	31	34
17	12	13	13	15	16	48	29	31	32	34	36
18	13	13	14	15	17	52	32	33	34	36	39
19	13	14	15	16	17	56	34	35	36	39	41
20	14	15	15	17	18	60	36	37	39	41	44
21	14	15	16	17	19	64	38	40	41	43	46
22	15	16	17	18	19	68	40	42	43	46	48
23	16	16	17	19	20	72	42	44	45	48	51
24	16	17	18	19	21	76	45	46	48	50	53
25	17	18	18	20	21	80	47	48	50	52	56
26	17	18	19	20	22	84	49	51	52	55	58
27	18	19	20	21	23	88	51	53	54	57	60
28	18	19	20	22	23	92	53	55	56	59	63
29	19	20	21	22	24	96	55	57	59	62	65
30	20	20	21	23	25	100	57	59	61	64	67
31	20	21	22	24	25	104	60	61	53	66	70
32	21	22	23	24	26	108	62	64	65	68	72
33	21	22	23	25	27	112	64	66	67	71	74
34	22	23	24	25	27	116	66	68	70	73	77
35	22	23	24	26	28	120	68	70	72	75	79

Примечания

1 Числа, указанные в таблице, определены по формулам биномиального распределения и являются точными. Для значений *n*, не включенных в таблицу, могут быть вычислены приближенные значения, что делается следующим образом: минимальное число правильных ответов *x* равняется ближайшему целому числу, бóльшему, чем

$$x = (n + 1)/2 + z\sqrt{0{,}25n},$$

где *z* — функция уровня значимости: $z = 1{,}28$ для $\alpha = 0{,}20$; $z = 1{,}64$ для $\alpha = 0{,}10$; $z = 1{,}96$ для $\alpha = 0{,}05$; $z = 2{,}58$ для $\alpha = 0{,}01$; $z = 3{,}29$ для $\alpha = 0{,}001$.

2 Значения $n < 18$ обычно не рекомендуется использовать при испытаниях на различие, проводимых по методу парного сравнения.

---

1) Значения, приведенные в таблице, рассчитаны по точной формуле биномиального распределения для параметра p = 0,5 и для n повторностей благодаря сведениям SAS, развитым в [6].

2) Приведенные в таблице значения соответствуют минимальному числу правильных ответов, необходимому для обеспечения статистической значимости при заданном уровне α-риска (см. соответствующую графу) и заданном числе испытателей n (см. соответствующую строку). Утверждение «нет различия» отвергается, если число согласующихся ответов больше или равно числу, указанному в таблице.

Таблица А.3 — Максимальное число правильных или согласующихся ответов, необходимое для заключения о том, что сравниваемые образцы подобны друг другу, на основании результатов теста парного сравнения[1), 2)]

n	β	*Необходимое максимальное число правильных или согласующихся ответов при* p_d					n	β	*Необходимое максимальное число правильных или согласующихся ответов при* p_d				
		10%	20%	30%	40%	50%			10%	20%	30%	40%	50%
18	0,001	—	—	—	—	—	72	0,001	—	—	—	37	40
	0,01	—	—	—	—	—		0,01	—	—	36	40	44
	0,05	—	—	—	—	9		0,05	—	—	39	43	47
	0,10	—	—	—	9	10		0,10	—	39	41	44	48
	0,20	—	—	9	10	11		0,20	—	39	42	46	50
24	0,001	—	—	—	—	—	78	0,001	—	—	—	40	44
	0,01	—	—	—	—	12		0,01	—	—	40	44	48
	0,05	—	—	—	12	13		0,05	—	39	43	47	51
	0,10	—	—	12	13	14		0,10	—	40	44	48	53
	0,20	—	—	13	14	15		0,20	—	42	46	50	54
30	0,001	—	—	—	—	—	84	0,001	—	—	—	44	48
	0,01	—	—	—	—	16		0,01	—	—	43	48	53
	0,05	—	—	—	16	17		0,05	—	42	46	51	55
	0,10	—	—	15	17	18		0,10	—	44	48	52	57
	0,20	—	15	16	18	20		0,20	—	46	50	54	59
36	0,001	—	—	—	—	—	90	0,001	—	—	—	48	53
	0,01	—	—	—	18	20		0,01	—	—	47	52	57
	0,05	—	—	18	20	22		0,05	—	45	50	55	60
	0,10	—	—	19	21	23		0,10	—	47	52	56	61
	0,20	—	18	20	22	24		0,20	45	49	54	58	63
42	0,001	—	—	—	—	21	96	0,001	—	—	—	52	57
	0,01	—	—	—	21	24		0,01	—	—	50	56	61
	0,05	—	—	21	23	26		0,05	—	49	54	59	64
	0,10	—	—	22	25	27		0,10	—	50	55	60	66
	0,20	—	22	24	26	28		0,20	48	53	58	62	68
48	0,001	—	—	—	—	25	102	0,001	—	—	—	55	61
	0,01	—	—	—	25	28		0,01	—	—	54	59	65
	0,05	—	—	25	27	30		0,05	—	52	57	63	68
	0,10	—	—	26	28	31		0,10	—	54	59	64	70
	0,20	—	25	27	30	33		0,20	51	56	61	67	72
54	0,001	—	—	—	—	29	108	0,001	—	—	54	59	65
	0,01	—	—	—	29	32		0,01	—	—	57	63	69
	0,05	—	—	28	31	34		0,05	—	55	61	67	72
	0,10	—	27	30	32	35		0,10	—	57	63	68	74
	0,20	—	28	31	34	37		0,20	54	60	65	71	76
60	0,001	—	—		—	33	114	0,001	—	—	57	63	69
	0,01	—	—	—	33	36		0,01	—	—	61	67	73
	0,05	—	—	32	35	38		0,05	—	59	65	71	77
	0,10	—	30	33	36	40		0,10	—	61	67	72	79
	0,20	—	32	35	38	41		0,20	57	63	69	75	81
66	0,001	—	—	—	—	37	120	0,001	—	—	61	67	73
	0,01	—	—	33	36	40		0,01	—	—	65	71	78
	0,05	—	—	35	39	43		0,05	—	62	68	75	81
	0,10	—	34	37	40	44		0,10	—	64	70	77	83
	0,20	—	35	39	42	46		0,20	60	67	73	79	85

[1)] Указанные в данной таблице значения получены благодаря использованию программы расчетов, основанной на вычислении доверительных интервалов по точной формуле биномиального распределения, см. [7].

[2)] Приведенные в таблице значения соответствуют максимальному числу правильных или согласующихся ответов, необходимому для заключения о «подобии» сравниваемых объектов при выбранных уровнях p_d, β и n. Утверждение «нет различия» принимается при уровне значимости 100 (1 — β) %, если число правильных или согласующихся ответов меньше или равно числу, указанному в таблице.

Окончание таблицы А.3

n	β	*Необходимое максимальное число правильных или согласующихся ответов при* p_d					n	β	*Необходимое максимальное число правильных или согласующихся ответов при* p_d				
		10%	20%	30%	40%	50%			10%	20%	30%	40%	50%
126	0,001	—		64	70	77	132	0,001	—	—	67	74	81
	0,01	—	—	68	75	82		0,01	—	65	72	79	86
	0,05	—	66	72	79	85		0,05	—	69	76	83	90
	0,10	—	68	74	81	87		0,10	—	71	78	85	92
	0,20	64	70	76	83	89		0,20	67	73	80	87	94

Примечания

1 Числа, указанные в таблице, являются точными значениями, поскольку определены по формуле биномиального распределения. Для значений n, отсутствующих в таблице, верхний доверительный интервал для p_d при уровне значимости 100 (1 — β) % может быть вычислен по формуле

$$[2(x/n) - 1] + 2 \cdot z_\beta \sqrt{(nx - x^2)/n^3},$$

где x — число правильных или согласующихся ответов;

n — число испытателей;

$z_\beta = 0,84$ для $\beta = 0,20$; $z_\beta = 1,28$ для $\beta = 0,10$; $z_\beta = 1,64$ для $\beta = 0,05$; $z_\beta = 2,33$ для $\beta = 0,01$; $z_\beta = 3,09$ для $\beta = 0,001$.

Если расчетное значение меньше, чем выбранное граничное значение p_d, то заявляют, что сравниваемые объекты подобны друг другу на уровне значимости β.

2 Значения $n < 30$ обычно не рекомендуется использовать при испытаниях на подобие, проводимых по методу парного сравнения.

3 Значения, соответствующие числу правильных ответов, меньшему, чем $n/2$, не указаны в данной таблице, вместо них поставлен знак пропуска (тире).

А.2 Статистический подход при определении числа испытателей на основе данных таблиц А.4 (односторонний критерий) и А.5 (двусторонний критерий)

Статистическая чувствительность теста зависит от трех величин: α-риска, β-риска и p_d — максимально допустимой доли «испытателей, обнаруживающих различие»[1)]. Перед началом проведения испытаний выбирают значения α, β и p_d, руководствуясь нижеследующими рекомендациями.

Как общее правило, статистически значимый результат при значении α-риска, в пределах:

- от 10 % до 5 % (от 0,10 до 0,05) соответствует ситуации, когда имеются слабые основания надеяться на то, что разница между образцами продукта действительно ощутима;
- от 5 % до 1 % (от 0,05 до 0,01) соответствует ситуации, когда имеются умеренные основания надеяться на то, что разница между продуктами действительно ощутима;
- от 1 % до 0,1 % (от 0,01 до 0,001) соответствует ситуации, когда имеются веские основания надеяться на то, что разница между образцами действительно ощутима,
- менее 0,1 % (менее 0,001) соответствует ситуации, когда имеются весьма веские основания надеяться, что разница между образцами действительно ощутима.

В случае, связанном с β-риском, весомость оснований заключить, что различие между образцами не выявлено, оценивается с помощью тех же критериев, что были оговорены выше (следует лишь заменить слова «действительно ощутима» на слова «не ощутима»).

Максимально допустимая доля испытателей, которые способны ощутить различие между сравниваемыми продуктами, p_d, может быть разделена на три диапазона:

- если p_d меньше 25 %, то это небольшие значения,
- если p_d больше 25 %, но меньше 35 %, то это средние значения,
- если p_d больше 35 %, то это большие значения.

[1)] В настоящем стандарте вероятность правильного ответа, p_c, определяется формулой:

$$p_c = 1 \cdot p_d + (1/2) \cdot (1 - p_d),$$

где p_d — это доля всей совокупности испытателей, таких, которые способны выявить различие между двумя образцами.

Применительно к методу парного сравнения может быть также использована психометрическая модель процесса принятия решения испытателем, такая как, например, модель Тюрстоуна-Юра (см. [8]).

Необходимое число испытателей выбирают таким образом, чтобы получить уровень чувствительности теста, отвечающий задаче испытаний. В таблице А.4 находят строку, соответствующую выбранному значению p_d, и графу, соответствующую выбранному значению β. Минимально необходимое число испытателей находят в строке, соответствующей выбранному значению α. Таблицу А.4 можно также использовать для нахождения таких величин p_d, β и α, которые обеспечивают приемлемую чувствительность теста, если число испытателей приходится устанавливать исходя из реальных практических возможностей. Этот подход подробно описан в [5].

Таблица А.4 — Число испытателей, необходимых для проведения теста парного сравнения при использовании одностороннего критерия[1), 2)]

α	p_d, %	*Число испытателей, необходимых для проведения теста парного сравнения с использованием одностороннего критерия, для значений риска β, равных*					
		0,50	0,20	0,10	0,05	0,01	0001
0,50	50	—	—	—	9	22	33
0,20		—	12	19	26	39	58
0,10		—	19	26	33	48	70
0,05		13	23	33	42	58	82
0,01		35	40	50	59	80	107
0,001		38	61	71	83	107	140
0,50	40	—	—	9	20	33	55
0,20		—	19	30	39	60	94
0,10		14	28	39	53	79	113
0,05		18	37	53	67	93	132
0,01		35	64	80	96	130	174
0,001		61	95	117	135	176	228
0,50	30	—	—	23	33	59	108
0,20		—	32	49	68	110	166
0,10		21	53	72	96	145	208
0,05		30	69	93	119	173	243
0,01		64	112	143	174	235	319
0,001		107	172	210	246	318	412
0,50	20	—	23	45	67	133	237
0,20		21	77	112	158	253	384
0,10		46	115	168	214	322	471
0,05		71	158	213	268	392	554
0,01		141	252	325	391	535	726
0,001		241	386	479	556	731	944
0,50	10	—	75	167	271	539	951
0,20		81	294	451	618	1006	1555
0,10		170	461	658	861	1310	1905
0,05		281	620	866	1092	1583	2237
0,01		550	1007	1301	1582	2170	2927
0,001		961	1551	1908	2248	2937	3812

Примечание — Прочерки соответствуют случаям, не представляющим практического интереса (высокие значения рисков α и β, принятые в расчет при выборе величины p_d).

1) Указанные в данной таблице значения заимствованы из [6] или рассчитаны по точной формуле биномиального распределения с параметром $p = 0,5$, для n откликов, благодаря программе SAS.

2) Число испытателей, необходимое для реализации теста парного сравнения с установленным уровнем чувствительности, определяются величинами p_d, α и β. В таблице А.4 находят выбранное значение p_d и графу, соответствующую выбранному значению β. Минимально необходимое число испытателей *находят на пересечении этих показателей в строке указанной графы*, соответствующей выбранному значению α.

Таблица А.5 — Число испытателей, необходимых для проведения теста парного сравнения при использовании двустороннего критерия[1), 2)]

α	p_d, %	*Число испытателей, необходимых для проведения теста парного сравнения с использованием двустороннего критерия, для значений риска β, равных*					
		0,50	0,20	0,10	0,05	0,01	0,001
0,50	50	—	—	—	23	33	52
0,20		—	19	26	33	48	70
0,10		—	23	33	42	58	82
0,05		17	30	42	49	67	92
0,01		26	44	57	66	87	117
0,001		42	66	78	90	117	149
0,50	40	—	—	25	33	54	86
0,20		—	28	39	53	79	113
0,10		18	37	53	67	93	132
0,05		25	49	65	79	110	149
0,01		44	73	92	108	144	191
0,001		48	102	126	147	188	240
0,50	30	—	29	44	63	98	156
0,20		21	53	72	96	145	208
0,10		30	69	93	119	173	243
0,05		44	90	114	145	199	276
0,01		73	131	164	195	261	345
0,001		121	188	229	267	342	440
0,50	20	—	63	98	135	230	352
0,20		46	115	168	214	322	471
0,10		71	158	213	268	392	554
0,05		101	199	263	327	455	635
0,01		171	291	373	446	596	796
0,001		276	425	520	604	781	1010
0,50	10	—	240	393	543	910	1423
0,20		170	461	658	861	1310	1905
0,10		281	620	866	1092	1583	2237
0,05		390	801	1055	1302	1833	2544
0,01		670	1167	1496	1782	2408	3203
0,001		1090	1707	2094	2440	3152	4063

Примечание — Прочерки соответствуют случаям, не представляющим практического интереса (высокие значения рисков α и β при определении p_d).

[1)] Значения, соответствующие минимальному числу испытателей, необходимому для реализации теста парного сравнения с установленным уровнем чувствительности, определяются величинами p_d, α и β. В таблице А.5 находят выбранное значение p_d и графу, соответствующую выбранному значению β. Минимально необходимое число испытателей определяют в строке, соответствующей выбранному значению α.

[2)] Указанные в данной таблице значения рассчитаны по точной формуле биномиального распределения с параметром p = 0,5 для n откликов, благодаря программе, приведенной в [6].

Приложение В
(справочное)

Примеры

В.1 Пример 1 — Тест парного сравнения с использованием одностороннего критерия, проводимый с целью подтвердить, что между двумя образцами продукта существует различие по интенсивности одной из его органолептических характеристик

В.1.1 Исходные сведения

В соответствии с некоторыми пожеланиями потребителей были проведены изменения в технологии производства бисквитов, которые должны были обеспечивать более хрустящий продукт, чем обычный ранее выпускавшийся. Перед проведением масштабного, с участием потребителей мероприятия по определению предпочтительности новой модификации продукта, руководство фирмы решило заранее убедиться, что новшество действительно принесло желаемый эффект. Руководство фирмы заинтересовано в том, чтобы свести к минимуму риск получения заключения «различие имеется», в то время как фактически оно отсутствует. С другой стороны, поскольку представляется возможным использовать иные, чем предложенные, технологические решения, руководство фирмы готово согласиться с высокой степенью риска не обнаружить разницу между продуктами, в то время как она фактически имеет место.

В.1.2 Цель испытаний

Цель испытаний — подтвердить, что новый продукт действительно является более хрустящим. Следовательно, это случай использования одностороннего критерия.

В.1.3 Число испытателей

Для того чтобы руководство фирмы имело достаточно высокую уверенность в том, что утверждение об улучшении качества продукта не оказалось ошибочным, руководитель *органолептических* испытаний предлагает выбрать значение допустимого α-риска, равным 0,05, процент испытателей, способных ощутить различие между продуктами p_d принять равным 30 %, и β-риска — 0,50. По таблице А.4, *приложение А*, он находит, что число испытателей должно быть не менее 30.

В.1.4 Проведение испытаний

Подготавливают 30 образцов бисквита «А» (*контрольный образец*) и 30 образцов бисквита «В» (*опытный образец*); все образцы кодируют, используя неодинаковые случайные трехзначные числа. 15 испытателям продукты предоставляют в порядке АВ, а остальным 15 испытателям — в порядке ВА. Образец бланка рабочего листа показан на рисунке Б.1.

ТЕСТ ПАРНОГО СРАВНЕНИЯ

Ф.И.О испытателя _____ Код испытателя _____ Дата _____

Инструкции:

Продегустируйте последовательно два образца продукта, начиная с того, который расположен слева от Вас. Укажите код того образца, который показался Вам более хрустящим (строка ниже).

Если у Вас нет уверенности, что один образец является более «хрустящим» по сравнению с другим, то действуйте наугад; в разделе «Комментарии» Вы можете отметить, что Ваш выбор сделан только на основе догадки.

Более «хрустящим» является образец № _____

Возможные комментарии _____

_____

_____

Рисунок В.1 — *Бланк рабочего листа*

В.1.5 Анализ и интерпретация результатов испытаний

21 испытатель отметили, что образец В является более «хрустящим». Обратившись к таблице А.1, в строке, соответствующей значению n = 30, и в графе, соответствующей α = 0,05, можно увидеть, что если имеются хотя бы 20 ответов, в которых правильно отмечено направление различия, то это достаточно для утверждения, что между двумя сравниваемыми образцами имеется значимое различие.

В.1.6 Отчет по испытаниям и заключения

Руководитель испытаний сообщает, что группе испытателей (n = 30, x = 21) *опытный* образец представляется более «хрустящим» при 5 %-ном уровне значимости. Поэтому можно изготовить бисквит по новой технологии для того, чтобы предоставить его для испытаний на предпочтение с участием потребителей.

В.2 Пример 2 — Тест парного сравнения с использованием одностороннего критерия, проводимый с целью подтвердить, что между двумя образцами продукта существует подобие по интенсивности оговоренной органолептической характеристики

В.2.1 Исходные сведения

Изготовителю известно, что продукт может содержать небольшие количества ингредиента, придающего ему посторонний травянистый флейвор. Поэтому он хотел бы определить то максимально допустимое количество этого ингредиента, когда разница во флейворе по сравнению с продуктом, не содержащим этого ингредиента (Т), едва-едва заметна и, следовательно, не будет чревата нежелательными последствиями.

В.2.2 Цель испытаний

Цель испытаний заключается в том, чтобы определить максимально допустимое количество данного ингредиента, когда разница во флейворе по сравнению с продуктом, не содержащим этого ингредиента, практически незаметна и не будет иметь нежелательных последствий.

В.2.3 Число испытателей

Изготовитель хочет быть вполне уверен в правильности установленных норм на содержание ингредиента, ответственного за травянистый посторонний флейвор продукта. Таким образом, в данном тесте риск не обнаружить наличие постороннего флейвора (β-риск) должен быть настолько низким, насколько это возможно. α-риск ошибочного заключения о наличии различия во флейворе продуктов, в то время как это различие отсутствует, имеет меньшую практическую важность, так как ошибочное заключение о наличии разницы во флейворе может стать причиной только ужесточения контроля. Поэтому устанавливают: $\beta = 0{,}05$, $\alpha = 0{,}50$, а процент испытателей, ощущающих разницу во флейворе p_d, принимают равным 20 %. По таблице А.4, *приложение А,* изготовитель находит, что потребуется участие в испытаниях по крайней мере 67 испытателей. Однако обратившись к таблице А.3, *приложение А,* он определит, что для выбранных значений β и p_d требуется участие минимум 78 испытателей, чтобы можно было использовать эту таблицу (в случаях, когда число испытателей меньше 78, максимальное число предлагаемых правильных или согласующихся ответов меньше, чем то, которое получается, если ответы выбирают наугад, т.е. $n/2$, и поэтому оно не приведено в таблице). В связи с этим изготовитель решает привлечь к испытаниям 78 испытателей.

В.2.4 Проведение испытаний

Принимая во внимание результаты предварительных испытаний и ранее известные сведения, определяют предельное (контрольное) значение концентрации С интересующего вещества. Готовят два раствора и каждый из них разливают в 78 пластиковых стаканчиков, которые кодируют, используя неодинаковые случайные числа.

39 испытателям образцы растворов подают в порядке ТС, а другим 38 испытателям образцы подают в порядке СТ.

Образец бланка рабочего листа показан на рисунке В.1, но в данном случае перед испытателем ставится задача сравнить образцы с точки зрения ощущения «травянистости».

В.2.5 Анализ и интерпретация результатов испытаний

78 испытателей принимают участие в испытаниях. Сорок один испытатель указал на образец С как на раствор, имеющий более ощутимый травянистый флейвор Из таблицы А.3 можно видеть, что для n = 78 и для p_d = = 20 % при $\beta = 0{,}05$ максимальное число, которое необходимо для заключения о том, что образцы подобны друг другу, составляет 39. Поскольку количество полученных при испытаниях правильных ответов было больше, чем это критическое значение, то руководитель испытаний вправе заявить, что не представляется возможным заключить, что между образцами *продукта* наблюдается подобие.

В.2.6 Отчет по испытаниям и заключения

Руководитель испытаний сообщает, что концентрация С используемого ингредиента слишком велика, чтобы можно было объявить, что она соответствует предельно допустимому значению, и рекомендует провести новый тест с использованием более низкой концентрации ингредиента.

В.3 Пример 3 — Тест парного сравнения с использованием двустороннего критерия, имеющий целью подтвердить, что между образцами продукта существует различие по интенсивности органолептической характеристики

В.3.1 Исходные сведения

Производитель супов хочет определить, какая из двух вкусовых добавок на основе поваренной соли обеспечивает более выраженный соленый вкус продукта. Новая вкусовая добавка была выбрана благодаря поступившей информации, что ее можно использовать в меньшей концентрации и поэтому потребуются меньшие затраты на сырье (цена за 1 кг обоих видов добавок одинакова). Если не будет обнаружено существенного различия, то не вызовет затруднений провести испытания продукта, изготовленного с использованием других ингредиентов.

В.3.2 Цель испытаний

Цель испытаний — определить, какая из двух вкусовых добавок обеспечивает более выраженный соленый вкус продукта (при использовании обеих добавок в равной концентрации). Здесь имеет место случай теста с использованием двустороннего критерия.

В.3.3 Число испытателей

Руководитель *органолептических* испытаний рассчитывает, что проведение теста позволит ему быть на 95 % уверенным, что высокая доля испытателей способна ощутить различие между продуктами. Поэтому принимается: $\alpha = 0{,}05$, $p_d = 50$ %. Однако если ошибочно будет заключено, что различие между продуктами не существует, то последствием этого будут дополнительные расходы на испытания с другими вкусовыми добавками. Поэтому руководитель испытаний назначает $\beta = 0{,}10$. Обратившись к таблице А.5, *приложение А*, находят, что потребуется по крайней мере 42 испытателя, и решают пригласить для участия в тесте 44 испытателя.

В.3.4 Проведение испытаний

Готовят два образца супа — А и В, различающиеся между собой единственно тем, что в них использованы разные добавки, обеспечивающие соленый вкус. Пробы образцов подают горячими в фарфоровых чашах, которые кодированы с применением неодинаковых случайных чисел. Для 22 испытателей пробы подают в порядке АВ, для остальных 22 испытателей — в порядке ВА. Форма бланка рабочего листа приведена на рисунке В.2.

ТЕСТ ПАРНОГО СРАВНЕНИЯ

Код теста: 845—2003

Испытатель № 14 Ф.И.О. _____ Дата _____

Инструкции:

Попробуйте на вкус два образца супа, начиная с левого от Вас. Выберите из них тот образец, который имеет более соленый вкус, отметив крестиком соответствующий квадратик:

842 ☐ 376 ☐ Комментарии _____

Если Вы имеете какие-либо замечания, разъясняющие, почему Вы сделали тот или иной выбор, то Вы можете изложить их в разделе Комментарии.

Рисунок В.2 — *Бланк рабочего листа*

В.3.5 Анализ и интерпретация результатов испытаний

44 испытателя принимали участие в испытаниях. 32 испытателя указали на образец А как более соленый, а 12 испытателей указали на образец В как более соленый. Обратившись к таблице А.2, *приложение А*, увидим, что при $n = 44$ и $\alpha = 0{,}05$ наибольшее из двух чисел, соответствующих числу испытателей, отдавших свой голос либо за образец А, либо за образец В, должно быть больше или равно 29 для того, чтобы можно было заявить, что два образца супа действительно различаются по степени солености. Данные таблицы А.2, *приложение А,* также свидетельствуют, что если наибольшее из полученных чисел не ниже чем 31, то можно утверждать, что существует различие между образцами супа даже при 1 %-ном уровне значимости.

В.3.6 Отчет по испытаниям и заключения

Руководитель органолептических испытаний сообщает в протоколе испытаний, что вкусовая добавка А обеспечивает более выраженный соленый вкус, чем вкусовая добавка В, при 1 %-ном уровне значимости. Таким образом, в будущем в производстве должна использоваться вкусовая добавка А.

В.4 Пример 4 — Тест парного сравнения с использованием двустороннего критерия, проводимый для подтверждения того, что два образца продукта подобны друг другу по данной характеристике

В.4.1 Исходные сведения

Изготовитель пластмасс, используемых в производстве автомашин для изготовления щитков, хочет из экономических соображений заменить обычный смазочный ингредиент новым, но не имеет целью изменить в бóльшую или меньшую сторону степень скользкости поверхности изделий.

В.4.2 Цель испытаний

Цель испытаний заключается в том, чтобы определить — при условии равенства концентраций нового и прежнего смазочного ингредиента — будет ли в результате новшества обеспечена та же «скользкость поверхности» изделий, что и в случае изделий, изготовленных по традиционной технологии.

В.4.3 Число испытателей

Изготовитель хочет надежно убедиться, что новый ингредиент обеспечивает тот же уровень достигаемой «скользкости поверхности» изделий, что и при использовании обычного смазочного ингредиента. Итак, в данном

тесте риск невыявления различия в скользкости поверхности изделий должен быть настолько низок, насколько это только возможно. α-риск ошибочного заключения о том, что существует различие в качестве изделий, в то время как это различие отсутствует, не играет значительной роли, потому что будет иметь последствием только то, что будет продолжать использоваться более дорогостоящий привычный смазочный ингредиент. Поэтому β-риск принимается равным 0,05, α-риск — *равным* 0,10, а процент испытателей, обнаруживающих различие между продуктами p_d, равным 30 %. Обратившись к таблице А.5, *приложение А*, находят, что для проведения теста необходимы по крайней мере 119 испытателей. Для того чтобы облегчить организацию работы, сбалансировав число испытателей, решено привлечь к проведению теста 120 человек.

В.4.4 Проведение испытаний

Готовят два вида пластиковых пластин А и В; их единственное различие состоит в том, что при производстве пластика используют разный смазочный ингредиент, что может иметь следствием получение различной степени скользкости поверхности изготовляемой продукции. По две пластины предоставляют испытателям в емкостях, имеющих коды, в виде неодинаковых случайных чисел. Для 60 испытателей образцы продукции предоставляют в порядке АВ, а для остальных 60 испытателей — в порядке ВА. Образец бланка рабочего листа приведен на рисунке В.3.

ТЕСТ ПАРНОГО СРАВНЕНИЯ

Код теста: 845—2003

Ф.И.О испытателя _____ Код испытателя _____ Дата _____

Инструкции:

Вам предлагается сопоставить два образца по степени скользкости их поверхности. Более скользкой считается поверхность пластины, которая не оказывает сопротивления при касательном движении руки по поверхности изделия. Пожалуйста, протестируйте оба образца изделия, начав с того образца, который находится слева от Вас, соблюдая следующие правила: двигайте руку слева направо, плашмя, оказывая небольшое давление.

Выберите тот из образцов, поверхность которого показалась Вам более скользкой, и укажите на него, поставив крестик в соответствующем квадратике:

192 ☐ 526 ☐

Рисунок В.3 — *Бланк рабочего листа*

В.4.5 Анализ и интерпретация результатов испытаний

В испытаниях принимали участие 120 приглашенных испытателей. Из них 67 испытателей решили, что образец продукции А имеет более скользкую поверхность, чем образец продукции В, а 53 испытателя указали на образец В как более скользкий. В таблице А.3, *приложение А,* для $n = 120$, $\beta = 0{,}05$ и $p_d = 30$ % находим число 68. Наибольшее значение, полученное в результате теста, составляет 67; это меньше, чем 68, и поэтому можно заключить, что имеет место подобие между образцами изделий при 95 %-ном уровне значимости и значении p_d, выбранном при организации теста.

В.4.6 Отчет по испытаниям и заключения

Руководитель сенсорных испытаний сообщает в отчете испытаний, что новый смазочный ингредиент обеспечивает уровень скользкости поверхности продукта, подобный тому, какой достигается при использовании традиционного смазочного ингредиента, и поэтому может быть использован в будущем при производстве продукции.

В.5 Пример 5 — Доверительные интервалы в методе парного сравнения

В.5.1 Исходные сведения

Если потребуется, аналитик может рассчитать доверительный интервал для доли испытателей, действительно способных различать два образца продукта (с поправкой на случайность, как говорится). Приведенный ниже способ расчета основан на приближении к нормальному распределению и состоит в следующем:

x — число ответов в предполагаемом направлении (односторонний критерий) или число согласующихся ответов (двусторонний критерий);

n — общее число ответов;

p_c — доля ответов в предполагаемом направлении или доля согласующихся ответов, равная x/n;

$\hat{p}_c$ — доля испытателей, чувствующих различие, $\hat{p}_d = 2\,p_c - 1$;

s_d — (стандартное отклонение для $\hat{p}_d$) $= 2\sqrt{p_c(1-p_c)/n} = 2\sqrt{(nx - x^2)/n^3}$;

$\hat{p}_d + z_\alpha s_d$ — верхний доверительный предел;

$\hat{p}_d - z_\alpha s_d$ — нижний доверительный предел,

где s_d — критическая величина для центрированного и нормированного нормального распределения, значения которой указаны в таблице В.1.

Таблица В.1

Уровень значимости, %	Тест парного сравнения	
	z_α в случае использования одностороннего критерия	z_α в случае использования двустороннего критерия
80	0,84	1,28
90	1,28	1,64
95	1,64	1,96
99	2,33	2,58
99,9	3,10	3,29

В.5.2 Анализ и интерпретация результатов

В примере 3 (метод парного сравнения, двусторонний критерий) получены следующие результаты теста: n = = 44, x = 32, α = 0,05, β = 0,10.

Из этого следует, что:

$p_c = x/n = 32/44 = 0,73$;

$\hat{p}_d$ (доля испытателей, способных почувствовать различие) = $2p_c - 1 = 2 \cdot 0,73 - 1 = 0,45$;

s_d (стандартное отклонение для $\hat{p}_d$) = $2\sqrt{(nx - x^2)/n^3} = 2\sqrt{(32 \cdot 44 - 32^2)/44^3} = 0,134$;

верхний 95 %-ный доверительный предел равен $\hat{p}_d + z_\alpha s_d = 0,45 + 1,96 \cdot 1,34 = 0,71$;

нижний 95 %-ный доверительный предел равен $\hat{p}_d - z_\alpha s_d = 0,45 - 1,96 \cdot 1,34 = 0,19$.

Поэтому руководитель испытаний может быть на 95 % обнадежен, что по крайней мере 19 % членов популяции способны различать образцы продукта, изготовленного по старой и по новой технологиям. Этот результат согласуется с заключением, приведенным в примере 3 (где сказано, что образец А представляется более соленым), поскольку расчет показал, что доверительный интервал, с одной стороны, не включает в себя нулевое значение и, с другой стороны, расположен за пределом p_d, равного 50 %, который был установлен первоначально в качестве фиксированного предела.

В примере 4 (метод парного сравнения, двусторонний критерий) получены следующие результаты теста:

n = 120, x = 67, α = 0,10, β = 0,05.

Отсюда следует, что:

$p_c = x/n = 67/120 = 0,56$;

$\hat{p}_d$ (доля испытателей, способных почувствовать различие) = $2p_c - 1 = 2 \cdot 0,56 - 1 = 0,12$;

s_d (стандартное отклонение для $\hat{p}_d$) = $2\sqrt{(nx - x^2)/n^3} = 2\sqrt{(120 \cdot 67 - 67^2)/120^3} = 0,09$;

верхний 95 %-ный доверительный предел = $\hat{p}_d + z_\alpha s_d = 0,12 + 1,96 \cdot 0,09 = 0,29$;

нижний 95 %-ный доверительный предел = $\hat{p}_d - z_\alpha s_d = 0,12 - 1,96 \cdot 0,09 = 0,06$.

Поэтому руководитель испытаний может быть на 95 % обнадежен, что фактически доля совокупности испытателей, способных различать образцы продукта, не превышает 29 %.

Рассматриваемые вместе доверительные интервалы разрешают 5 %-ную вероятность ошибки как для верхнего, так и для нижнего предела, поэтому руководитель испытаний может быть обнадежен на 90 % в том, что фактическая доля людей, различающих образцы продукта, находится в пределах от 0 до 29 %. Это меньше, чем первоначально назначенный предел ($\hat{p}_d$ = 30 %). Поэтому руководитель испытаний может заключить, что в данном случае между образцами продукта нет заметного различия.

Библиография

[1] Künert, J. and Meyners, M. On the triangle test with replications. Food Quality and Preference, 10, 1999

[2] Künert, J. On repeated difference testing. Food Quality and Preference, 12, 2001, pp. 385—391

[3] Ennis, D.M. and Bi, J. The Beta-Binomial Model: Accounting for Inter-trial Variation in Replicated Difference and Preference Tests. Journal of Sensory Studies, 13, № 4, 1998, pp. 389—412

[4] Brockhoff, P.B. and Schlich, P. Handling replications in discrimination tests. Food Quality and Preference, 9, № 5, 1998, pp. 303—312

[5] ISO 6658:2005 Sensory analysis — Methodology — General guidance

[6] Schlich, P. Risk Tables for Discrimination Tests. Food Quality and Preference, 4, 1993, pp. 141—151

[7] MacRae, A. Confidence intervals for triangle test. Food Quality and Preference, 6, 1995, pp. 61—67

[8] Frijters, J.E.R. Three-Stimulus Procedure in Olfactory Psycholophysics: An Experimental Comparison of Thurston-Ura and Three-Alternative Forced-Choice Models of Signal Detection Theory. Perception and Psychophysics, 28, № 5, 1980, pp. 390—397

[9] ISO 3534-1:2006 Statistics — Vocabulary and symbols — Part 1: General statistical terms and terms used in probability

УДК 633.11:006.354 ОКС 67.240 Н91

Ключевые слова: органолептический анализ, органолептические испытания, испытания на различие, испытания на подобие, метод парного сравнения

Редактор *Р.Г. Говердовская*
Технический редактор *В.Н. Прусакова*
Корректор *Т.И. Кононенко*
Компьютерная верстка *Л.А. Круговой*

Сдано в набор 25.06.2009. Подписано в печать 08.07.2009. Формат 60 × 84 1/8. Бумага офсетная. Гарнитура Ариал. Печать офсетная. Усл. печ. л. 2,79. Уч.-изд. л. 2,50. Тираж 248 экз. Зак. 410.

ФГУП «СТАНДАРТИНФОРМ», 123995 Москва, Гранатный пер., 4.

gost.e-ecolog.ru gost.e-ecolog.ru

Набрано во ФГУП «СТАНДАРТИНФОРМ» на ПЭВМ.
Отпечатано в филиале ФГУП «СТАНДАРТИНФОРМ» — тип. «Московский печатник», 105062 Москва, Лялин пер., 6.

gost.e-ecolog.ru