

**МИНОБРНАУКИ РОССИИ**

**федеральное государственное бюджетное образовательное учреждение  
высшего профессионального образования  
«Вологодский государственный университет»  
(ВоГУ)**

**Кафедра автоматики и вычислительной техники**

# **ТЕСТИРОВАНИЕ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ**

**Методические указания к выполнению  
практических работ «Оценка качества  
программных продуктов»**

**Факультет электроэнергетический**

**Направления подготовки:**

**231000.62 – Программная инженерия**

**230100.62 – Информатика и вычислительная  
техника**

**Вологда 2015**

УДК 681.3.06

Тестирование программного обеспечения: методические указания к выполнению практических работ «Оценка качества программных продуктов». - Вологда, ВоГУ, 2014. - 14 с.

Методические указания предназначены для проведения двух практических работ. В них имеется необходимый теоретический материал, включены задачи для самостоятельной работы и вопросы для самоконтроля.

Утверждено редакционно – издательским советом ВоГУ.

Составитель: Сергушичева А.П., канд.техн. наук, доц. каф. АВТ,

Рецензент: Швецов А.Н., декан ФЗДО, доктор техн. наук, профессор

# Практическая работа №1. Показатели качества программных продуктов и методы их оценки

## Основные теоретические положения

В соответствии с терминологией, принятой Международной организацией по стандартизации (ISO), **качество** – это свойство объекта, относящееся к потребностям и ожиданиям и направленное на их удовлетворение. Под объектом подразумевается все, что мыслимо и ощутимо: продукты, процессы, системы. Качество описывается рядом свойств, количественные характеристики которых называются показателями. Показатели качества включают показатели назначения, надежности, технологичности, безопасности, эстетические, эргономические и другие показатели. Показатели качества программных продуктов (ПП) весьма разнообразны. Так, одними из основных показателей качества являются претензии пользователей. В силу уникальности программных продуктов и организаций-разработчиков, качество и его показатели будут в каждом конкретном случае индивидуальными.

**Показатели качества программных продуктов** устанавливают ГОСТ 28195 «Оценка качества программных средств. Общие положения» и ГОСТ Р ИСО/МЭК 9126-93 «Информационная технология. Оценка программной продукции. Характеристика качества и руководства по их применению».

ГОСТ Р ИСО/МЭК 9126-93 определяет **качество (quality)** как весь объем признаков и характеристик *продукции или услуги* (а **качество программного обеспечения (software quality)** как весь объем признаков и характеристик *программной продукции*), который относится к их способности удовлетворять установленным или предполагаемым потребностям. Для описания показателей качества в стандарте применяются следующие **термины**:

*Оценка (assessment)* - действие по применению конкретного задокументированного критерия оценки к конкретному программному модулю, пакету или продукции с целью обусловленной приемки или выпуска программного модуля, пакета или продукции.

*Признаки (показатели) (features)* - признаки, определяющие свойства программной продукции, которые могут быть отнесены к характеристикам качества. Примеры признаков включают длину маршрута, модульность, структуру программы и комментарии.

*Уровень качества функционирования (level of performance)* - степень, в которой удовлетворяются потребности, представленные конкретным набором значений для характеристик качества.

*Измерение (measurement)* - действие по применению показателя качества программного обеспечения к конкретной программной продукции.

*Ранжирование (рейтинг) (rating)* - действие по отнесению измеренного значения к соответствующему уровню ранжирования. Используется для определения уровня ранжирования программного обеспечения по конкретной характеристике качества.

*Уровень ранжирования (rating level)* - диапазон значений в масштабе, позволяющем классифицировать (ранжировать) программное обеспечение в соответствии с установленными или предполагаемыми потребностями. Соответствующие уровни ранжирования могут быть связаны с различными представлениями о качестве, то есть для пользователей, руководителей или разработчиков.

*Критерий оценки качества программного обеспечения (software quality assessment criteria)* - набор определенных и задокументированных правил и условий, которые используются для решения о приемлемости общего качества конкретной программной продукции. Качество представляется набором установленных уровней, связанных с программной продукцией.

*Характеристики качества программного обеспечения (software quality characteristics)* - набор свойств (атрибутов) программной продукции, по которым ее качество описывается и оценивается. Характеристики качества программного обеспечения могут быть уточнены на множестве уровней комплексных показателей (подхарактеристик).

*Метрика качества программного обеспечения (software quality metric)* - количественный масштаб и метод, которые могут быть использованы для определения значения признака, принятого для конкретной программной продукции.

**ГОСТ Р ИСО/МЭК 9126** отражает представление пользователя о качестве ПО и предназначен для характеристик, связанных с приобретением, разработкой, эксплуатацией, поддержкой, сопровождением или проверкой программного обеспечения.

### **1.1. Основные характеристики качества**

В числе основных характеристик, которые с минимальным дублированием описывают качество ПП и образуют основу для дальнейшего уточнения и описания качества программного обеспечения, следующие:

**1. Функциональные возможности (Functionality)** - это атрибуты, описывающие набор функций и их конкретные свойства. Этот набор атрибутов характеризует возможности ПП, выполняемые для удовлетворения установленных (при наличии контракта, технического задания) или предполагаемых (должны быть установлены и формально определены какими-либо нормативными документами (стандартами, техническими условиями и пр.)) потребностей. Данная характеристика описывает свойства ПО в части полноты удовлетворения требований пользователя и в этом смысле является определяющей для потребительских свойств ПО,

Функциональные возможности определяют с помощью подхарактеристик (примитивов качества):

- *Пригодность* (Suitability) – характеризует наличие и соответствие набора функций конкретным задачам;
- *Правильность* (Accuracy) – показывает обеспечение правильности или соответствия результатов (эффектов). Например, она включает необходимую степень точности вычисленных значений.
- *Способность к взаимодействию* (Interoperability) – атрибуты программного обеспечения, относящиеся к способности его взаимодействовать с конкретными системами. - Способность к взаимодействию используется вместо совместимости для того, чтобы избежать возможной путаницы с взаимозаменяемостью.
- *Согласованность* (Compliance) – атрибуты программного обеспечения, которые заставляют программу придерживаться соответствующих стандартов, соглашений, положений законов или подобных рекомендаций.
- *Защищенность* (Security) – атрибуты программного обеспечения, относящиеся к его способности предотвращать несанкционированный доступ, случайный или преднамеренный, к программам и данным.

Вместо характеристики функциональные возможности могут применяться **показатели корректности**, которые показывают степень соответствия ПП требованиям, установленным в техническом задании, требованиям к обработке данных и общесистемным требованиям. К ним относятся:

- *Полнота реализации* – характеризует полноту реализации заданных функций ПС и достаточность их описания в программной документации;
- *Согласованность* – однозначное, непротиворечивое описание и использование тождественных объектов, функций, терминов, определений, идентификаторов и т. д. в различных частях программных документов и текста программы;
- *Логическая корректность* – функциональное и программное соответствие процесса обработки данных при выполнении задания общесистемным требованиям;
- *Проверенность* – полнота проверки возможных маршрутов выполнения программы в процессе тестирования.

**2. Надежность (Reliability)** – набор атрибутов, относящихся к способности программного обеспечения сохранять свой уровень качества функционирования при установленных программными документами условиях за установленный период времени, в том числе и в условиях возникновения отклонений в среде функционирования, вызванных сбоями технических средств, ошибками во входных данных, ошибками обслуживания и другими дестабилизирующими воздействиями. ПП не

подвержены старению и износу и ограничения надежности проявляются из-за ошибок в требованиях, проекте и реализации. Отказы из-за этих ошибок зависят от способа использования программного обеспечения и ранее выбранных версий программ. Надежность характеризуют примитивами:

- *Работоспособность* – способность программы функционировать в заданных режимах и объемах обрабатываемой информации в соответствии с программными документами при отсутствии сбоев технических средств
- *Стабильность* (Maturity) – атрибуты, показывающие частоту отказов при ошибках в программном обеспечении;
- *Устойчивость к ошибке* (Fault tolerance) (устойчивость функционирования) – атрибуты программного обеспечения, относящиеся к его способности поддерживать определенный уровень качества функционирования в случаях возникновения отклонений, вызванных сбоями технических средств, ошибками во входных данных и ошибками обслуживания. Определенный уровень качества функционирования включает возможность отказобезопасности;
- *Восстанавливаемость* (Recoverability) – атрибуты программного обеспечения, описывающие его возможности восстанавливать уровень качества функционирования и восстанавливать данные, непосредственно поврежденные в случае отказа, а также к времени и усилиям, необходимым для этого.

**3. Практичность (Usability) (удобство применения)** – набор атрибутов, относящихся к объему работ, требуемых для использования ПП и индивидуальной оценки такого использования определенным или предполагаемым кругом пользователей (в эргономике, как составные части практичности, рассматриваются так же эффективность и неэффективность). Пользователи разного уровня подготовленности предъявляют разные (часто взаимоисключающие) требования. Круг пользователей может включать операторов, конечных пользователей и косвенных пользователей, на которых влияет данное программное обеспечение или которые зависят от его использования. Практичность должна рассматриваться во всем разнообразии условий эксплуатации пользователем, включая подготовку к использованию и оценку результатов. Практичность выражают примитивами, которые характеризуют свойства ПП, способствующие его быстрому освоению, применению и эксплуатации с минимальными трудозатратами и с учетом характера решаемых задач и требований к квалификации обслуживающего персонала:

- *Понятность* (Understandability) (легкость освоения) – характеризует усилия пользователя по пониманию общей логической концепции функционирования программы в целом и ее частей и ее применимости;

- *Обучаемость* (Learnability) – атрибуты программного обеспечения, относящиеся к усилиям пользователя по обучению его применению (например оперативному управлению, вводу, выводу);
- *Доступность эксплуатационных программных документов* – характеризует понятность, наглядность и полноту описания взаимодействия пользователя с программой в эксплуатационных программных документах и соответствие процесса обработки данных и форм представления результатов характеру решаемых задач;
- *Простота использования* (Operability) (удобство эксплуатации и обслуживания) – показывает усилия пользователя по эксплуатации и оперативному управлению.

**4. Эффективность (Efficiencies)** - набор атрибутов, относящихся к соотношению между уровнем качества функционирования программного обеспечения (удовлетворения потребности пользователя в обработке данных) и объемом используемых ресурсов (экономических, вычислительных, людских) при установленных условиях. Ресурсы могут включать другие программные продукты, технические средства, материалы (например, бумага для печати, гибкие диски) и услуги эксплуатирующего, сопровождающего или обслуживающего персонала. Оценка данной характеристики также критически зависит от требований пользователя. Программное обеспечение может выглядеть неэффективным не в силу плохого кодирования, а в силу противоречивости и нереальности исходных требований. Например, требования к ПП выполнять функции на технических средствах минимальной (по объему оперативной и дисковой памяти, тактовой частоте и пр.) конфигурации компьютера противоречит требованиям о высоком быстродействии. Эффективность определяют с помощью подхарактеристик:

- *Характер изменения во времени* (Time behavior) (временная эффективность) – атрибуты программного обеспечения, отражающие способность программы выполнять заданные действия в заданный интервал времени (времена отклика и обработки, скорость выполнения его функций).
- *Характер изменения ресурсов* (Resource behavior) – определяет объем используемых ресурсов и продолжительности такого использования при выполнении функции. (*Ресурсоемкость* – минимально необходимые вычислительные ресурсы и число обслуживающего персонала для эксплуатации ПП)
- *Уровень автоматизации* – отражает уровень автоматизации процесса обработки данных с учетом рациональности функциональной структуры программы с точки зрения взаимодействия с ней пользователя и использования вычислительных ресурсов

**5. Сопровождаемость (Maintainability)** - характеризует технологические аспекты, обеспечивающие простоту устранения ошибок в

программе и программных документах и поддержания ПП в актуальном состоянии, отражает объем работ, требуемых для проведения конкретных изменений (модификаций) ПП. Изменение может включать исправления, усовершенствования или адаптацию программного обеспечения, изменения в окружающей обстановке, требованиях и условиях функционирования. Сопровождаемость выражают через примитивы:

- *Анализируемость* (Analysability) – показывает усилия, необходимые для диагностики недостатков или случаев отказов или определения составных частей для модернизации;
- *Изменяемость* (Changeability) – определяет усилия, необходимые для модификации, устранения отказа, изменения условий эксплуатации;
- *Устойчивость* (Stability) – атрибуты программного обеспечения, относящиеся к риску от непредвиденных эффектов модификации;
- *Тестируемость* (Testability) – характеризует усилия, необходимые для проверки модифицированного программного обеспечения.
- Структурность – организация всех взаимосвязанных частей программы в единое целое с использованием логических структур "последовательность", "выбор", "повторение"
- Простота конструкции – построение модульной структуры программы наиболее рациональным с точки зрения восприятия и понимания образом
- Наглядность – наличие и представление в наиболее легко воспринимаемом виде исходных модулей ПП, полное их описание в соответствующих программных документах
- Повторяемость – степень использования типовых проектных решений или компонентов, входящих в ПП

**6. Мобильность (Portability)** - набор атрибутов, относящихся к способности программного обеспечения быть перенесенным из одного окружения (организационного, технического или программного) в другое. Мобильность характеризуют примитивами:

- *Адаптируемость* (Adaptability) – показывает удобство адаптации к различным конкретным условиям эксплуатации, без применения других действий или способов, кроме тех, что предназначены для этого в рассматриваемом ПП.
- *Простота внедрения* (Installability) – атрибуты программного обеспечения, относящиеся к усилиям, необходимым для внедрения его в конкретное окружение.
- *Соответствие* (Conformance) – определяет соответствие ПП стандартам или соглашениям, относящимся к мобильности.
- *Взаимозаменяемость* (Replaceability) – характеризует простоту и трудоемкость применения ПП вместо другого конкретного программного средства в среде этого средства. Взаимозаменяемость с конкретным программным средством не



предполагает, что данное средство заменимо рассматриваемым программным средством. Она может включать атрибуты простоты внедрения и адаптируемости. Понятие было введено в качестве отдельной подхарактеристики из-за его важности.

Вместо характеристики мобильность могут применяться ***показатели универсальности***, которые показывают адаптируемость ПП к новым функциональным требованиям, возникающим вследствие изменения области применения или других условий функционирования. К ним относятся:

- *Гибкость* – возможность использования ПП в различных областях применения;
- *Мобильность* – возможность применения ПП. без существенных дополнительных трудозатрат на ЭВМ аналогичного класса;
- *Модифицируемость* – обеспечение простоты внесения необходимых изменений и доработок в программу в процессе эксплуатации.

Перечисленные характеристики могут применяться к любому виду программного обеспечения. Последним двум характеристикам обычно не уделяется достаточно внимания со стороны пользователя, так как они связаны с долгосрочным планированием развития ПП (и эксплуатирующей его организации). Улучшение сопровождаемости и мобильности может, вообще говоря, повысить стоимость программного обеспечения в настоящий момент времени, что (многократно) окупается лишь через несколько лет.

Определения характеристик и соответствующая модель процесса оценки качества, приведенные в стандарте ГОСТ Р ИСО/МЭК 9126, применимы тогда, когда определены требования для программной продукции и оценивается ее качество в процессе жизненного цикла. Данный стандарт не определяет также методы измерения, ранжирования и оценки. Существует ряд моделей качества, описанных в литературе, и применяемых на практике. Степень завершенности этих моделей, терминов и определении пока еще не позволяет включить их в стандарт.

Таким образом, стандарты, с одной стороны, требуют измерения показателей качества, а с другой – в них отсутствуют конкретные перечни показателей и методы их оценки. Также отсутствуют необходимые методические рекомендации по анализу соответствующей информации. Специальные международные и отечественные нормативные документы тоже не дают ответа на этот вопрос. В результате для всех этапов жизненного цикла программных продуктов разработчик должен самостоятельно разрабатывать комплексы показателей качества. Он также самостоятельно должен выявлять факторы, влияющие на качество. Выбор и обоснование факторов должен осуществляться комплексно, в зависимости от принятой системы показателей качества, выдвинутых гипотез о причинах изменения качества программного продукта, установленных целей статистического анализа и других причин. Всегда есть вероятность того, что не все

существенные факторы включены в рассмотрение, а значения некоторых из них неточны. Поэтому при выборе факторов большую роль играет опыт эксперта. Необходимо учитывать, что факторы оказывают различное воздействие на значения показателей качества. Согласно принципу Парето, при анализе системы существенны лишь некоторые факторы, причем 20% из них определяют 80% свойств системы. Только структурированный индивидуальный подход к выбору и обоснованию показателей и факторов обеспечивает эффективный контроль и управление качеством.

## 1.2. Методы оценки качества программного средства

При оценке качества используются все *виды измерительных шкал*:

- 1) номинальная – соответствует набору категорий;
- 2) порядка – соответствует упорядоченному набору делений шкалы;
- 3) интервалов – соответствует упорядоченной шкале с равноудаленными делениями;
- 4) отношений – соответствует упорядоченной шкале с равноудаленными делениями, оцененными в относительных единицах (относительно некоторой абсолютной величины).

Измерения показателей качества могут быть *инструментальными и безинструментальными*. В большинстве случаев предпочтение следует отдавать *инструментальным* измерениям. Однако, список показателей качества, для которых существуют средства измерения, весьма ограничен. Например, в их числе время наработки на отказ, показатели производительности (потребление ресурсов центрального процессора, потребление оперативной памяти, время выполнения запроса), потребление сетевых ресурсов. При *безинструментальных* измерениях роль «нетехнического» средства измерений играет человек (эксперт). При этом полагают, что меру соответствующего свойства он создает в своем воображении. Человек не сохраняет свои способности к оцениванию постоянными. Он учится, набирается опыта, приспосабливается, подвергается внешним влияниям, помнит и забывает. И, тем не менее, экспертные оценки для многих показателей качества технической и другой продукции является единственно возможными и применяется достаточно широко. Экспертная оценка выставляется по итогам изучения программ и документации на ПП, в ходе которого эксперт составляет свое мнение об обладании программным продуктом требуемым примитивом качества. При этом группа экспертов должна руководствоваться конкретизацией этого примитива и указанием о способе его оценки, сформулированными в спецификации качества аттестуемого ПС.

**Эксперт** (от лат. *expertus* — опытный) — специалист, приглашаемый или нанимаемый за вознаграждение для выдачи квалифицированного заключения или суждения по вопросу, рассматриваемому или решаемому

другими людьми, менее компетентными в этой области. Вполне очевидно, что во всех без исключения случаях экспертиза должна проводиться грамотными, высококвалифицированными, вполне компетентными в рассматриваемых вопросах и достаточно опытными специалистами. Компетентность эксперта в отношении объекта исследования — профессиональная компетентность, а в отношении методологии принятия экспертного решения исследуемой задачи — это экспертная компетентность.

**Экспертная комиссия.** С целью повышения достоверности, точности, надежности и воспроизводимости экспертных оценок экспертизу осуществляют путем принятия группового решения. Создается экспертная комиссия, состоящая из экспертной и рабочей групп. Члены рабочей группы осуществляют сбор и обработку информации, подготовку документов и т.п. В экспертную группу включаются высококвалифицированные и специально подготовленные работники. Желательно, чтобы для оценок однотипной продукции экспертная комиссия формировалась из постоянных экспертов и членов рабочей группы. Это связано с тем, что в процессе работы относительно постоянной комиссии накапливается опыт работы, происходит обучение ее членов, вырабатываются общие подходы и принципы, а это повышает эффективность работы экспертной комиссии.

Число экспертов, входящих в группу, зависит от требуемой точности средних оценок и составляет от семи до двадцати человек (и более при социологических исследованиях, например, при опросе потребителей продукции). Желательно специальное предварительное обучение и обязателен инструктаж экспертов. На завершающем этапе формирования экспертной группы целесообразно протестировать экспертов, выявить их самооценку (экспертные группы с высокими показателями самооценки экспертов ошибаются в меньшей степени) и взаимооценку (если они уже работали вместе), оценить надежность и согласованность мнений.

Тестированием устанавливают компетентность и профпригодность экспертов. При этом им предлагается решить одну или несколько задач, подобных реальным, с известными (не экспертам) ответами. Для определения самооценки эксперты отвечают на вопросы специально составленной анкеты. При наличии опыта совместной работы эксперты оценивают друг друга (в баллах). Показателем (степенью) надежности эксперта называют отношение числа случаев, когда мнение его совпало с результатами экспертизы, к общему числу экспертиз, в которых он участвовал. За меру согласованности мнений экспертов принимается так называемый коэффициент конкордации.

**Коэффициент конкордации** Кендалла (от лат. *concordare* - привести в соответствие, упорядочить) определяют по формуле:

$$W = \frac{12S}{n^2(m^3 - m)};$$

где  $S$  — сумма квадратов отклонений суммы рангов каждого объекта экспертизы от среднего арифметического рангов;  $n$  — число экспертов;  $m$  — число объектов экспертизы. В зависимости от степени согласованности мнений экспертов коэффициент конкордации может принимать значения от 0 (при отсутствии согласованности) до 1 (при полном единодушии). Если степень согласованности мнений экспертов оказывается неудовлетворительной, принимают специальные меры для ее повышения. Сводятся они, в основном, к проведению тренировок с обсуждением результатов и разбором ошибок

**Пример 1.1.** Необходимо определить степень согласованности мнения пяти экспертов, результаты ранжирования которыми семи объектов приведены в таблице 1.1.

*Таблица 1.1 - Данные для оценки согласованности мнений пяти экспертов*

Номер объекта экспертизы	Оценка эксперта					Сумма рангов	Отклонение от среднего	Квадрат отклонения
	1	2	3	4	5			
1	4	6	4	4	3	21	1	1
2	3	3	2	3	4	15	-5	25
3	2	2	1	2	2	9	11	121
4	6	5	6	5	6	28	8	64
5	1	1	3	1	1	7	-13	169
6	5	4	5	6	5	25	5	25
7	7	7	7	7	7	35	15	225

**Решение:**

1) Оцениваем среднеарифметическое число рангов:

$$Q_{cp} = (21 + 15 + 9 + 28 + 7 + 25 + 35)/7 = 20.$$

2) оцениваем сумму квадратов отклонений от среднего:  $S = 630$ .

3) Определяем величину коэффициента конкордации:

$$W = 12 * 630 / 25 * (343 - 7) = 0,9.$$

**Вывод:** Степень согласованности мнений экспертов можно считать удовлетворительной.

**Экспертные методы** оценивания показателей качества применяют тогда, когда использование технических средств измерения, невозможно, сложно или экономически не оправдано. Они основаны на использовании обобщенного опыта и интуиции специалистов-экспертов. Методы экспертных оценок представляют комплекс логических и математико-статистических методов и процедур, связанных с деятельностью экспертов по переработке необходимой для анализа и принятия решений информации. Организует процедуру опроса экспертов, анализирует полученные результаты и составляет заключение рабочая группа экспертной комиссии.

**Опрос эксперта** может проводиться путем анкетирования, интервьюирования, открытого или тайного голосования и т.п.

По тому, в какой форме эксперты выражают свое мнение, т.е. по способу проведения экспертизы, различают методы:

- экспертный (непосредственное измерение);
- ранжирования;
- попарного сопоставления;
- последовательного сопоставления.

#### 1.2.1. Экспертный (непосредственное измерение) метод

При **непосредственных измерениях** экспертным методом значения показателей качества определяются сразу в установленных единицах (в баллах, нормо-часах, рублях, единицах условного топлива и т.д.). Балльные оценки даются экспертами непосредственно или получаются в результате формализации процесса оценки. Назначение балльных оценок производится экспертами независимо друг от друга или в процессе обсуждения. Возможен такой сценарий экспертизы: сначала эксперты выставляют оценки; затем следует открытое обсуждение этих оценок; после чего эксперты вновь независимо друг от друга дают оценки каждому параметру качества. Впоследствии по скорректированным индивидуальным оценкам рассчитывают общую экспертную оценку.

Для оценки показателей качества обычно используют пяти-, семи- или десятибалльную шкалы. Однако эта оценка может производиться и по двухбалльной системе ("обладает" - "не обладает"). Обобщенный показатель качества, определяемый экспертным методом по балльной системе исчислений, находят как среднее арифметическое значение оценок, поставленных всеми экспертами. Если при экспертизе качества оценку (опрос) проводят в несколько туров, то в этом случае значение показателя качества определяют как среднеарифметическое значение оценок, полученных в каждом туре опроса экспертов.

Такие измерения могут проводиться как по шкале отношений, так и по шкале интервалов или шкале порядка. Измерения по шкале отношений требуют наличия эталонов. Критерии, по которым осуществляется экспертиза качества, подразделяются на общие и конкретные. К общим критериям относятся сложившиеся в обществе ценностные ориентиры, представления и нормы. Конкретные критерии для эксперта — это реальные требования к качеству продукции данного вида, установленные в нормативно-технических и других обязательных для исполнения документах, базовые значения показателей качества, характеристики реально существующих высококачественных изделий, изготавливаемых в стране или за рубежом.

Экспертным методом, например, часто пользуются при выборе техники, представленной несколькими предприятиями на тендерные конкурсы (торги).

### 1.2.2. Метод ранжирования

**Ранжирование** состоит в расстановке объектов измерений или показателей в порядке их предпочтения, по важности или весомости. Место, занятое при такой расстановке, называется рангом. Чем выше ранг, тем предпочтительней объект, весомее, важнее показатель. Порядок построения ранжированного ряда следующий:

1. Все объекты оценки (изделия, свойства) нумеруются произвольно;
2. Эксперты ранжируют объекты по шкале порядка: численное значение ранга в ряду возрастающей шкалы порядка увеличивается от 1 до  $n$  ( $n$  — количество оцениваемых объектов);
3. Ранжированные ряды объектов, составленные экспертами, сопоставляются, определяются суммы рангов каждого из объектов экспертной оценки;
4. На основании полученных сумм рангов строят обобщенный ранжированный ряд.
5. Обобщенные экспертные оценки качества рассматриваемых объектов экспертизы, т.е. коэффициенты их весомости, рассчитываются по формуле:

$$g_i = \frac{\sum_{j=1}^n G_{i,j}}{\sum_{i=1, j=1}^{n,m} G_{i,j}},$$

где  $n$  — количество экспертов,  $m$  — число «взвешиваемых» показателей,  $G_{i,j}$  — коэффициент весомости  $j$ -го показателя (ранг) в баллах, данный  $i$ -м экспертом. Коэффициенты весомости позволяют не только указать, какой объект лучше или хуже других, но и на сколько.

**Пример 1.2.** Результаты ранжирования пятью экспертами семи объектов экспертизы следующие:

Первый эксперт:  $Q5 < Q3 < Q2 < Q1 < Q6 < Q4 < Q7$ ;

Второй эксперт:  $Q5 < Q3 < Q2 < Q6 < Q4 < Q1 < Q7$ ;

Третий эксперт:  $Q3 < Q2 < Q5 < Q1 < Q6 < Q4 < Q7$ ;

Четвертый эксперт:  $Q5 < Q3 < Q2 < Q1 < Q4 < Q6 < Q7$ ;

Пятый эксперт:  $Q5 < Q3 < Q1 < Q2 < Q6 < Q4 < Q7$ .

Построить обобщенный ранжированный ряд и определить весомость членов ряда.

**Решение:**

1) Сумма рангов:  $Q1$  равна  $4+6+4+4+3=21$

$Q2$  равна  $3+3+2+3+4=15$

Q3 равна  $2+2+1+2+2=9$   
 Q4 равна  $6+5+6+5+6=28$   
 Q5 равна  $1+1+3+1+1=7$   
 Q6 равна  $5+4+5+6+5=25$   
 Q7 равна  $7+7+7+7+7=35$

Обобщенный ранжированный ряд  $Q5 < Q3 < Q2 < Q1 < Q6 < Q4 < Q7$ .

$$2) \sum_{i=1, j=1}^{n, m} G_{i, j} = 21 + 15 + 9 + 21 + 7 + 25 + 35 = 140$$

Весовые коэффициенты:

$g_1 = 21/140 = 0,15$ ;  $g_2 = 15/140 = 0,11$ ;  $g_3 = 9/140 = 0,06$ ;  $g_4 = 28/140 = 0,2$ ;  
 $g_5 = 7/140 = 0,05$ ;  $g_6 = 25/140 = 0,18$ ;  $g_7 = 35/140 = 0,25$ ;

**Вывод:** Лучшим по качеству признан продукт №5, на 0,01 отстает от него продукт №3.

### 1.2.3. Метод попарного сопоставления

Попарное сопоставление самое простое и наиболее оправданное с психологической точки зрения. Оно лежит в основе любого выбора. Измеряемые размеры сравниваются друг с другом, и для каждой пары результат сравнения выражается в форме «больше – меньше», «лучше – хуже». Сравнить две величины между собой легче, чем сразу определить их место на шкале порядка. Иногда учитывается равноценность. Ранжирование производится на основе результатов попарного сопоставления.

Порядок попарного сопоставления следующий:

1. Составляется таблица, в которую эксперт заносит результаты сопоставления рассматриваемых объектов. При этом каждый  $i$ -й объект сопоставляется с другими  $j$ -ми объектами сравнения. Если  $i$ -й объект признается качественнее  $j$ -го, то это обозначается цифрой 1, противоположная оценка обозначается  $-1$ , а равнокачественные объекты отмечаются в таблице цифрой 0 (ноль);

2. Данные о предпочтениях эксперта суммируются и строится ранжированный ряд;

3. При групповой экспертизе строится обобщенный ранжированный ряд и рассчитываются весовые коэффициенты.

**Пример 1.3.** Результат дегустации пищевых продуктов, обозначенных номерами от 1 до 6 представлен в таблице 1.2, где предпочтению  $i$  –го продукта над  $j$  –м соответствует 1, а противоположному отношению – 0. Требуется расставить продукты по качеству.

Таблица 1.2 - Результат дегустации пищевых продуктов

Продукт <b>i</b>	<b>j</b>						Итого
	1	2	3	4	5	6	
1		1	0	1	1	1	4

2	0		0	1	1	1	3
3	1	1		1	1	1	5
4	0	0	0		0	0	0
5	0	0	0	1		0	1
6	0	0	0	1	1		2

**Решение:** Ранжированный ряд имеет вид:  $Q_3 > Q_1 > Q_2 > Q_6 > Q_5 > Q_4$   
или  $Q_4 < Q_5 < Q_6 < Q_2 < Q_1 < Q_3$

**Вывод:** Лучшим по качеству признан продукт №3, худшим – продукт №4.

**Пример 1.4.** Результат попарного сопоставления мастерства певцов, выступавших на конкурсе вокалистов, представлен в таблице 1.3, где предпочтению  $i$  –го солиста над  $j$  –м соответствует 1, противоположному отношению –1, равноценное выступление – 0. Требуется определить победителей конкурса.

Таблица 1.3 - Результат попарного сопоставления мастерства певцов

Певец	j						Итого
i	1	2	3	4	5	6	
1	0	1	-1	1	1	1	3
2	-1	0	-1	0	1	0	-1
3	1	1	0	1	1	1	5
4	-1	0	-1	0	1	0	-1
5	-1	-1	-1	-1	0	-1	-5
6	-1	0	-1	0	1	0	-1

**Решение:** Ранжированный ряд имеет вид:  $Q_3 > Q_1 > Q_2 = Q_4 = Q_6 > Q_5$ .

**Вывод:** победителем конкурса стал певец под №3, второе место у певца №1, третье место поделили певцы с номерами 2,4,6.

Приведенные в примерах 1.2 и 1.3 таблицы являются избыточными. При попарном сопоставлении достаточно данных, приведенных в таблицах по одну сторону от диагонали. Предпочтение при этом выражается указанием номера предпочтительного объекта так, как это показано в таблице 1.4.

Таблица 1.4 - Результат попарного сопоставления объектов экспертизы

№ объекта экспертизы	1	2	3	4	5	6
1	X	1	3	1	1	1
2		X	3	2	2	2
3			X	3	3	3
4				X	5	6



5					X	6
6						X

Балл  $j$  - го объекта или весомость  $j$  - го показателя рассчитываются по формуле:

$$G_{i,j} = \frac{F_{i,j}}{C}$$

где  $F_{i,j}$  – частота предпочтения  $i$  - м экспертом  $j$  - го объекта экспертизы;  $C$  - общее число суждений одного эксперта, связанное с числом объектов экспертизы  $m$  (числом измеряемых показателей или коэффициентов весомости) соотношением:

$$C = \frac{m(m-1)}{2}$$

**Пример 1.5.** Предположим для простоты, что пять экспертов, выразили свое мнение о шести объектах экспертизы одинаково: так, как это представлено в таблице 1.4. Определить весомость каждого объекта и построить ранжированный ряд.

**Решение:**

1) Частоты предпочтений

$$F_{i1} = 4/5=0,8; \quad F_{i2} = 3/5=0,6; \quad F_{i3} = 5/5=1,0;$$

$$F_{i4} = 0/5=0,0; \quad F_{i5} = 1/5=0,2; \quad F_{i6} = 2/5=0,4.$$

2) Общее число суждений каждого эксперта  $C = (6(6-1))/2=15$ .

3) Балл или весомость первого объекта экспертизы по общему мнению пяти экспертов:  $G_1 = 0,8/15+0,8/15+0,8/15+0,8/15+0,8/15 = 0,27$

Аналогично для остальных объектов  $G_2 = 0,20$ ;  $G_3 = 0,33$ ;  $G_4 = 0$ ;  $G_5 = 0,07$ ;  $G_6 = 0,13$ . Сумма  $G_1+G_2+G_3+G_4+G_5+G_6=0,27+0,20+0,33+0+0,07+0,13=1$

4) Ранжированный ряд объектов экспертизы имеет вид: № 3; № 1; №2; №6; №5; №4.

**Вывод:** лучшим оказался объект экспертизы под №3, худшим - объект №4.

Опыт попарного сопоставления показывает, что в силу особенностей человеческой психики эксперты иногда бессознательно отдают предпочтение не тому объекту в очередной рассматриваемой паре, который важнее, а тому, который стоит в перечне первым. Чтобы избежать этого, попарное сопоставление проводят дважды: 1) в прямом (сначала первого объекта со вторым, третьим, четвертым и т.д., затем второго с первым, третьим, четвертым, ... и так до последнего) и 2) в обратном (последнего с предпоследним, и до первого; предпоследнего с последним, предыдущим ... и вновь до первого) порядке. При сопоставлении в обратном порядке используют свободную часть таблицы (см. табл.1.4). Таким образом, каждая пара объектов сопоставляется дважды, в разном порядке и по истечении некоторого времени. Такое сопоставление, называют полным или двойным,

Оно позволяет выявить экспертов, небрежно относящихся к своим обязанностям или не имеющих определенной точки зрения и не допустить случайных ошибок. Т.е, двойное попарное сопоставление обладает более высокой надежностью, чем однократное. Порядок расчетов при нем остается прежним, за исключением того, что  $C = t(t-1)$ .

#### 1.2.4. Метод последовательного сопоставления

**Последовательное сопоставление** каждого объекта экспертизы с совокупностью всех тех, которые ниже рангом, позволяет откорректировать ранжированный ряд, уточнить позиции входящих в него объектов с учетом их важности. Оно имеет смысл тогда, когда несколько объектов экспертизы можно рассматривать как один составной объект той же природы. Порядок последовательного сопоставления следующий:

1. Объекты экспертизы располагаются в порядке их предпочтения (ранжируются):

2. Наиболее важному объекту приписывается балл или весовой коэффициент, равный 1; всем остальным в порядке уменьшения их относительной значимости — баллы или весовые коэффициенты от 1 до 0;

3. Первый объект сопоставляется с совокупностью всех остальных. Если, по мнению эксперта, он предпочтительнее, чем совокупность всех остальных вместе взятых, то результат его измерения корректируется в сторону увеличения с таким расчетом, чтобы он стал больше суммы баллов (весовых коэффициентов) всех остальных объектов экспертизы, которые ниже рангом. В противном случае результат измерения первого объекта корректируется в сторону уменьшения так, чтобы он оказался меньше суммы баллов или весовых коэффициентов остальных объектов;

4. Аналогично второй объект сопоставляется с совокупностью всех остальных, стоящих ниже рангом. Корректируется результат его измерения (при этом нужно следить, чтобы не нарушилось предпочтение первого объекта перед совокупностью всех остальных, если оно установлено на предыдущем этапе). Такая процедура сопоставлений и корректировок продолжается вплоть до предпоследнего объекта;

5. Полученные результаты измерений или весовые коэффициенты нормируют, т.е. делят на общую сумму баллов или весовых коэффициентов. После этого они принимают значения в пределах от 0 до 1, а их сумма становится равной 1.

#### 1.2.5. Социологический метод

Существует так называемый социологический метод оценки качества продукции. Этот метод, как и экспертный, основан на опросах, на мнениях, но не специальных экспертов, а различных потребителей оцениваемой продукции (фактических или потенциальных). Сбор мнений потребителей производится опросом или с помощью распространения и заполнения специальных анкет-вопросников, а также путем организации конференций,

выставок, аукционов, опытно-показательной эксплуатации и т.п. Социологический метод часто используется для решения маркетинговых задач.

## Задачи

**Задача 1.1.** Сформировать список предполагаемых потребностей пользователя по показателям

- 1) назначения (функциональные возможности);
- 2) удобства применения

для программы: пригодной а) для создания текстовых документов, б) для создания рисунков в) для создания вычисляемых таблиц.

**Задача 1.2.** Сформировать 5-бальную шкалу для показателей практичности ПП и оценить по ней программы, перечисленные в таблице 1.5

*Таблица 1.5 - Список прикладных программ:*

Вариант	1	2	3	4	5	6	7
Про- грамма	Блок- нот	Word Pad	Rpad 32	OpenOffice. org Writer	Word	OpenOffice. org Draw	PAINT

**Задача 1.3.** Для программ, перечисленных в таблице 1.5 построить ранжированный ряд по показателям назначения и удобства применения их для создания а) текстовых документов, б) рисунков в) вычисляемых таблиц

**Задача 1.4.** Используя метод попарного сопоставления расположить текстовые редакторы (табл.1.5) в порядке возрастания их а) функциональных возможностей, б) удобства применения

**Задача 1.5.** По результатам решения задачи 1.4 несколькими студентами (экспертами) построить обобщенный ранжированный ряд и определить весомость членов ряда.

**Задача 1.6.** По результатам решения задачи 1.4 несколькими студентами (экспертами) определить степень согласованности мнения экспертов.

**Задача 1.7.** Определить степень согласованности мнения экспертов по результатам ранжирования объектов экспертизы, приведенным в примере 1.2.

## Контрольные вопросы

1. Какими стандартами регламентируется оценка качества программных продуктов?
2. Что понимают под качеством программного обеспечения?
3. С помощью каких примитивов описывают надежность?
4. Какие виды измерительных шкал используют при оценке качества программных продуктов?
5. Что понимают под термином «надежность эксперта»?
6. Что показывает коэффициент конкордации Кендалла?
7. В чем заключается метод попарного сопоставления?