

## 6. THE FORMATION OF AN EXTERNAL CRITERION FOR MACHINE LEARNING BASED ON MEDICAL-BIOLOGICAL DATA

*Nella Alexeevna Shchukina, researcher, Military Medical Academy, Russia, St.Petersburg, Akademika Lebedeva str., house 6, BOX 194044, tshukina.nella@gmail.com.*

**Annotation.** *In medical-biological research, a serious obstacle to the use of machine learning methods is the ambiguity of external criteria as the basis of a “well-understood” image in the pattern recognition tasks. In the work, the absence of a clear external criterion is compensated by the creation of a meaningful scale that establishes the correspondence between a quantitative trait and its qualitative assessments. To select respondents for a specialty by the level of their physical fitness, an external criterion based on the qualitative assessment of their anaerobic abilities is modeled. The article describes the stages of empirical modeling of the external criterion by a method of mediate scaling the physiological properties of respondents.*

**Keywords.** *Percentile, sten, anaerobic capacity, attribute, scaling, categorization, classification.*

## ФОРМИРОВАНИЕ ВНЕШНЕГО КРИТЕРИЯ ДЛЯ МАШИННОГО ОБУЧЕНИЯ НА ОСНОВЕ МЕДИКО-БИОЛОГИЧЕСКИХ ДАННЫХ

*Нэлла Алексеевна Щукина, научный сотрудник, Военно-медицинская академия имени С.М.Кирова, Россия, Санкт-Петербург, ул. Академика Лебедева, д. 6, индекс 194044, tshukina.nella@gmail.com.*

**Аннотация.** *В медико-биологических исследованиях серьёзным препятствием для применения методов машинного обучения является неопределённость внешнего критерия как основы «хорошо понимаемого образа» в задачах распознавания. В работе отсутствие чёткого внешнего критерия компенсируется созданием содержательной шкалы, устанавливающей соответствие между количественным признаком и его качественными оценками. Для отбора респондентов на специальность по уровню их физической подготовленности моделируется внешний критерий, основанный на качественной*

оценке их анаэробных способностей. В статье описаны этапы эмпирического моделирования внешнего критерия методом опосредованного шкалирования физиологического свойства респондентов.

**Ключевые слова.** Процентиль, стен, анаэробная способность, признак, шкалирование, категоризация, классификация.

## Введение

Медико-биологические исследования — та область, где исследователь ведёт поиск закономерностей, опираясь на принципы теории планирования эксперимента, на существующие методы работы с малыми выборками, использует возможности *bootstrap*<sup>6.1</sup> и другие методы.

При этом исследователь решает задачи по сопоставлению экспериментальных данных, выявлению связей, оцениванию процессов и явлений, прогнозированию результатов, распознаванию образов и др. То есть в целом осуществляет содержательный анализ данных, в ходе которого вырабатываются решающие правила по отношению к совокупности изучаемых объектов (субъектов) методами машинного обучения.

Медико-биологические показатели человека составляют основной предмет исследования в физиологии военного труда, так как успех выполнения работы, которая, как правило, характеризуется большими физическими нагрузками и повышенным психоэмоциональным напряжением, зависит, прежде всего, от состояния здоровья и уровня работоспособности военнослужащих. И профессиональному отбору придаётся большое значение.

---

<sup>6.1</sup>*bootstrap*, *bootstrapping* (встречающиеся варианты перевода: бутстрап, бутстреп) — практический компьютерный метод исследования статистик вероятностных распределений, основанный на многократной генерации выборок методом Монте-Карло на базе имеющейся выборки [6.11]. Предложен в 1977 году Брэдли Эфроном (первая публикация относится к 1979 году). Суть метода состоит в том, чтобы по имеющейся выборке построить эмпирическое распределение. Используя это распределение как теоретическое распределение вероятностей можно с помощью датчика псевдослучайных чисел генерировать практически неограниченное количество псевдовыборок произвольного размера, например, такого же, как у исходной выборки. На множестве псевдовыборок можно не только оценить статистические характеристики, но и изучить их вероятностные распределения.

В задачах отбора на специальность, наряду с профессионально важными психологическими качествами, экспертной оценке подлежат определённые физиологические и психофизиологические свойства (способности) кандидатов. При этом для некоторых специальностей в качестве критерия пригодности может рассматриваться именно физиологическое свойство, которое:

- либо является измеряемым и нормируемым;
- либо оценивается опосредованно (через другие измеряемые и нормируемые показатели).

Для измеряемых и нормируемых физиологических и психофизиологических свойств людей существуют медицинские (физиологические) нормы, которые, часто носят достаточно условный характер. Эти нормы не всегда обеспечивают ранжирование кандидатов по степени пригодности к требованиям специальности (кто пригоден в первую очередь, кто – во вторую, кто пригоден условно, кто – непригоден), поэтому не могут полноценно выполнять функцию внешнего критерия.

В случае недостаточно формализованных измеряемых свойств респондентов рекомендуется разработать подходящую альтернативу – целевой показатель, выполняющий функцию внешнего критерия. В данной работе для измеряемого в интервальной шкале физиологического свойства  $X$  предлагается разработать категориальную форму  $Y$ , содержательно отражающую степень (уровень) развития соответствующего свойства у респондентов. При этом под целевым показателем понимается  $Y$  как совокупность категорий, определяющих пригодность респондентов к той или иной деятельности и используемых в дальнейшем как внешний критерий.

То есть отсутствие чёткого внешнего критерия для распознавания образов при машинном обучении компенсируется предварительным созданием содержательной шкалы, устанавливающей соответствие между количественными медико-биологическими данными и их качественными оценками (уровнями, категориями).

Совокупность качественных оценок может быть задана заранее (по нормативам).

В том и другом случае должно быть сформулировано решающее правило  $X \rightarrow Y$  перехода от измеренных значений физиологического свойства к значениям целевого показателя (уровням, категориям).

Оба варианта относятся к области управляемого машинного обучения.

В более сложных случаях, когда значения измеряемого свойства по каким-либо объективным причинам получить не представляется возможным, предварительно необходимо подобрать релевантную физиологическую замену (аналог) и обосновать возможность замены. После чего на подобранном аналоге осуществить шкалирование и сформировать искомый целевой показатель.

В данной главе мы будем заниматься именно этим случаем.

Идея состояла в разработке способа эмпирического моделирования целевого показателя  $Y$  как внешнего критерия, необходимого для применения управляемого машинного обучения в задачах профессионального отбора на специальность, предъявляющую требования к трудно измеряемым физиологическим способностям респондентов. В работе предложен алгоритм доказательной замены одного физиологического показателя пригодности респондентов к деятельности другим физиологическим показателем с целью последующего построения решающих правил  $X \rightarrow Y$ .

Реализация идеи осуществлена на примере замены скоростного критерия физической подготовленности (внешний критерий) респондентов тестовым индексом максимальной анаэробной мощности организма (физиологический аналог, принимаемый за  $X$ ). Предложен способ шкалирования физиологического аналога  $X$  и описано соответствующее правило отображения  $X \rightarrow Y$ . Разработан вариант оценки пригодности респондентов к специальности, предъявляющей требования к их физической подготовленности.

Глава структурирована следующим образом.

В первом разделе приведено описание постановки задачи по отбору кандидатов на специальность, предъявляющую требования к уровню их физической подготовленности.

Во втором разделе представлен общий план решения поставленной задачи.

В третьем разделе подобран метод шкалирования измеряемых физиологических показателей.

В четвёртом разделе дано описание эксперимента по разработке решающего правила отбора кандидатов на специальность, включающее:

- доказательное обоснование альтернативной замены внешнего критерия физической подготовленности респондентов соответствующим физиологическим показателем;
- разработку целевого показателя методом непрямого шкалирования;
- разработку критерия оценки уровня тестируемого физиологического свойства;
- формулирование решающего правила оценки пригодности респондентов к специальности.

В пятом разделе обобщается методика шкалирования измеряемых физиологических свойств (способностей) респондентов.

### **6.1. Постановка задачи по отбору кандидатов на специальность**

Обратимся к физиологии военного труда. *Задача.* Предстоит осуществлять отбор кандидатов на специальность, требующую выполнения относительно непродолжительной, но очень интенсивной работы (работы большой мощности в минимальное время). Необходимо разработать руководство по отбору с учётом того, что условия проведения отбора могут быть затруднены (сложная горная местность, нахождение на плавсредстве и т. п.).

Очевидно, что при отборе будут предъявляться повышенные требования к уровню физической подготовленности кандидатов. Тестирование уровня физической подготовленности военнослужащих производится согласно действующему в настоящее время наставлению НФП-2009 по физической подготовке (далее — НФП-2009).

- В соответствии с НФП-2009 оцениваются такие физические качества как:
- выносливость (способность продолжительное время выполнять работу на высоком уровне без снижения её эффективности);
  - сила (способность преодолевать внешнее сопротивление или противодействовать ему за счёт мышечных усилий);

- быстрота (способность совершать двигательные действия в минимальное время);
- ловкость (способность целесообразно согласовывать, соподчинять движения и организовывать их в единое целое).

Среди перечисленных качеств в рамках поставленной задачи первоочередное значение отводится такому качеству как быстрота. В соответствии с НФП-2009 критерием оценки данного качества принимается время пробега 100-метровой дистанции  $T$ , которое измеряется в секундах.

При этом бег на короткую дистанцию с максимальной скоростью (спринтерский бег) с полным основанием можно отнести к относительно непродолжительной, но очень интенсивной работе — работе большой мощности, выполняемой в короткое время. То есть свойственное кандидату значение показателя быстроты  $T$ , по сути, является *прямым показателем* его способности выполнить такую работу в течение короткого времени <sup>6.2</sup>.

Однако для получения указанного показателя придётся всякий раз при отборе организовывать массовые мероприятия по бегу на короткие дистанции, что не всегда выполнимо (см. условие задачи).

Необходимо разработать способ отбора кандидатов по уровню требуемого качества методически осуществимый в стационарных условиях.

*Но как иначе измерять и оценивать этот уровень? По какому правилу осуществлять обследование и дифференциацию кандидатов? Наконец, что вообще можно рассматривать в качестве аналога указанного показателя?*

В физиологии установлено, что некоторые качества человека зависимы от врождённых задатков. Причём наибольшая наследственная обусловленность выявлена для морфологических показателей организма человека, меньшая — для физиологических параметров и наименьшая — для психологических признаков [6.13—6.15].

В частности, под выраженным генетическим контролем находится такой морфофункциональный признак как *анаэробные возможности* [6.14] <sup>6.3</sup>.

<sup>6.2</sup>Для простоты оставляем за скобками дискуссию о том, есть ли необходимость учёта восстановительных способностей и резервных возможностей организма кандидатов в рамках настоящей работы.

<sup>6.3</sup>Анаэробная возможность — это способность выполнять кратковременную работу большой мощности (анаэробную нагрузку) в течение короткого времени.

В то же время генетическому контролю в наибольшей степени подчинены быстрые движения. Об этом свидетельствуют высокие коэффициенты наследуемости  $H$ , полученные исследователями для различных элементарных проявлений качества быстроты (таблица 6.1) [6.13—6.15].

Таблица 6.1

**Показатели влияния наследственности ( $H$ ) на физические качества человека**

Физическое качество человека	Коэффициент наследуемости ( $H$ )
Скорость двигательной реакции	0,80
Теппинг-тест	0,85
Скорость элементарных движений	0,64
Скорость спринтерского бега	0,70
Максимальная статическая сила	0,55
Взрывная сила	0,68
Координация движений рук	0,45
Суставная подвижность (гибкость)	0,75
Локальная мышечная выносливость	0,50
Общая выносливость <sup>10</sup>	0,65

Так, в [6.14; 6.15] говорится, что с помощью близнецового и генеалогического методов подтверждена высокая зависимость от врождённых свойств ( $H = 0,70 - 0,90$ ) показателей скоростного бега на короткие дистанции, теппинг-теста, кратковременного педалирования на велоэргометре в максимальном темпе, прыжков в длину с места и других скоростных и скоростно-силовых упражнений.

Другими словами, быстрота и гибкость — это наименее тренируемые физические качества, а значит наиболее стабильные (мало изменяемые) показатели физической подготовленности человека.

В основе качества быстроты лежит анаэробное обеспечение энергией мышечной деятельности<sup>6.4</sup>.

<sup>6.4</sup>Индивидуальные различия в проявлении быстроты связываются и с особенностями нервной системы, которые также в основном генетически обусловлены. Но это выходит за рамки данной работы.

Анаэробный механизм обеспечения мышечной деятельности также испытывает значительное влияние генетических факторов. Коэффициент наследуемости этого механизма, согласно данным большинства исследователей, составляет от 70% до 80%. Более того, многие авторы указывают на то, что наследуемость анаэробной работоспособности может составлять до 90% и выше [6.13—6.15].

Анаэробная работоспособность определяет выносливость, проявляемую человеком в процессе непродолжительной, но очень интенсивной работы. Причём практически каждый человек использует лишь часть своих анаэробных возможностей, уровень которых косвенно можно оценивать в стационарных условиях. Для определения аэробных возможностей организма в лабораторных условиях используют моделирование реальной мышечной деятельности — нагрузочные тесты с применением различных средств (трекбаны, велоэргометры и др.). Тестовые оценки зависят от способа тестирования, а также от принятых единиц измерения. Поэтому в разных источниках (методиках) они различны. В простейшем случае для оценки максимальной анаэробной мощности (далее — *МАМ*) организма требуются: секундомер, несколько ступеней лестницы и рулетка (показатель *МАМ*, кгм/с).

Итак, анаэробные возможности человека обусловлены генетически. Они, в свою очередь, определяют его скоростные способности. Оба эти качества, будучи связанными, зависимы от врождённых задатков.

Формализованной (аналитической) формы столь непростой связи не существует.

Но если прямой показатель быстроты *T* и косвенный показатель анаэробных возможностей *МАМ* действительно находятся во взаимном соответствии, то для отбора кандидатов в стационарных условиях вместо времени пробега 100-метровой дистанции был бы вполне пригоден показатель физиологического теста *МАМ*<sup>6.5</sup>.

Однако на этот счёт в настоящее время отсутствуют какие-либо руководства и наставления, что усложняет задачу.

---

<sup>6.5</sup>Идея заменить спринтерский бег тестированием военнослужащих по физиологическому тесту *МАМ* принадлежит Авдюшенко С.А. — начальнику НИЛ (физиологии военного труда) научно-исследовательского центра ВМедА.



Для обоснования выдвинутого предположения и возможности его практического применения предстоит:

- 1) доказать, что в рамках поставленной задачи время пробега 100-метровой дистанции можно заменить показателем методики *МAM*;
- 2) разработать критерий оценки уровня анаэробных возможностей по значениям показателя *МAM*;
- 3) сформировать правило дифференциации кандидатов по уровню анаэробных возможностей.

## 6.2. Этапы разработки способа отбора кандидатов на специальность

Без эксперимента выполнить вышеприведённые требования не представляется возможным. С учётом этого обстоятельства процесс разработки способа отбора кандидатов на специальность должен выполняться в два этапа (рис. 6.1).

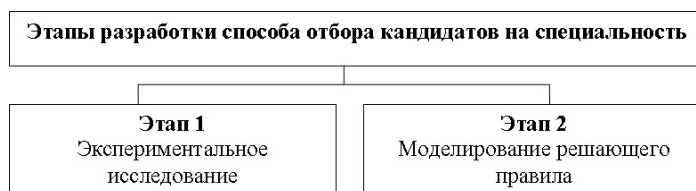


Рис. 6.1. Схема поэтапной разработки способа отбора кандидатов на специальность

На первом этапе проводится экспериментальное исследование:

- формируется эталонная выборка кандидатов с учётом возраста и пола ( $n \geq 200$  по требованию к объёму эталонных выборок);
- организуются мероприятия по получению типичного для каждого кандидата времени пробега 100-метровой дистанции  $T$ ;
- осуществляется индивидуальное тестирование кандидатов по физиологической методике *МAM*;
- по экспериментальным данным анализируются особенности корреляции между показателями  $T$  и *МAM*;
- при выявлении явного взаимного соответствия прямой показатель  $T$  способности кандидатов к выполнению работы большой мощности в

короткое время заменяется на показатель *МАМ* уровня анаэробных возможностей .

*На втором этапе* разрабатывается правило дифференциации кандидатов именно по уровню показателя *МАМ*.

*Принимается допущение* о том, что распределение показателя анаэробных возможностей у людей подчиняется нормальному закону. Допущение вполне естественное для способности выполнять кратковременную работу большой мощности (анаэробную нагрузку) в течение короткого времени. Кроме того, основная масса трудоспособных людей наделена уровнями анаэробных возможностей, которые условно можно рассматривать как интервал нормы, у менее значительной части людей эти возможности выражены слабее или, напротив, сильнее нормы<sup>6.6</sup>. Правда, истинные параметры этого нормального распределения неизвестны. Но их оценкой могут служить параметры эталонной (репрезентативной) выборки. Эта выборка принимается за выборку стандартизации. На ней разрабатывается искомый стандарт с помощью процедуры нелинейной нормализации.

В соответствии с принятым допущением найдутся некие пороги показателя *МАМ*, позволяющие дифференцировать его значения как высокие, средние или низкие. Пороги будут разделять диапазон возможных значений *МАМ* у кандидатов на несколько интервалов (рисунок 6.2).

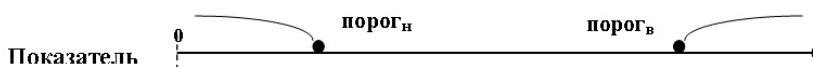


Рис. 6.2. Интервалы значений показателя *МАМ* у кандидатов

На этой основе можно разработать критерий оценки уровня анаэробных возможностей у кандидатов по схеме, представленной в таблице 6.2, и осуществить дифференциацию кандидатов на группы по индивидуальному уровню *МАМ*.

Но для правила отнесения кандидата к какой-либо группе, требуется прежде определить сами эти группы, т.е. сначала сформировать категории показателя *МАМ*. Для этого диапазон возможных значений параметра раз-

<sup>6.6</sup>Ровно на этом же основании закон распределения какого-либо психологического показателя генеральной совокупности принято считать нормальным.

**Критерий группирования кандидатов по уровню анаэробных возможностей**

Значение <i>МАМ</i>	Уровень <i>МАМ</i>	Категория индивидуальной анаэробной мощности
$МАМ < \text{порог}_H$	низкий	3
$\text{порог}_H \leq МАМ < \text{порог}_B$	средний	2
$МАМ \geq \text{порог}_B$	высокий	1

бить на подмножества, для которых ввести порядок предпочтения. При этом результаты разбиения не будут носить универсальный характер, так как для разных контингентов границы категорий могут не совпадать. То есть категории показателя не являются синонимом классов в строгом смысле, поскольку изначально отсутствует разбиение на блоки эквивалентности. Кроме того границы различных категорий являются нечёткими, размытыми, и обычно сама категория понимается не через её формальное определение, а лишь в сравнении с другими категориями.

Определение критических значений показателя *МАМ* ( $\text{порог}_H$  — нижнее (меньшее) критическое значение, а  $\text{порог}_B$  — верхнее (большее) критическое значение) требует помимо теоретических допущений привлечения ряда статистических положений.

В целом формирование категорий *МАМ* (*целевого показателя*) представляет отдельную, непростую задачу шкалирования и связано со статистическими вопросами, решаемыми на релевантной базе экспериментальных данных.

### 6.3. Выбор шкалы для разработки целевого показателя

*Шкалирование* (от лат. *scala* — «лестница») совмещает в себе черты количественного и качественного изучения исследуемого свойства. Количественный аспект шкалирования состоит в том, что в его процедуру входит числовое представление данных. Качественный аспект выражается в том, что шкалирование позволяет манипулировать не только числовыми данными, но и данными, не имеющими единиц измерения, и включает в себя элементы ка-

качественных методов таких как классификация, типологизация, систематизация [6.8].

Итогом шкалирования является построение шкалы, то есть некоторой знаковой (числовой) модели исследуемой реальности (в нашем случае степени соответствия измеряемого свойства кандидатов определённым требованиям отбора), посредством которой можно эту реальность измерить [6.8].

Сведения о шкалировании (как методе моделирования реальных процессов с помощью числовых систем) и его модификациях достаточно широко представлены в работах многих авторов. В работах [6.2—6.5; 6.7; 6.9; 6.10; 6.12; 6.17; 6.20; 6.21] рассматриваются основные процедуры шкалирования, такие как парное сравнение объектов, отнесение объектов к категориям и др., описываются характеристические особенности шкал. Шкалы измерений принято классифицировать по типам измеряемых данных, которые определяют допустимые для данной шкалы математические преобразования, а также типы отношений, отображаемых соответствующей шкалой. Современная классификация включает 4 типа измерительных шкал:

- номинативную (номинальную или шкала наименований);
- порядковую (ординальную);
- интервальную (шкалу равных интервалов);
- равных отношений.

Измерения с помощью первых двух измерительных шкал считаются качественными, это шкалы неметрические; измерения с помощью третьей и четвёртой измерительных шкал считаются количественными, эти шкалы — метрические.

В нашей постановке задачу шкалирования нужно рассматривать как процедуру перехода от интервальной (метрической) шкалы диапазона значений физиологического показателя к номинативной шкале категорий.

*В соответствии с принятым допущением распределение показателя анаэробных возможностей у людей подчиняется нормальному закону.*

Однако, помимо данного теоретического допущения конструирование шкалы категорий для показателя анаэробных возможностей требует ещё и ряда статистических манипуляций, то есть будет осуществляться опосредованным путём. Применяемый при этом метод непрямого шкалирования —

метод категорий предполагает исследование частотного распределения значений показателя. При нормальном (близком к нормальному) законе распределения исходных значений показателя имеется возможность выбора релевантной стандартной шкалы.

В мировой практике применяется ряд стандартных шкал, разработанных на базе нормального закона распределения (рисунок 6.3)<sup>6.7</sup>.

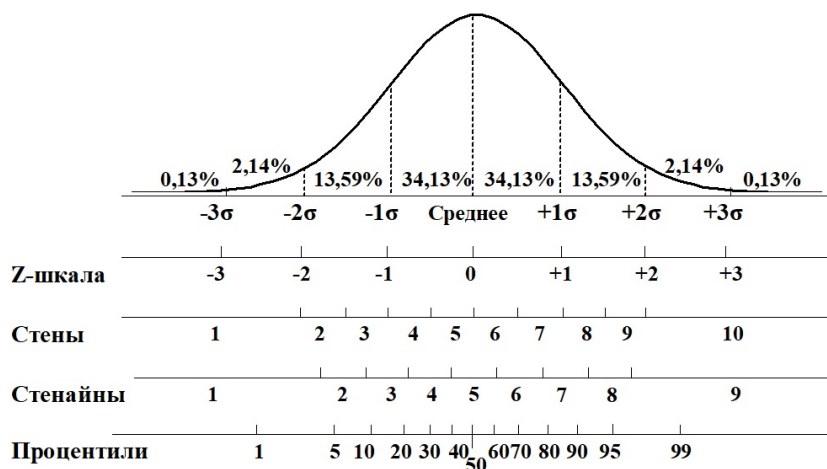


Рис. 6.3. Нормальная кривая и стандартные шкалы [6.6]

Для дальнейшей разработки нам необходимы две из них — шкала процентиля и шкала стен<sup>6.8</sup>.

Чтобы исключить разночтение терминов, следует уточнить несколько существенных моментов:

- а) процентиль — это определённое значение из упорядоченной нормально распределённой совокупности, которому на шкале процентиля соответствует одна и только одна точка, называемая процентильным уровнем (первая процентиль, вторая процентиль, и т. д. до сотой процентиля). Процентильные уровни могут быть не только целыми, например 2,28. В этом случае произносится «процентиль уровня 2,28»;
- б) процентильный уровень — это не значение самой процентиля, а лишь ее условный код (аналог порядкового номера) на непрерывной шкале.

<sup>6.7</sup> В статье мы неоднократно будем обращаться к этому рисунку.

<sup>6.8</sup> Пишется «стен», произносится «стэн».

Под одним и тем же процентильным уровнем в разных нормально распределенных совокупностях оказываются разные числовые значения (процентили);

- с) стен — это некоторый интервал значений в упорядоченной нормально распределённой совокупности. Ему на шкале стен соответствует числовой отрезок с определенным порядковым номером (первый стен, второй стен, и т. д. до десятого стена)<sup>6.9</sup>. Над порядковыми номерами стен бессмысленно выполнять какие-либо арифметические действия (в том числе определять среднее значение);
- д) расчёт процентилей и расчёт стен производится только на предварительно упорядоченной совокупности (*по возрастанию*)<sup>6.10</sup>.

Между шкалой процентилей и шкалой стен существует однозначное соответствие (это отражено на рисунке 6.3): стенам соответствуют проценти́ли с неизменными процентильными уровнями [6.18].

Поскольку каждая проценти́ль (как и ее процентильный уровень) — это всего лишь число, а каждый стен — это определённый числовой интервал со своим номером, то их соответствие нуждается в уточнении. С этой целью из процентильных уровней выделяют те, которые соответствуют границам стен (таблица 6.3). Данное обстоятельство обладает значительной практической ценностью, которая будет раскрыта позже.

Процентиль через ее процентильный уровень обычно «привязывают» к правой границе соответствующего стена [6.18], хотя исследователь имеет полное право переназначить привязку (рисунок 6.4)<sup>6.11</sup>.

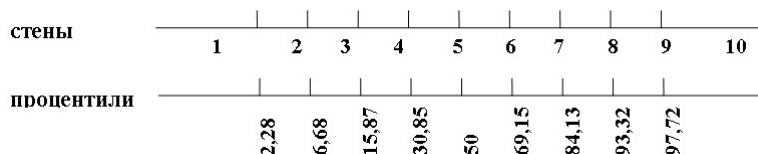


Рис. 6.4. Схема соответствия между стенами и процентильными уровнями

<sup>6.9</sup>Первому стелу и десятому стелу условно соответствует числовой луч.

<sup>6.10</sup>Расположение данных именно в порядке возрастания соотносится с геометрическим представлением эмпирического распределения и его описательных характеристик (статистик).

<sup>6.11</sup>На рисунке 6.3 отмечены только целые процентильные уровни. Рисунок 6.4 является дополнительным поясняющим фрагментом к рисунку 6.3.

Соответствие между стенами и процентильными уровнями

Номер стены (i)	Процентильный уровень ( $P_i$ )
1	2,28
2	6,68
3	15,87
4	30,85
5	50
6	69,15
7	84,13
8	93,32
9	97,72
10	

Значения наиболее показательных процентилей (таблица 6.4) и границы стандартных числовых интервалов (от первого стена до десятого) вычисляются по определенным соотношениям для любой нормально распределенной генеральной совокупности на основании ее основных характеристик положения и рассеяния — математического ожидания  $\mu$  и среднеквадратического отклонения  $\sigma$  [6.16].

Таблица 6.4

Процентильный уровень	Расчётное соотношение
2,5	$\mu - 2\sigma$
16	$\mu - \sigma$
50	$\mu$
84	$\mu + \sigma$
97,5	$\mu + 2\sigma$

В силу того, что истинные значения данных характеристик невозможно указать ни для какой конкретной генеральной совокупности, практически применяются их статистические оценки: среднее значение  $\bar{X}$  и стандартное отклонение  $S$  репрезентативной выборки, а также используются правила, основанные на описательных статистиках репрезентативной выборки.

Правило расчёта процентилей следующее.

1. Репрезентативная выборка значений исследуемого показателя упорядочивается по возрастанию<sup>6.12</sup>.
2. Рассчитывается порядковый номер процентиля в упорядоченной выборке по формуле:

$$j = \frac{n}{100} \times P, \quad (6.1)$$

где  $j$  — номер искомой процентиля в упорядоченной выборке<sup>6.13</sup>;  $P$  — заданный процентильный уровень<sup>6.14</sup>;  $n$  — объем выборки.

3. По номеру  $j$  в упорядоченной выборке находится «сырое» значение, являющееся искомой процентилью.

Таким образом, умея рассчитывать порядковые номера процентилей по формуле (6.1) и определять их фактические значения, можно для исследуемого показателя указать правые границы стандартных интервалов (стен) до девятого стена включительно (все, что правее — это уже десятый стен). Затем выделить ту часть, которую предполагается считать типичной (нормой) и ту, которая по мнению исследователя таковой не является.

Почему в данном случае так важны процентиля?

Во-первых, даже хорошо подготовленные (выверенные) экспериментальные данные практически не подчиняются нормальному (Гауссовскому) закону распределения, а всего лишь приближаются к нему<sup>6.15</sup>.

*Но для экспериментальной репрезентативной выборки существует гипотетическое нормальное распределение как теоретический образ, к которому она близка и на которое может отобразиться*<sup>6.16</sup>.

Во-вторых, отображение экспериментальной выборки на гипотетическое нормальное распределение как раз и производится через поиск в выборке тех

<sup>6.12</sup> Каждое «сырое» значение в выборке является процентилью какого-то процентильного уровня.

<sup>6.13</sup> Рассчитанное значение  $j$  округляется до целых по обычному правилу округления.  $j$  — это порядковый номер места того «сырого» значения в упорядоченной выборке, которое является процентилью уровня.

<sup>6.14</sup> Выборочные процентильные уровни от нуля до 100 (при необходимости и дробные значения, например, 2,28 и т. п.) задаёт исследователь, исходя из потребностей решаемой задачи.

<sup>6.15</sup> Для нормального распределения выполняются как минимум следующие условия: медиана и мода совпадают со средним значением; асимметрия и эксцесс равны нулю, а всё рассеяние значений укладывается в известных пределах по «правилу трёх сигма».

<sup>6.16</sup> Нормально распределенная выборка, если таковая все-таки найдётся, отображается сама на себя.



«сырых» процентильных значений, которые будут соответствовать стандартным интервалам (стенам). Эту процедуру называют нелинейной нормализацией экспериментальных данных.

Таким образом, постоянство связи между процентилями и стенами (см. таблицу 6.3) обеспечивает возможность выхода на гипотетические тестовые нормы с помощью представленной выше процедуры нелинейной нормализации. *Это важно в экспериментальном исследовании.*

#### **6.4. Описание эксперимента по разработке решающего правила отбора кандидатов на специальность**

##### ***6.4.1. Обоснование альтернативной замены показателя скорости***

Утверждение о том, что альтернативой показателю скорости  $T$  может служить показатель  $MAM$  при оценке способности кандидатов выполнять работу большой мощности в короткое время, нуждается в обосновании. *Обоснование должно носить доказательный характер.*

С этой целью планируется и проводится эксперимент на эталонной выборке респондентов. В рамках настоящей публикации мы вынуждены пойти на упрощение данных: сформировали группу из 50 кандидатов, отобранных по возрасту и полу. При этом сознательно нарушили основное требование репрезентативности, касающееся объёма выборки ( $n \geq 200$ ). То есть нашу выборку нельзя принять как эталонную.

Ограничение объёма экспериментальной выборки до приемлемого размера, конечно же, отразится на точности искомых пороговых значений, и мы не вправе будем считать их достоверными<sup>6.17</sup>.

Но сейчас мы преследуем другую цель – показать, как логически построить доказательное обоснование, продемонстрировать необходимый минимум расчётов.

---

<sup>6.17</sup> Рассчитанные далее пороговые значения можно считать «пилотными», требующими уточнения посредством повторения всех операций на эталонной выборке стандартизации.

Итак, сформировав группу кандидатов, мы спланировали и провели эксперимент<sup>6.18</sup>:

- организовали и провели мероприятия по получению типичного индивидуального времени пробега 100-метровой дистанции каждым кандидатом;
- в стационарных условиях осуществили тестирование кандидатов по физиологической методике *МАМ*

В таблице 6.5 представлены значения показателей *T* и *МАМ*, полученные в ходе эксперимента<sup>6.19</sup>. Три последние строки таблицы представляют среднее значение, медиану и моду соответственно.

Таблица 6.5

Экспериментальные значения исследуемых показателей ( $n = 50$ )

Кандидат, ФИО	<i>T</i> , с	<i>МАМ</i> , кгм/с
ФИО-01	14,0	80,0
ФИО-02	13,7	84,0
ФИО-03	12,9	111,0
ФИО-04	13,0	88,0
ФИО-05	12,3	100,0
ФИО-06	13,6	88,0
ФИО-07	14,1	80,5
ФИО-08	13,1	97,0
ФИО-09	13,4	110,0
ФИО-10	13,2	99,7
ФИО-11	13,7	91,0
ФИО-12	14,2	78,0
ФИО-13	12,3	103,0
ФИО-14	13,0	100,0
ФИО-15	13,4	96,0
ФИО-16	12,3	104,0
ФИО-17	14,2	89,0

<sup>6.18</sup>Подробности об условиях проведения и методическом обеспечении эксперимента мы опускаем ввиду их неактуальности для нашего обучающего примера.

<sup>6.19</sup>Из данных НИЛ (физиологии военного труда). ФИО-1, ФИО-2 и т. д. — это условные имена кандидатов. Не будем путать их с порядковыми номерами элементов выборки.

Продолжение таблицы 6.5

Кандидат, ФИО	$T$ , с	$MAM$ , кгм/с
ФИО-18	12,6	125,0
ФИО-19	13,9	104,0
ФИО-20	13,3	98,6
ФИО-21	12,5	114,6
ФИО-22	12,6	97,0
ФИО-23	13,1	99,8
ФИО-24	12,6	137,0
ФИО-25	13,4	90,0
ФИО-26	14,0	75,0
ФИО-27	13,6	92,6
ФИО-28	12,9	115,0
ФИО-29	12,4	113,0
ФИО-30	12,6	107,0
ФИО-31	13,2	86,0
ФИО-32	13,2	96,2
ФИО-33	13,4	93,0
ФИО-34	14,8	79,0
ФИО-35	13,8	94,5
ФИО-36	13,5	96,0
ФИО-37	14,0	79,0
ФИО-38	12,8	110,0
ФИО-39	14,1	81,0
ФИО-40	13,6	92,0
ФИО-41	14,4	64,0
ФИО-42	13,0	100,0
ФИО-43	12,0	142,0
ФИО-44	12,5	120,5
ФИО-45	13,2	98,3
ФИО-46	12,0	140,0
ФИО-47	13,2	80,0
ФИО-48	14,1	70,0
ФИО-49	13,2	78,0

Кандидат, ФИО	$T$ , с	$MAM$ , кгм/с
ФИО-50	13,0	104,0
$\bar{X}$	13,3	97,4
$Me$	13,2	96,6
$Mo$	13,2	100,0

Согласно таблице 6.5 для экспериментальных значений каждого исследуемого показателя  $T$  и  $MAM$  было характерно незначительное различие между средним  $\bar{X}$ , медианой  $Me$  и модой  $Mo$ . Это свидетельствовало о слабо выраженной «скошенности» и несущественной «остро-плосковершинности» соответствующих кривых распределения<sup>6.20</sup>, то есть о близости к нормальному закону распределения. В рамках данной работы нет необходимости в более детальном анализе и описании вида распределения экспериментальных данных.

Оценим корреляцию между показателями  $T$  и  $MAM$ .

По результатам произведённых расчётов коэффициент линейной корреляции  $r$ -Пирсона ( $r = -0,796$ ) является статистически значимым ( $p = 4,8 \cdot 10^{-12}$ ).

Коэффициент ранговой корреляции  $\rho$ -Спирмена ( $\rho = -0,814$ ) также статистически значим ( $p = 1,0 \cdot 10^{-6}$ ).

Таким образом между показателями  $T$  и  $MAM$  с высокой вероятностью существует обратная ( $r < 0$ ;  $\rho < 0$ ) и достаточно тесная ( $|r| > 0,7$ ;  $|\rho| > 0,7$ ) связь, которую можно считать линейной ( $r \approx \rho$ )<sup>6.21</sup>.

Ограничившись этим анализом, мы поступили бы недальновидно, поскольку планируем разработать правило дифференциации кандидатов по уровню  $MAM$  более чем на две группы и заинтересованы в том, чтобы именно лучшим бегунам соответствовали лучшие значения  $MAM$ , худшим — худшие, а средним — средние. И наоборот.

<sup>6.20</sup>Для показателей  $T$  и  $MAM$  коэффициент асимметрии наблюдаемых значений составлял 0,1 и 0,7 соответственно, коэффициент эксцесса составлял -0,6 и 0,7 соответственно, что укладывалось в рамки допустимого отклонения от нуля для распределения, близкого к нормальному.

<sup>6.21</sup>Подробнее на интерпретации расчётных значений « $p$ » остановимся ниже при выполнении сравнительного статистического анализа.

Для обоснования соответствия между категориями показателей  $T$  и  $MAM$  был разработан перекрёстный сравнительный анализ (рисунок 6.5).

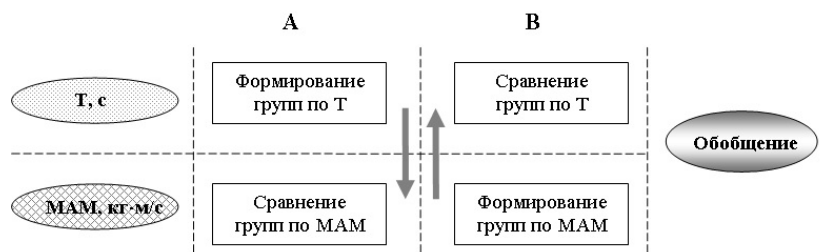


Рис. 6.5. Условная схема перекрёстного сравнительного анализа

Согласно представленной на рисунке 6.5 схеме поступим следующим образом.

Сначала выясним, как проявляет себя уровень анаэробных возможностей  $MAM$  в группах кандидатов с различными достижениями  $T$  по спринтерскому бегу на 100-метровую дистанцию (*шаг А*).

Затем выясним, как проявляет себя показатель  $T$  бега на 100-метровую дистанцию у групп кандидатов с различным уровнем  $MAM$  анаэробных возможностей (*шаг В*).

При обнаружении явного взаимного соответствия между категориями показателей  $T$  и  $MAM$  у кандидатов мы сможем приступить к разработке решающего правила дифференциации кандидатов по показателю  $MAM$ , приняв его за приемлемую замену прямого показателя  $T$ .

Шаг «А» и шаг «В» предусматривают предварительную стратификацию выборки стандартизации по определенным характеристикам кандидатов.

#### 6.4.1.1. Шаг А

На этапе А требуется произвести стратификацию по уровню достижений кандидатов в спринтерском беге  $T$ .

Прежде всего упорядочим выборку экспериментальных данных (таблица 6.5) по возрастанию показателя  $T$  и пронумеруем полученный ряд наблюдений.

Упорядоченный вариант данных представлен в таблице 6.6.

Таблица 6.6

Экспериментальные данные, упорядоченные по возрастанию показателя  $T$  ( $n = 50$ )

Присвоенный номер	Кандидат, ФИО	$T$ , с	$MAM$ , кг.м/с
1	ФИО-43	12,0	142,0
2	ФИО-46	12,0	140,0
3	ФИО-05	12,3	100,0
4	ФИО-13	12,3	103,0
5	ФИО-16	12,3	104,0
6	ФИО-29	12,4	113,0
7	ФИО-21	12,5	114,6
8	ФИО-44	12,5	120,5
9	ФИО-18	12,6	125,0
10	ФИО-22	12,6	97,0
11	ФИО-24	12,6	137,0
12	ФИО-30	12,6	107,0
13	ФИО-38	12,8	110,0
14	ФИО-03	12,9	111,0
15	ФИО-28	12,9	115,0
16	ФИО-04	13,0	88,0
17	ФИО-14	13,0	100,0
18	ФИО-42	13,0	100,0
19	ФИО-50	13,0	104,0
20	ФИО-08	13,1	97,0
21	ФИО-23	13,1	99,8
22	ФИО-47	13,2	80,0
23	ФИО-49	13,2	78,0
24	ФИО-10	13,2	99,7
25	ФИО-31	13,2	86,0
26	ФИО-32	13,2	96,2
27	ФИО-45	13,2	98,3
28	ФИО-20	13,3	98,6
29	ФИО-25	13,4	90,0
30	ФИО-09	13,4	110,0
31	ФИО-15	13,4	96,0

Продолжение таблицы 6.6

Присвоенный номер	Кандидат, ФИО	$T$ , с	$MAM$ , кг.м/с
32	ФИО-33	13,4	93,0
33	ФИО-36	13,5	96,0
34	ФИО-06	13,6	88,0
35	ФИО-27	13,6	92,6
36	ФИО-40	13,6	92,0
37	ФИО-02	13,7	84,0
38	ФИО-11	13,7	91,0
39	ФИО-35	13,8	94,5
40	ФИО-19	13,9	104,0
41	ФИО-01	14,0	80,0
42	ФИО-26	14,0	75,0
43	ФИО-37	14,0	79,0
44	ФИО-07	14,1	80,5
45	ФИО-39	14,1	81,0
46	ФИО-48	14,1	70,0
47	ФИО-12	14,2	78,0
48	ФИО-17	14,2	89,0
49	ФИО-41	14,4	64,0
50	ФИО-34	14,8	79,0

Выделим крайние группы кандидатов: с лучшими (наименьшими) и худшими (наибольшими) значениями показателя  $T$ , применив процентильный подход <sup>6.22</sup>, предполагающий наличие гипотетического нормального распределения, на которое может отобразиться выборка стандартизации.

Дальнейшие рассуждения будем связывать именно с этим распределением, которое мы подразумеваем и в которое как бы встраиваем нашу экспериментальную выборку.

Руководствуясь таблицей 6.3, примем в качестве целевых значений (порогов) две процентиля — процентиль уровня 15,87 и процентиль уровня 84,13

<sup>6.22</sup>Для эталонной выборки способы выделения групп по успешности бега на 100-метровую дистанцию с помощью процентильного подхода или по фактическим оценкам, рассчитанным с помощью соответствующего наставления по физической подготовке НФП-2009, дадут сопоставимые результаты.

— и найдём их значения, они «отсекут хвосты», между которыми останется средняя часть (см. рисунок 6.3).

Для этого определим номера искомым процентилей в упорядоченной выборке, по которым и отыщем их значения в таблице 6.6.

В соответствии с формулой (6.1) имеем:

- 1)  $j = \frac{50}{100} \times 15,8 \approx 8$ . То есть перцентиль уровня  $P = 15,87$  находится в таблице 6.6 под порядковым номером восемь и равна 12,5 с.;
- 2)  $j = \frac{50}{100} \times 84,13 \approx 42$ . То есть перцентиль уровня  $P = 84,13$  находится в таблице 6.6 под порядковым номером 42 и равна 14,0 с.

Таким образом, в нашем исследовании группу «лучших достижений» по бегу на 100-метровую дистанцию представляли кандидаты с табличными номерами от первого до восьмого включительно (гр. 1), группу «худших достижений» представляли кандидаты с табличными номерами от 42-го до 50-го включительно (гр. 3). В общей сложности 17 кандидатов (см. таблицу 6.6).

Средняя группа на первый взгляд должна включать 33 кандидата ( $50 - 17 = 33$ ). Но это не совсем так.

Обратимся к рисунку 6.3. На нем схематично изображена кривая плотности гипотетического распределения. При этом плотность «хвостов» явно несопоставима с плотностью центральной части («нормой»). И мы понимаем, что между этими участками слева и справа невозможно провести единственным образом чёткое разграничение, поскольку наблюдаются зоны неопределённости, которые однозначно не назовёшь «нормой», либо «не нормой».

В нашем случае в зоны неопределённости попадает часть наблюдений после 8-го и перед 42-ым порядковыми номерами (см. таблицу 6.6). То есть значения, ближайšie к гр. 1, и значения, ближайšie к гр. 3, будут вносить элемент «зашумления» в процедуру сравнения групп.

Поэтому в качестве средней группы мы сформировали гр. 2, сопоставимую по численности с группами 1 и 3 и включающую те значения показателя  $T$ , которые сосредоточены около средневывборочного значения ( $\bar{X} = 13,3$  с)



и медианы ( $Me = 13,2$  с). Это будут значения с порядковыми номерами примерно от 21-го до 29-го включительно<sup>6.23</sup>.

Принцип группирования схематично представлен на рисунке 6.6.

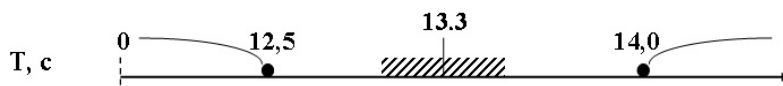


Рис. 6.6. Схема группирования по значениям показателя  $T$ , с

Мы целенаправленно формировали типовые группы 1, 2, 3 по уровню достижений кандидатов в беге на 100-метровую дистанцию (лучшие – средние – худшие) и, тем самым, априори заложили высокую вероятность различия кандидатов (наличествующих и предполагаемых) по показателю  $T$ .

*Вопрос:* обусловливает ли вообще заведомо качественное различие кандидатов по показателю  $T$  одновременно их качественное различие по показателю  $MAM$ ?

В рамках генеральной совокупности, из которой наша выборка (а в полном варианте — эталонная выборка стандартизации), возникает альтернатива:

- различие по показателю  $MAM$  у кандидатов является *делом случая*, результатом вмешательства случайных факторов (нулевая гипотеза  $H_0$ );
- различие по показателю  $MAM$  у кандидатов сопряжено с разным уровнем их достижений в спринтерском беге (рабочая гипотеза  $H_1$ ).

Чтобы доказать справедливость гипотезы  $H_1$ , надо опровергнуть гипотезу  $H_0$  на уровне генеральной совокупности.

Здесь уместно напомнить, что все практическое применение теории вероятностей основывается на «практически невозможных» событиях, «практически достоверных» событиях и «принципе практической уверенности» [6.1; 6.19]<sup>6.24</sup>. И если два случайных события составляют полную группу событий,

<sup>6.23</sup>Для последующего анализа целесообразно формировать группы, не только обладающие определенным качественным уровнем исследуемого показателя (лучшее, среднее, худшее), но и сопоставимые по объёму. С этой целью из наиболее многочисленной средней группы выбиралась та ее часть, которая подчёркивала присущее группе качество исследуемого показателя.

В нормально распределенной эталонной выборке обязательного объёма ( $n \geq 200$ ) средневыворочное значение окажется к тому же центральным значением (медианой). Либо, по крайней мере, различие между средним и медианой будет пренебрежимо малым.

<sup>6.24</sup>Не путать с невозможными ( $p = 0$ ) и достоверными ( $p = 1$ ) событиями.

при этом вероятность одного из них близка к нулю ( $p_1 \approx 0$ ), то это событие классифицируется как практически невозможное, а другое — как практически достоверное ( $p_2 \approx 1$ )<sup>6.25</sup>.

То есть, когда на уровне генеральной совокупности вероятность гипотезы  $H_0$  ничтожно мала, можно классифицировать гипотезу как «практически невозможную».

К сожалению, нам неизвестно истинное значение этой вероятности. А вопрос о том, насколько малым оно должно быть, выходит за рамки математической теории и регулируется лишь принципом практической уверенности исследователя [6.19].

Поэтому мы будем рассматривать не собственно вероятность гипотезы  $H_0$ , а меру доверия к ней и сами должны определиться с *порогом ее малости*, который принято называть «уровнем значимости» и обозначать  $\alpha$ .

Назначая  $\alpha$ <sup>6.26</sup>, мы допускаем, что в генеральной совокупности имеются случаи несоответствия уровней  $T$  и  $MAM$  у кандидатов. Однако предполагаем количество этих случаев таким, что вероятность их появления в многочисленных испытаниях будет пренебрежимо малой и не окажет влияния на закономерность исходов.

Под многочисленными испытаниями подразумевается процесс многократного (теоретически исчисляемого тысячами) повторения опыта, состоящего в *случайном* выборе кандидата из генеральной совокупности (с его обязательным возвращением назад для соблюдения одних и тех же условий опыта в процедуре их многократного повторения)<sup>6.27</sup>.

Если предположить, что в генеральной совокупности доля кандидатов с несоответствием уровней  $T$  и  $MAM$  составляет, скажем, 0,01 (условно говоря, один такой кандидат на сотню), то из тысяч исходов случайного выбора кан-

---

<sup>6.25</sup>Заметим, что именно на этом фундаментальном принципе построен классический статистический частотный анализ.

<sup>6.26</sup>Величина  $\alpha$  назначается исследователем настолько малой (близкой к нулю), чтобы возможно было полагать, что случайное событие, наступление которого в многочисленных испытаниях ожидается с вероятностью ещё меньшей, чем  $\alpha$ , в единичном испытании не должно наступить (считается практически невозможным согласно принципу практической невозможности маловероятных событий [6.1; 6.19]). Выбор степени малости  $\alpha$  зависит всецело от личного опыта исследователя, его знаний об изучаемом событии (явлении), а также от характера решаемой задачи, объёма и качества экспериментальных данных.

<sup>6.27</sup>Аналогичные задачи являются типовыми в теории математической статистики.

дидата приблизительно 0,01 доля будет приходиться на рассматриваемый случай (статистическая вероятность составит 0,01). То есть на уровне генеральной совокупности закономерным исходом с высокой степенью вероятности окажется как раз выбор кандидата с сопоставимыми уровнями  $T$  и  $MAM$ .

Именно такую закономерность нам и предстоит обосновать, располагая лишь выборкой из генеральной совокупности.

Разумеется истинная доля кандидатов с несоответствием уровней  $T$  и  $MAM$  в генеральной совокупности нам также неизвестна. Но мы имеем основание полагать эту долю весьма незначительной.

Итак, состоятельность гипотезы  $H_0$  будем рассматривать в контексте *исключения неконтролируемого случая* в объяснении результатов сравнительного анализа групп 1, 2, 3 по показателю  $MAM$ <sup>6.28</sup>.

Сравнительный анализ проведём с учётом следующих ограничений и допущений:

- для малых выборок применимы только непараметрические критерии;
- уровень значимости  $\alpha$  определим равным 0,01;
- расчётные  $p$ -значения округлим до четвёртого знака после запятой.

В ходе проверки гипотезы будем руководствоваться следующими основными положениями:

- 1)  $\alpha$  — теоретический минимум вероятности нулевой гипотезы в многочисленных испытаниях на уровне генеральной совокупности, ниже которого гипотеза ( $H_0$ ) расценивается уже как практически невозможная;
- 2)  $p$ -значение служит рассчитываемой *мерой доверия* к нулевой гипотезе ( $H_0$ ) в диапазоне всех возможных значений  $MAM$ , обусловленном именно выборкой стандартизации<sup>6.29</sup>. Эта мера является свидетельством либо против нулевой гипотезы, либо в ее пользу. Под  $p$ -значением *подразумевается* вероятность того, что в пределах этого диапазона у кандидатов (наличествующих и предполагаемых), имеющих разные уровни достижений в спринтерском беге  $T$ , соответствующее качественное раз-

---

<sup>6.28</sup>Проверка статистической гипотезы означает проверку согласования исходных выборочных данных с выдвинутой гипотезой.

<sup>6.29</sup>В случае эталонной выборки диапазон возможных значений  $MAM$  будет сопоставим с аналогичным диапазоном в генеральной совокупности со всеми вытекающими последствиями.

личие анаэробных возможностей *МАМ* является делом случая, но не закономерностью;

- 3) если «случай» не даст правдоподобного объяснения результатов сравнения кандидатов по показателю *МАМ* (рассчитанное  $p$ -значение окажется меньше назначенного нами уровня значимости  $\alpha = 0,01$ ), то более вероятным объяснением становится рабочая гипотеза  $H_1$ .

При этом классифицировать гипотезу  $H_0$  как практически невозможную, конечно, не сможем, поскольку рассчитанное  $p$ -значение не является вероятностью этой гипотезы. Но имеем достаточное основание *отклонить* ее, признав гипотезу  $H_1$  хотя и не практически достоверной, но *статистически значимой*<sup>6.30</sup>.

Результаты сравнительного анализа приведены в таблице 6.7 и на рисунке 6.7<sup>6.31</sup>. Выборки представлены через медиану  $Me$  и размах  $[x_{min}; x_{max}]$ . Символом «\*» отмечены случаи выявления статистической значимости различия на уровне значимости  $\alpha = 0,01$ .

Здесь уместно пояснить, что хотя доверительные интервалы предпочтительнее при описании результатов исследования, однако в рамках наших упрощённых данных это некорректно ввиду намеренно малого объёма сравниваемых групп.

Расчёты показали (см. таблицу 6.7), что уровень доверия  $p$  к предположению о том, что обнаруженное различие групп 1, 2, 3 по исследуемым показателям является делом случая ( $H_0$ ), меньше принятого уровня значимости ( $p \ll \alpha = 0,01$ ). Это свидетельствует в пользу гипотезы  $H_1$ <sup>6.32</sup>.

Таким образом, в соответствии с основными положениями проверки выдвинутой гипотезы (см. выше) применительно к результатам проведённого сравнительного анализа (таблица 6.7 и рисунок 6.7) можно констатировать, что:

- все три группы кандидатов 1, 2, 3 статистически значимо (в данном случае *практически достоверно*) различались по скорости бега на 100-

<sup>6.30</sup> Вполне очевидно, что от репрезентативности выборки стандартизации зависит не только уровень доверия к нулевой гипотезе, но и уровень доверия к результатам исследования в целом.

<sup>6.31</sup> Расчёты выполнены в пакете SPSS. Выбрано асимптотическое приближение.

<sup>6.32</sup> Математический символ « $\ll$ » обозначает «много меньше». Хотя принципиально важно не то, насколько меньше, а то, что именно «меньше».

**Сравнительная характеристика групп кандидатов 1, 2, 3 по исследуемым показателям  
(критерий Манна-Уитни)**

Показатель	Описательные статистики данных $Me[x_{min}; x_{max}]$			Статистики парных сравнений		
	гр. 1 (n=8)	гр. 2 (n=9)	гр. 3 (n=9)	$p_{1,2}$	$p_{1,3}$	$p_{2,3}$
$T$ , сек.	12,3 [12;12,5]	13,2 [13,1;13,4]	14,1 [14;14,8]	0,0004*	0,0005*	0,0003*
$MAM$ , кгм/с	113,8 [100;142]	96,2 [78;99,8]	79,0 [64;89]	0,0015*	0,0005*	0,0031*

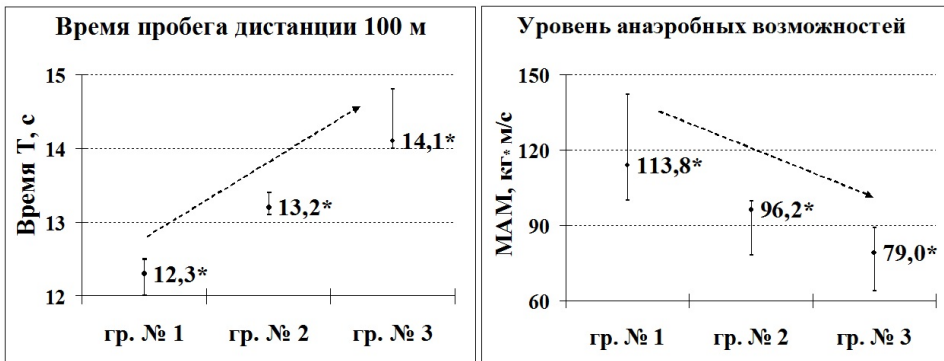


Рис. 6.7. Характер связи между показателем быстроты  $T$  при беге на 100-метровую дистанцию и уровнем анаэробных возможностей  $MAM$  у кандидатов

метровую дистанцию ( $T$ , медианы 12,3 с, 13,2 с, 14,1 с соответственно;  $p_{1,2} = 0,0004 < \alpha$ ,  $p_{1,3} = 0,0005 < \alpha$ ,  $p_{2,3} = 0,0003 < \alpha$ ), что вполне согласуется с принципом их формирования;

- чем хуже были групповые достижения кандидатов в спринтерском беге (время  $T$  больше), тем ниже (меньше) был уровень анаэробных возможностей ( $MAM$ , медианы 116,5 кг·м/с, 98,0 кг·м/с, 79,0 кг·м/с соответственно;  $p_{1,2} = 0,0015 < \alpha$ ,  $p_{1,3} = 0,0005 < \alpha$ ,  $p_{2,3} = 0,0031 < \alpha$ ).

Более того, с учётом поправки Бонферрони ( $\alpha' = 0,01/3 \approx 0,0033$ ) по данным той же таблицы 6.7 и рисунка 6.7 можно утверждать, что среди трёх групп кандидатов 1, 2, 3 совсем не случайно:

- группа 1 оказалась лучшей не только по значениям показателя  $T$  (по факту формирования групп), но и по уровню  $MAM$ . Так как  $p_{1,2} = 0,0015 < \alpha'$  и  $p_{1,3} = 0,0005 < \alpha'$ , то этот эффект различия невозможно объяснить волей случая;
- группа 3 оказалась худшей не только по значениям показателя  $T$  (по факту формирования групп), но и по уровню  $MAM$ . Так как  $p_{1,3} = 0,0005 < \alpha'$  и  $p_{2,3} = 0,0031 < \alpha'$ , то и этот эффект различия также невозможно объяснить волей случая.

То есть при принятой нами классификации кандидатов выявлена высокая степень вероятности следующей закономерности:

- кандидаты с лучшими достижениями в спринтерском беге обладают лучшими анаэробными возможностями;
- кандидаты с худшими достижениями в спринтерском беге обладают худшими анаэробными возможностями;
- кандидаты со средними достижениями в спринтерском беге обладают средними анаэробными возможностями.

*Что и требовалось обосновать.*

#### 6.4.1.2. Шаг В

Теперь осуществим обратное сравнение по уже выстроенной последовательности рассуждений (см. шаг А). При этом некоторые подробности будем опускать, чтобы не повторяться.

На данном этапе требуется произвести стратификацию по уровню анаэробных возможностей кандидатов  $MAM$ , кгм/с.

Упорядочим выборку экспериментальных данных (см. таблицу 6.5) по возрастанию показателя  $MAM$  и пронумеруем полученный ряд наблюдений. Упорядоченный вариант данных представлен в таблице 6.8.

Таблица 6.8

**Экспериментальные данные, упорядоченные по возрастанию показателя  $MAM$  (n=50)**

Присвоенный номер	Кандидат, ФИО	$MAM$ , кгм/с	$T$ , с
1	ФИО-41	64,0	14,4

Продолжение таблицы 6.8

Присвоенный номер	Кандидат, ФИО	МAM, кгм/с	T, с
2	ФИО-48	70,0	14,1
3	ФИО-26	75,0	14,0
4	ФИО-12	78,0	14,2
5	ФИО-49	78,0	13,2
6	ФИО-34	79,0	14,8
7	ФИО-37	79,0	14,0
8	ФИО-01	80,0	14,0
9	ФИО-47	80,0	13,2
10	ФИО-07	80,5	14,1
11	ФИО-39	81,0	14,1
12	ФИО-02	84,0	13,7
13	ФИО-31	86,0	13,2
14	ФИО-04	88,0	13,0
15	ФИО-06	88,0	13,6
16	ФИО-17	89,0	14,2
17	ФИО-25	90,0	13,4
18	ФИО-11	91,0	13,7
19	ФИО-40	92,0	13,6
20	ФИО-27	92,6	13,6
21	ФИО-33	93,0	13,4
22	ФИО-35	94,5	13,8
23	ФИО-15	96,0	13,4
24	ФИО-36	96,0	13,5
25	ФИО-32	96,2	13,2
26	ФИО-08	97,0	13,1
27	ФИО-22	97,0	12,6
28	ФИО-45	98,3	13,2
29	ФИО-20	98,6	13,3
30	ФИО-10	99,7	13,2
31	ФИО-23	99,8	13,1
32	ФИО-05	100,0	12,3
33	ФИО-14	100,0	13,0

Продолжение таблицы 6.8

Присвоенный номер	Кандидат, ФИО	МAM, кгм/с	T, с
34	ФИО-42	100,0	13,0
35	ФИО-13	103,0	12,3
36	ФИО-16	104,0	12,3
37	ФИО-19	104,0	13,9
38	ФИО-50	104,0	13,0
39	ФИО-30	107,0	12,6
41	ФИО-38	110,0	12,8
40	ФИО-09	110,0	13,4
42	ФИО-03	111,0	12,9
43	ФИО-29	113,0	12,4
44	ФИО-21	114,6	12,5
45	ФИО-28	115,0	12,9
46	ФИО-44	120,5	12,5
47	ФИО-18	125,0	12,6
48	ФИО-24	137,0	12,6
49	ФИО-46	140,0	12,0
50	ФИО-43	142,0	12,0

Выделим крайние группы кандидатов: с худшими (наименьшими) и лучшими (наибольшими) значениями показателя *МAM*. Для этого рассчитаем порядковые номера процентилей с процентильными уровнями 15,87 и 84,13 по формуле (6.1) и найдём их значения в таблице 6.8<sup>6.33</sup>:

- 1) перцентиль уровня 15,87 находится в таблице 6.8 на восьмом месте и равна 80,0 кг.м/с;
- 2) перцентиль уровня 84,13 находится в таблице 6.8 на 42-ом месте и равна 111,0 кг.м/с.

Таким образом, группу «худших возможностей» по показателю *МAM* представили кандидаты с порядковыми номерами от первого до восьмого включительно (гр. 4), группу «лучших возможностей» представили кандида-

<sup>6.33</sup>Для выборки *МAM* также предполагается гипотетическое нормальное распределение, на которое она может отобразиться.



ты с порядковыми номерами от 42-го до 50-го включительно (гр. 6). В общей сложности 17 кандидатов (см. таблицу 6.8).

В качестве средней группы мы сформируем группу 5, сопоставимую по численности с группами 4 и 6 и включающую те значения показателя  $MAM$ , которые сосредоточены около средневыборочного значения ( $\bar{X} = 97,4$  кгм/с) и медианы ( $Me = 97,4$  кгм/с: это значения с порядковыми номерами примерно от 23-го до 30-го включительно).

Принцип группирования схематично представлен на рисунке 6.8.

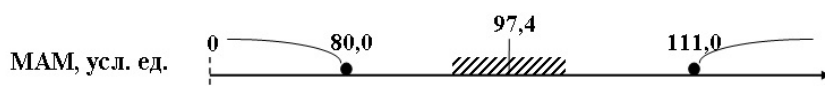


Рис. 6.8. Схема группирования по значениям показателя  $MAM$

Мы целенаправленно формировали типовые группы 4, 5, 6 по уровню анаэробных возможностей у кандидатов (*худшие – средние – лучшие*) и, тем самым, априори заложили высокую вероятность различия кандидатов (наличествующих и предполагаемых) по показателю  $MAM$ .

*Вопрос:* обусловливает ли заведомо качественное различие кандидатов по показателю  $MAM$  одновременно их качественное различие по показателю  $T$ ?

В рамках генеральной совокупности, из которой наша выборка (а в полном варианте — эталонная выборка стандартизации), возникает альтернатива:

- а) различие по достижениям в спринтерском беге  $T$  у кандидатов является *делом случая*, результатом вмешательства случайных факторов (нулевая гипотеза  $H_0$ );
- б) различие по достижениям в спринтерском беге  $T$  является *закономерным следствием* разного качественного уровня анаэробных возможностей у кандидатов (рабочая гипотеза  $H_1$ ).

Чтобы доказать справедливость гипотезы  $H_1$ , надо опровергнуть гипотезу  $H_0$  на уровне генеральной совокупности.

Если на уровне генеральной совокупности вероятность гипотезы  $H_0$  ничтожно мала, можно классифицировать ее как «практически невозможную».

Однако нам неизвестно истинное значений этой вероятности. Поэтому будем рассматривать не собственно вероятность гипотезы  $H_0$ , а меру доверия к ней  $p$  и сами определимся с *порогом ее малости* <sup>6.34</sup>.

Как уже было отмечено, чем меньше в генеральной совокупности доля кандидатов с несоответствием уровней  $MAM$  и  $T$ , тем меньше вероятность их появления в многочисленных испытаниях, состоящих в случайном выборе кандидата. Будем придерживаться ранее принятой версии о том, что в генеральной совокупности доля кандидатов с несоответствием уровней  $MAM$  и  $T$  составляет примерно 0,01. Из тысяч исходов случайного выбора кандидата из такой совокупности *закономерным исходом* с высокой долей вероятности можно ожидать выбор кандидата наоборот с сопоставимыми уровнями  $MAM$  и  $T$ . Именно эту закономерность нам предстоит обосновать, располагая выборкой из генеральной совокупности.

Проверку гипотезы  $H_0$  выполним в контексте исключения неконтролируемого случая в объяснении результатов сравнительного анализа групп 4, 5, 6 по показателю  $T$  с учётом принятых ранее ограничений и допущений.

В ходе проверки гипотезы будем руководствоваться ранее указанными основными положениями относительно  $\alpha$ ,  $p$ -значения, нулевой  $H_0$  и и рабочей  $H_1$  гипотез.

Результаты сравнительного анализа представлены в таблице 6.9 и на рисунке 6.9.

Таблица 6.9

**Сравнительная характеристика групп кандидатов 4, 5, 6 по исследуемым показателям (критерий Манна-Уитни)**

Показатель	Описательные статистики данных $Me[x_{min}; x_{max}]$			Статистики парных сравнений		
	гр. 4 (n=8)	гр. 5 (n=9)	гр. 6 (n=9)	$p_{4,5}$	$p_{4,6}$	$p_{5,6}$
$MAM$ , кгм/с	78 [64;80]	97 [96;99,7]	120,5 [111;142]	0,0008*	0,0005*	0,0014*
$T$ , сек.	14,1 [13,2;14,8]	13,2 [12,6;13,5]	12,5 [12;12,9]	0,0015*	0,0005*	0,0031*

<sup>6.34</sup>Выше (шаг А) были приведены необходимые теоретические пояснения.

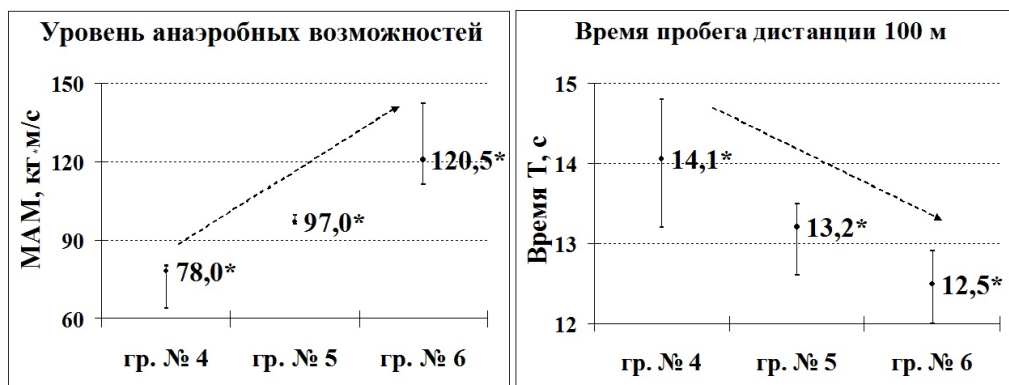


Рис. 6.9. Характер связи между уровнем анаэробных возможностей  $MAM$  и показателем быстроты  $T$  при беге на дистанцию 100 м у кандидатов

Расчёты показали (см. таблицу 6.9), что уровень доверия  $p$  к предположению о том, что обнаруженное различие групп 4, 5, 6 по исследуемым показателям является делом случая ( $H_0$ ), меньше принятого уровня значимости ( $p \ll \alpha = 0,01$ ). Это свидетельствует в пользу рабочей гипотезы ( $H_1$ ).

Таким образом, в соответствии с основными положениями проверки гипотезы (см. выше) применительно к результатам проведённого сравнительного анализа (таблица 6.9 и рисунок 6.9) можно констатировать следующее:

- все три группы кандидатов 4, 5, 6 статистически значимо (в данном случае *практически достоверно*) различались по показателю анаэробных возможностей ( $MAM$ , медианы 78,0 кг.м/с, 97,0 кг.м/с, 120,5 кг.м/с соответственно,  $p_{4,5} = 0,0008 < \alpha$ ,  $p_{4,6} = 0,0005 < \alpha$ ,  $p_{5,6} = 0,0005 < \alpha$ ), что вполне согласуется с принципом их формирования;
- чем лучше (больше) был уровень анаэробных возможностей у кандидатов  $MAM$ , тем выше их достижения (меньше  $T$ ) в беге на 100-метровую дистанцию ( $T$ , медианы 14,1 с, 13,2 с, 12,5 с соответственно;  $p_{4,5} = 0,0033 < \alpha$ ,  $p_{4,6} = 0,0005 < \alpha$ ,  $p_{5,6} = 0,0014 < \alpha$ ).

С учётом поправки Бонферрони ( $\alpha' = 0,01/3 \approx 0,0033$ ) по данным той же таблицы 6.9 можно утверждать, что среди трёх групп кандидатов 4, 5, 6 совсем не случайно:

- группа 4 оказалась худшей не только по уровню  $MAM$  (по факту формирования групп), но и по значениям показателя  $T$ . Так как  $p_{4,5} = 0,0033 \equiv$

$\alpha'$  и  $p_{4,6} = 0,0005 < \alpha'$ , то этот эффект различия невозможно объяснить волей случая;

- группа 6 оказалась лучшей не только по уровню МАМ (по факту формирования групп), но и по значениям показателя  $T$ . Так как  $p_{4,6} = 0,0005 < \alpha'$  и  $p_{5,6} = 0,0014 < \alpha'$ , то и этот эффект различия также невозможно объяснить волей случая.

То есть при принятой нами классификации кандидатов выявлена высокая степень вероятности следующей закономерности:

- кандидатам с худшими анаэробными возможностями присущи худшие достижения в спринтерском беге;
- кандидатам с лучшими анаэробными возможностями присущи лучшие достижения в спринтерском беге;
- кандидатам со средними анаэробными возможностями присущи средние достижения в спринтерском беге.

*Что и требовалось обосновать.*

Итак, результаты перекрёстного сравнительного анализа (шаг А, шаг В) свидетельствуют о том, что между уровнем анаэробных возможностей кандидатов и уровнем достижений в беге на короткие дистанции существует взаимное соответствие. Поэтому при отборе кандидатов на специальность, требующую выполнения работы большой мощности в минимальное время, *мы можем бег на короткие дистанции заменить тестированием по физиологической методике МАМ в стационарных условиях.*

#### **6.4.2. Разработка критерия оценки уровня МАМ у кандидатов**

Разработаем критерий оценки *уровня* анаэробных возможностей у кандидатов по значениям показателя МАМ.

С этой целью произведём шкалирование диапазона возможных значений МАМ. За основу примем шкалу стенов.

Прежде всего, выделим выборку стандартизации, упорядоченную по показателю МАМ, на котором мы разрабатываем целевой показатель (таблица 6.10).

Таблица 6.10

## Выборка стандартизации (n=50)

Присвоенный номер	Кандидат, ФИО	МММ, кгм/с
1	ФИО-41	64,0
2	ФИО-48	70,0
3	ФИО-26	75,0
4	ФИО-12	78,0
5	ФИО-49	78,0
6	ФИО-34	79,0
7	ФИО-37	79,0
8	ФИО-01	80,0
9	ФИО-47	80,0
10	ФИО-07	80,5
11	ФИО-39	81,0
12	ФИО-02	84,0
13	ФИО-31	86,0
14	ФИО-04	88,0
15	ФИО-06	88,0
16	ФИО-17	89,0
17	ФИО-25	90,0
18	ФИО-11	91,0
19	ФИО-40	92,0
20	ФИО-27	92,6
21	ФИО-33	93,0
22	ФИО-35	94,5
23	ФИО-15	96,0
24	ФИО-36	96,0
25	ФИО-32	96,2
26	ФИО-08	97,0
27	ФИО-22	97,0
28	ФИО-45	98,3
29	ФИО-20	98,6
30	ФИО-10	99,7
31	ФИО-23	99,8

Продолжение таблицы 6.10

Присвоенный номер	Кандидат, ФИО	МММ, кгм/с
42	ФИО-03	111,0
32	ФИО-05	100,0
33	ФИО-14	100,0
34	ФИО-42	100,0
35	ФИО-13	103,0
36	ФИО-16	104,0
37	ФИО-19	104,0
38	ФИО-50	104,0
39	ФИО-30	107,0
40	ФИО-09	110,0
41	ФИО-38	110,0
43	ФИО-29	113,0
44	ФИО-21	114,6
45	ФИО-28	115,0
46	ФИО-44	120,5
47	ФИО-18	125,0
48	ФИО-24	137,0
49	ФИО-46	140,0
50	ФИО-43	142,0

Затем в соответствии с таблицей 6.3 по приведённым в ней процентильным уровням рассчитаем значения  $j$  — порядковые номера соответствующих процентилей в упорядоченной выборке стандартизации. По рассчитанным номерам идентифицируем в упорядоченной выборке стандартизации искомые процентиля. В таблице 6.11 представлена вся последовательность расчётов.

Далее «привяжем» рассчитанные процентиля к правым границам стандартных интервалов, то есть стен.

На рисунке 6.10 наглядно представлено это соответствие<sup>6.35</sup>.

<sup>6.35</sup>Стандартные интервалы (стены) разграничены вертикальными штрихами и пронумерованы. Под правым концом каждого стена (до девятого включительно) указано соответствующее значение процентиля («сырого» значениями МММ).

Расчёт процентилей в выборке стандартизации ( $n=50$ )

Номер стена (i)	Соответствующий процентильный уровень ( $P_i$ )	Порядковый номер процентили в выборке ( $j$ )	Значение процентили, кг·м/с
1	2,28	$j = 50/100 \times 2,28 \approx 1$	64,0
2	6,68	$j = 50/100 \times 6,68 \approx 3$	75,0
3	15,87	$j = 50/100 \times 15,87 \approx 8$	80,0
4	30,85	$j = 50/100 \times 30,85 \approx 15$	88,0
5	50	$j = 50/100 \times 50 \approx 25$	96,2
6	69,15	$j = 50/100 \times 69,15 \approx 35$	103,0
7	84,13	$j = 50/100 \times 84,13 \approx 42$	111,0
8	93,32	$j = 50/100 \times 93,32 \approx 47$	125,0
9	97,72	$j = 50/100 \times 97,72 \approx 49$	140,0
10	—	не требуется	—

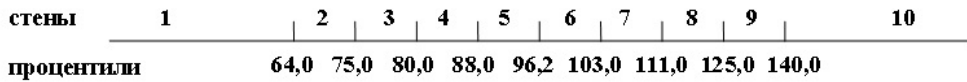


Рис. 6.10. Схема соответствия между стенами и процентиями показателя МАМ в выборке стандартизации

Теперь ещё раз обратимся к рисунку 6.3, мысленно спроецируем на него наш рисунок 6.10 и акцентируем внимание на следующих важных моментах.

«Хвосты» МАМ будут охватывать:

- слева первый, второй и третий стены (это худшие значения; нижний порог 80,0 кгм/с);
- справа восьмой, девятый и десятый стены (это лучшие значения; верхний порог 111,0 кгм/с).

«Норма» укладывается в пятый и шестой стены (от 88,0 кгм/с до 103,0 кгм/с).

Зону неопределённости составят значения четвёртого и седьмого стен.

Таким образом, мы имеем своеобразное «расслоение» всех возможных значений МАМ на интервалы, что и позволяет разработать критерий оцен-

ки уровня индивидуальных анаэробных возможностей кандидатов по номинальной шкале (шкале категорий).

Необходимо уточнить место четвёртого и седьмого стандартных интервалов в общей картине категорий. При этом *обеспечить непротиворечивость оценки.*

Поскольку седьмой стен (см. рисунок 6.1) составляет зону неопределённости между «нормой» и очень хорошими значениями, то логично будет отнести значения *МАМ*, встроенные в седьмой стен, к «норме», что *не порождает противоречия.*

Так как четвёртый стен составляет зону неопределённости между худшими значениями и «нормой», то значения *МАМ*, встроенные в четвёртый стен, логично интерпретировать как дополнительную категорию, что *исключит противоречие.*

На основании вышеизложенного формулируем правило оценки уровня анаэробных возможностей у кандидатов (таблица 6.12).

Таблица 6.12

**Критерий оценки уровня анаэробных возможностей у кандидатов**

Значение <i>МАМ</i> , кг.м/с	Уровень анаэробных возможностей
$МАМ \leq 80,0$	низкий
$80,0 < МАМ \leq 88,0$	ниже среднего
$88,0 < МАМ \leq 111,0$	средний
$МАМ > 111,0$	высокий

В соответствии с таблицей 6.12 представленный *метод позволяет перевести интервальную шкалу МАМ в надёжную номинативную и разработать на эталонной выборке стандарт категорий (целевой показатель) для оценки анаэробных возможностей кандидатов.*

Этот стандарт можно затем использовать как внешний критерий в различных задачах исследования медико-биологических данных методами управляемого машинного обучения (регрессия, классификация с учителем). А также, как частный вариант, решать задачу отбора кандидатов на специальность, требующую выполнения работы большой мощности в минимальное время.



### 6.4.3. Решающее правило отбора кандидатов на специальность по уровню анаэробных возможностей

Оценка пригодности кандидатов к специальности, требующей выполнения работы большой мощности в минимальное время, производится в соответствии с уровнем их индивидуальных анаэробных возможностей.

В целом решающее правило дифференциации кандидатов по уровню их пригодности к выполнению работы большой мощности в минимальное время включает:

- получение тестового значения *МAM* кандидата;
- оценку качественного уровня *МAM* у кандидата по полученному тестовому значению;
- отнесение индивидуальных анаэробных возможностей кандидата к определённой категории;
- по полученной категории *МAM* присвоение кандидату соответствующей группы пригодности.

Предлагается следующая классификация пригодности кандидатов к специальности (таблица 6.13)<sup>6.36</sup>.

Таблица 6.13

#### Правило дифференциации кандидатов по способности к выполнению работы высокой мощности в минимальное время (пригодность к специальности)

Значение <i>МAM</i> , кгм/с	Уровень анаэробных возможностей	Категория анаэробной мощности	Группа пригодности к специальности
$MAM \leq 80,0$	низкий	4	4 — не пригоден
$80,0 < MAM \leq 88,0$	ниже среднего	3	3 — условно пригоден
$88,0 < MAM \leq 111,0$	средний	2	2 — пригоден
$MAM > 111,0$	высокий	1	1 — пригоден в первую очередь

<sup>6.36</sup>Ещё раз напоминаем, что критические значения 80,0, 88,0 и 111,0 нельзя считать достоверными, поскольку они были получены не на эталонной выборке, а потому требуют уточнения. Нумерацию категорий и групп пригодности разработчик устанавливает сам.

Необходимо оценить прогностическую способность разработанного решающего правила. Известно, что для верификации помимо эталонной (обучающей) выборки требуется контрольная выборка. Обе выборки могут формироваться на одной базе экспериментальных данных достаточного объёма, причём могут быть взаимопроникающими. Однако это уже отдельный этап исследования, не предусмотренный нами в рамках данной работы.

### **6.5. Обобщение методики для шкалирования измеряемых физиологических показателей**

Изложим кратко схему моделирования категорий для отдельного измеряемого свойства (физиологического показателя) военнослужащих, основанную на использовании стандартных шкал.

- 1) Сформировать эталонную выборку стандартизации показателя  $X$  объёмом  $n \geq 200$ .
- 2) Упорядочить эталонную выборку по возрастанию значений показателя  $X$ .
- 3) Определить процентиля («сырые» значения показателя  $X$ ), соответствующие правым границам стен. Для этого заполнить таблицу 6.14.
- 4) Выделить интервалы (категории) качества показателя  $X$  (таблица 6.14):
  - 1-ый, 2-ой и 3-ий стены отнести к низкому уровню показателя. Этот интервал качества завершается процентилем уровня 15,87;
  - 4-ый стен отнести к уровню «ниже среднего». Этот интервал качества начинается сразу после процентиля уровня 15,87 и завершается процентилем уровня 30,85;
  - 5-ый, 6-ой и 7-ой стены отнести к *среднему уровню* показателя. Этот интервал качества начинается сразу после процентиля уровня 30,85 и завершается процентилем уровня 84,13;
  - 8-ой, 9-ый и 10-ый стены отнести к *высокому уровню* показателя. Этот интервал качества начинается сразу после процентиля уровня 84,13.

Схема расчёта процентилей в выборке стандартизации объёмом  $n$ 

Номер стена (i)	Уровень проценти (P <sub>i</sub> )	Порядковый номер проценти в выборке (j)	Значение проценти, ед. изм.	Уровень показателя, соответствующий стени
1	2,28	$j = n/100 \times 2,28$	$x_j$	низкий
2	6,68	$j = n/100 \times 6,68$	$x_j$	низкий
3	15,87	$j = n/100 \times 15,87$	$x_j$	низкий
4	30,85	$j = n/100 \times 30,85$	$x_j$	ниже среднего
5	50	$j = n/100 \times 50$	$x_j$	средний
6	69,15	$j = n/100 \times 69,15$	$x_j$	средний
7	84,13	$j = n/100 \times 84,13$	$x_j$	средний
8	93,32	$j = n/100 \times 93,32$	$x_j$	высокий
9	97,72	$j = n/100 \times 97,72$	$x_j$	высокий
10	—	не требуется	—	высокий

Данная модель предназначена для перевода «сырых» значений показателя в стени, что позволяет установить место каждого респондента в общем континууме возможных значений показателя.

В приведённой схеме моделирования особое значение отводится построению шкалы категорий измеряемого показателя.

В таблице 6.14 процентильные уровни ( $P=15,87$ ,  $P=30,85$  и  $P=84,13$ ) служат своего рода неизменными маркерами для различения категорий показателя.

В соответствии с этими маркерами процедуру шкалирования показателя можно свернуть:

- сформировать эталонную выборку стандартизации показателя  $X$  объёмом  $n \geq 200$ ;
- упорядочить эталонную выборку по возрастанию значений показателя  $X$ ;
- рассчитать порядковые номера  $j$  трёх целевых процентилей (таблица 6.15);
- отыскать в упорядоченной выборке «сырые» значения показателя  $X$  с этими порядковыми номерами — проценти;

- разметить интервальную шкалу показателя  $X$ , используя в качестве порогов отсечения выбранные процентиля (рисунок 6.11).

Таблица 6.15

**Определение целевых процентилей в выборке стандартизации**

Уровень процентили ( $P$ )	Порядковый номер процентили в выборке ( $j$ )	Процентиль ( $j$ -тое сырое значение)
15,87	$j = n/100 \times 15,87$	$x_j$
30,85	$j = n/100 \times 30,85$	$x_j$
84,13	$j = n/100 \times 84,13$	$x_j$

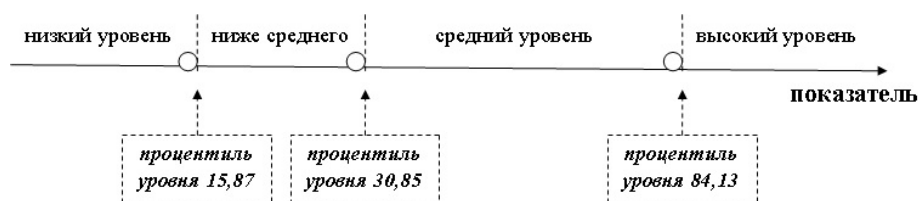


Рис. 6.11. Шкала категорий показателя

Подчеркнём, что приведённый алгоритм (см. таблицу 6.14) позволяет перевести интервальную шкалу измеряемого физиологического свойства кандидатов в номинативную шкалу (см. рисунок 6.11), на которой исследователем могут быть назначены другие пороги отсечения для оценки *уровня* этого свойства. Выбор процентильных уровней  $P$  для порогов отсечения зависит от решаемой задачи.

Стандарт, разработанный по предложенной схеме на выборке стандартизации измеряемого свойства кандидатов, может применяться затем в качестве внешнего критерия в медико-биологических исследованиях с применением методов управляемого машинного обучения. Например, при классификации «с учителем» (дискриминантный анализ) в задачах оценки и прогнозирования успешности военно-профессиональной деятельности разработанные уровни (категории) анаэробных возможностей военнослужащих будут служить известными классами.

## Выводы

Результативность машинного обучения при распознавании объектов (субъектов) по их свойствам в значительной степени зависит от того, насколько *хорошо понимаем образ*, то есть от полноты информации о свойствах объектов.

Неполнота информации о вариативности некоторого *измеряемого* свойства (*недостаточная формализованность данных*) может быть восполнена предварительным конструированием *целевого* (оценочного) показателя с помощью какого-либо релевантного способа. Один из способов извлечения дополнительной информации о подобном измеряемом свойстве — номинативное шкалирование его возможных значений.

В данной работе процесс шкалирования рассматривается в контексте определённого подхода к моделированию внешнего критерия для задач профессионального отбора.

Использовано существующее соответствие между двумя стандартными шкалами — шкалой степеней и шкалой процентилей — в качестве математического инструмента для категориального представления диапазона возможных значений измеряемого физиологического свойства, к которому предъявляются определённые требования при отборе кандидатов на специальность.

Предложен алгоритм доказательной замены одного физиологического показателя другим, пригодным для формирования внешнего критерия методом опосредованного шкалирования, с возможностью дальнейшей разработки решающих правил.

Предложенный подход может применяться для решения задач дифференциации респондентов (кандидатов на обучение, курсантов, слушателей и т. п.) по любому измеряемому свойству, нормально распределённому в генеральной совокупности (гипотетически).

Изложена методика перевода в номинативную шкалу диапазона возможных значений некоторого измеряемого свойства человека, позволяющая разработать на эталонной выборке стандарт категорий для оценки этого свойства и затем использовать этот стандарт как внешний критерий в медико-биологических исследованиях с применением методов управляемого машин-

ного обучения. Например, при классификации «с учителем» (дискриминантный анализ) в задачах оценки и прогнозирования успешности военно-профессиональной деятельности разработанные уровни (категории) анаэробных возможностей военнослужащих могут служить известными классами.

### Библиографический список

- 6.1. *Borovkov A. A.* Course of probability theory: textbook. — Moscow: Nauka, 1972. — 287 p. — (In Russian).
- 6.2. *Davison M.* Multidimensional scaling. — M.: Mir, 1986. — 224 p. — (In Russian).
- 6.3. *Davison M.* Multidimensional scaling. Methods of visual presentation of data. — M.: Finance, statistics, 1986. — 254 p. — (In Russian).
- 6.4. *Gusev E. K.* Psychophysics: textbook. — L. : Leningrad state University, 1987. — 84 p. — (In Russian).
- 6.5. *Lewis R., Galanter E.* Psychophysical scales. In sat.: Psychological measurement. — M.:Mir, 1967. — 196 p. — (In Russian).
- 6.6. *Nasledov A. D.* Mathematical methods of psychological research. Analysis and interpretation of data: a training manual. Vol. 392. — 3rd ed. — 2008. — (In Russian).
- 6.7. *Nikandrov V. V.* A method of modeling in psychology. — St. Petersburg, 2003. — 167 p. — (In Russian).
- 6.8. *Nikandrov V. V.* Experimental psychology. — SPb.:Publishing House "Rech", 2003. — 480 p. — (In Russian).
- 6.9. *Nikandrov V. V.* Nonempirical methods of psychology: a training manual. — SPb.:Rech, 2003. — 53 p. — (In Russian).
- 6.10. *Orlov A. I.* Econometrics. — Moscow: Exam, 2002. — 576 p. — (In Russian).

- 6.11. *Orlov A. I.* On real possibilities of bootstap as a statistical method // Factory laboratory. — 1987. — Vol. 53, no. 10. — P. 82–85. — (In Russian).
- 6.12. *Pfanzagl I., Baumann V., Huber H.* Measurement Theory. — Moscow: Exam, 1976. — 248 p. — (In Russian).
- 6.13. *Schwartz B. V., Khrushchev S. V.* Data analysis using multidimensional scaling methods. — M.: FiS, 1984. — 246 p. — (In Russian).
- 6.14. *Solodkov A. S., Sologub E. B.* Human physiology. Sports. Age: a textbook for higher educational institutions of physical culture. — M.: Terra-Sport, 2001. — 520 p. — (In Russian).
- 6.15. *Sologub E. B., Taymazov V. A.* Sports genetics: textbook for higher educational institutions of physical culture. — M.: Terra-Sport, 2000. — 127 p. — (In Russian).
- 6.16. *Stanton A., Glantz P. D.* Medico-biological statistics. — M.: Practice, 1999. — 459 p. — (In Russian).
- 6.17. *Terekhina Y.* Data analysis using multidimensional scaling methods. — M.: Nauka, 1986. — 168 p. — (In Russian).
- 6.18. The use of mathematical methods in the tasks of professional selection and distribution of personnel: a handbook / S. V. Wenceslaw [et al.]. — L.: Military Medical Academy, 1987. — 140 p. — (In Russian).
- 6.19. *Wentzel E. S.* Theory of Probability: textbook for higher technical educational Institutions. — Moscow: Nauka, 1964. — 572 p. — (In Russian).
- 6.20. *Zarochentsev K. D.* Fundamentals of psychometrics. — SPb, 1996. — 324 p. — (In Russian).
- 6.21. *Zarochentsev K. D., Khudyakov A. I.* Experimental psychology: textbook. — M.: TK Velbi, Publishing house Prospect, 2005. — 208 p. — (In Russian).