

## **МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ ЛОКОМОТИВНЫХ БРИГАД ДИРЕКЦИИ ТЯГИ ОКТЯБРЬСКОЙ ЖЕЛЕЗНОЙ ДОРОГИ ФИЛИАЛА ОАО «РЖД»**

**Седов М.С., Гадасина Л.В., Куранова О.Н.**

*ФГБОУ ВПО Петербургский государственный университет путей сообщения Императора Александра I, г. Санкт-Петербург, (190031, С.-Петербург, Московский пр., 9), maarcss@gmail.com*

Был проведен анализ результатов массового тестирования персонала ТЧЭ (локомотивных бригад) Октябрьской железной дороги по основной инструкции «Инструкция по движению поездов и маневровой работе на железнодорожном транспорте РФ - Приложение №8 к приказу Минтранса России от 04.06.2012г. №162» и построена математическая модель оценивания результатов на основе использованием методов многомерного статистического анализа. Результатов применения предложенной модели может стать повышение эффективности обучения сотрудников, а также для построения карты их профессиональных навыков. Использование описанного метода также позволит выявлять плохо подготовленные кадры, в случае, когда модель будет использоваться в связке с системой дистанционного обучения. Модель может быть применена не только для оценивания результатов тестирования локомотивных бригад, но также и для оценки обучения других массовых специальностей на железной дороге.

Ключевые слова: статистический анализ, оценивание, дистанционное обучение, локомотивные бригады, тестирование, тест.

## **MATHEMATICAL MODEL OF EVALUATING THE TEST RESULTS RECEIVED BY OCTOBER RAILWAY LOCOMOTIVE CREWS**

**Sedov M.S., Gadasina L.V., Kuranova O.N.**

*Petersburg State University of Railway Transport of the Emperor Alexander I, St. Petersburg, (190031, St. Petersburg, Moscow Avenue., 9), maarcss@gmail.com*

On the basis of testing October Railway locomotive crews was developed the mathematical model of evaluating. This model could help to improve level of technical education on October Railway. The use of proposed method is less labor-intensive than the expert approach. The model is based on methods of multivariate statistical analysis. Method can be used to assess the skills of employees for the map of competences. In addition, this model will identify poorly trained staff when working in conjunction with e-learning system. Model can be applied not only for the evaluation of test results locomotive crews, but also to assess the learning of other employees of railway.

Keywords: Statistical analysis, evaluating, e-learning, locomotive crew, railway, testing, test

На основании Распоряжения от 14 февраля 2008 г. N 300р об утверждении стандарта ОАО "РЖД" в локомотивных депо Октябрьской железной дороги проводится техническая учеба сотрудников. При этом в соответствии с Указанием Т-17/у от 22 апреля 2011 г. разрешено проведение технических занятий с использованием системы дистанционного обучения (СДО) на базе ЦИИТО (Центра интернет/интранет технологий в обучении) ФГБОУ ВПО ПГУПС. Каждый сотрудник имеет доступ к учебным материалам через сети передачи данных ОАО «РЖД» и/или Интернет, что позволяет ему проходить обучение в удобное для него время и в удобном месте.

Сотрудникам локомотивных бригад необходимо присутствовать на технических занятиях несколько раз в месяц и проходить тестирование по итогам этих занятий. При этом

важной является задача анализа и оценивания результатов обучения. Это связано как с организационными проблемами, так и с экономическими, поскольку сотрудники, успешно завершившие изучение месячных тем, получают денежное вознаграждение.

Машинисты локомотивов имеют возможность повысить уровень своей заработной платы при условии повышения класса. На основании приказа 23Ц «О порядке проведения испытаний, выдачи свидетельств на право управления локомотивом, моторвагонным подвижным составом на путях общего пользования и присвоения класса квалификации машинистам локомотивов и моторвагонного подвижного состава» от 11 ноября 1997 г. для получения нового класса необходимо провести проверку знаний законодательства Российской Федерации, приказов, указаний, инструкций и иных актов МПС России, а также проверку знаний по порядку действий в нестандартных ситуациях. В процессе обучения с использованием СДО проверка знаний осуществляется постоянно.

В работе [5] был проведен анализ и оценка проходного балла тестирования. Однако, данный метод не учитывал то, что вопросы могут быть неравнозначными. Одним из способов решения этой проблемы является оценка значимости вопросов с помощью экспертной группы. Однако, такой способ требует существенных финансовых и организационных затрат и является фактически неприменимым на практике.

С другой стороны, мы имеем большой банк данных результатов тестирования машинистов, имеющих большой опыт работы. Качество их ответов позволит придать вопросам соответствующие веса, определяющие их значимость. Предлагаемый метод использует эти данные в качестве эталона для проведения оценки результатов тестирования претендентов на повышение класса.

В данной работе предлагается построение математической модели оценивания результатов тестирования, использующей дискриминантный анализ – метод классификации многомерных наблюдений в ситуации, когда имеется несколько обучающих выборок данных.

Первой выборкой является набор результатов тестирования машинистов, которые рассматриваются как эксперты. Второй выборкой – набор результатов помощников машинистов, заведомо не прошедших тестирование. Поскольку в работе [5] был получен проходной балл в интервале от 70% до 80%, то в качестве верхнего порогового значения нами выбран средний балл в 60%. Для помощников, набравших средний балл более 60%, проведем более детальный анализ, используя упомянутый выше дискриминантный метод многомерной классификации.

Использование дистанционного обучения при проведении технических занятий позволяет использовать достаточно широкий функционал современных информационных

технологий. Тестирование, проводимое при помощи СДО, дает возможность хранить результаты тестирования всех сотрудников, а также проводить анализ полученных результатов. Кроме того существует возможность дифференцирования полученных ответов на основе должности сотрудника.

В качестве результатов тестирования рассматриваются результаты проведенного теста по одной из основных инструкций на железном транспорте – «Инструкция по движению поездов и маневровой работе на железнодорожном транспорте РФ - Приложение №8 к приказу Минтранса России от 04.06.2012г. №162». В результате исследования были получены две обучающие выборки, построенные по одному и тому же набору вопросов.

Задача дискриминантного анализа относится к разделу многомерного статистического анализа. Она представляет собой задачу классификации многомерных наблюдений в случае, когда имеются так называемые обучающие выборки.

В общем виде задача формулируется следующим образом.

Пусть имеется 2 совокупности  $\pi_1$  и  $\pi_2$ , каждая из которых содержит  $n_i k$ -мерных наблюдений  $X_1^{(i)}, X_2^{(i)}, \dots, X_{n_i}^{(i)}$ ,  $i=1,2$ . Пусть вероятность того, что наблюдение ведется над индивидуумом из генеральной совокупности  $\pi_1$  равна  $\alpha$ , тогда соответствующая вероятность для генеральной совокупности  $\pi_2$  –  $(1-\alpha)$ .

Новым наблюдением является  $k$ -мерный вектор  $Y = (Y_1, Y_2, \dots, Y_k)$ . Необходимо установить правило, по которому этот вектор будет отнесен к одной из возможных совокупностей  $\pi_1$  или  $\pi_2$ .

В нашем случае обучающими выборками являются:

$\pi_1$  : Результаты тестирования машинистов.

$\pi_2$  : Результаты тестирования помощников машинистов, набравших средний балл 60% и ниже.

Требуется принять решение по данному результату (при сдаче новым сотрудником) для помощника машиниста, считать ли прохождение теста успешным (будет ли отнесен результат к совокупности  $\pi_1$ ).

Для решения поставленной задачи используем байесовский подход.

В случае, когда распределения вероятностей обучающих выборок известны, и плотности распределений равны  $p_1(\bar{x})$  и  $p_2(\bar{x})$ ,  $\bar{x} = (x_1, x_2, \dots, x_k)$ , соответственно, критерий принятия решения выглядит следующим образом.

Область значений  $R$  вектора  $Y$  разбивается на две части  $R = R_1 \cup R_2$ , определяемые как

$$\begin{aligned} R_1: & \alpha \cdot p_1(Y) \geq (1-\alpha) \cdot p_2(Y), \\ R_2: & \alpha \cdot p_1(Y) < (1-\alpha) \cdot p_2(Y). \end{aligned} \quad (1)$$

В случае, если наблюдение  $Y \in R_1$ , отнесем его к совокупности  $\pi_1$ , иначе - к совокупности  $\pi_2$ .

В задаче тестирования априорные вероятности неизвестны. Однако и в этом случае, метод Байеса является допустимым, т.е. таким, который нельзя улучшить.

### Экспериментальное исследование

Для теста, состоящего из 10 вопросов, выбранных из общего пула, были собраны результаты ответов машинистов и помощников, затем была проведена сравнительная оценка вероятностей правильных ответов на вопросы для этих двух групп. Полученные результаты приведены на рис.1. По диаграмме видно, что не все машинисты дают 100% ответы на вопросы, что позволяет на основе эмпирических данных судить о значимости каждого из вопросов теста для экспертной группы.

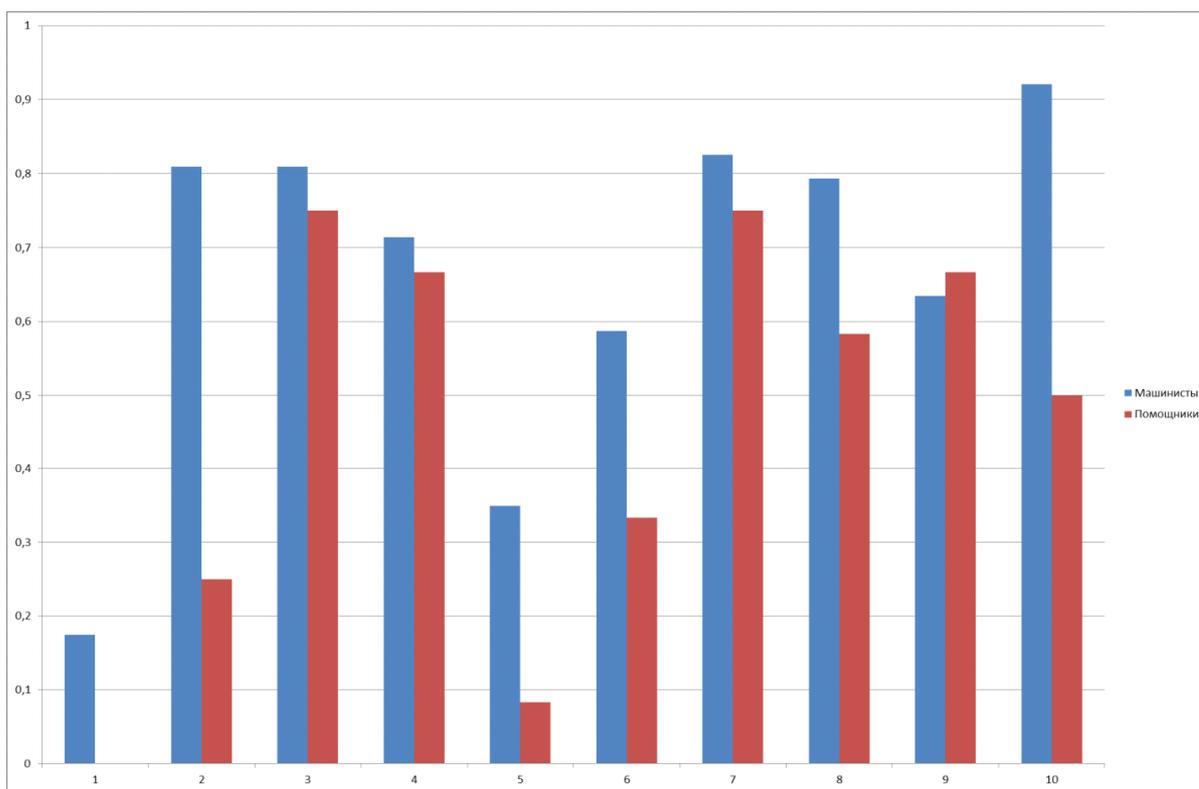


Рис.1. Сравнение оценок вероятностей получения верных ответов

В данном случае результаты тестирования являются случайными величинами дискретного типа, следовательно, вместо плотностей будем рассматривать распределения  $f_1(\bar{x}) = P(X_1^{(1)} = x_1, X_2^{(1)} = x_2, \dots, X_k^{(1)} = x_k)$  и  $f_2(\bar{x}) = P(X_1^{(2)} = x_1, X_2^{(2)} = x_2, \dots, X_k^{(2)} = x_k)$ . В ходе анализа будем считать выбранные вопросы тестирования независимыми между собой. Тогда каждое наблюдение состоит из 10 независимых случайных величин, имеющих распределение Бернулли таких, что

$$X_j^{(1)} = \begin{cases} 1, & p_{1j} \\ 0, & 1 - p_{1j} \end{cases}, \quad X_j^{(2)} = \begin{cases} 1, & p_{2j} \\ 0, & 1 - p_{2j} \end{cases}, \quad j = 1, 2, \dots, 10.$$

Вероятности правильных ответов  $p_{kj}, k = 1, 2, j = 1, 2, \dots, 10$  оцениваются по имеющимся выборкам данных.

Заметим, что никто из помощников машинистов не дал правильного ответа на первый вопрос, что дало  $p_{21} = 0$ . Для проведения дальнейшего анализа подправим полученное нулевое значение вероятности, заменив его на малое значение  $p_{21} = 10^{-4}$ . Тогда эмпирически оцененные распределения  $\hat{f}_1(\bar{x})$  и  $\hat{f}_2(\bar{x})$  будут иметь вид

$$\hat{f}_1(\bar{x}) = \prod_{j=1}^{10} p_{1j}^{x_j^{(1)}} (1 - p_{1j})^{1 - x_j^{(1)}}, \quad \hat{f}_2(\bar{x}) = \prod_{j=1}^{10} p_{2j}^{x_j^{(2)}} (1 - p_{2j})^{1 - x_j^{(2)}}.$$

Используя критерий (1), отнесем наблюдение к соответствующей совокупности. Поскольку у нас не имеется никаких априорных условий, позволяющих отнести нового претендента к одной из групп, положим в критерии (1) значение  $\alpha = \frac{1}{2}$ .

Проведя анализ результатов тестирования помощников машинистов, мы увидели в частности, что часть результатов со средним баллом 70% являются более близкими к эталонной группе и данные результаты теста приняты нашим методом как успешные. Тогда как другие результаты с тем же баллом были признаны неуспешными. Проведя оценку ошибки дискриминации с помощью следующей формулы

$$\frac{1}{2} \sum_{\bar{x} \in R_2} \hat{f}_1(\bar{x}) + \frac{1}{2} \sum_{\bar{x} \in R_1} \hat{f}_2(\bar{x}),$$

мы получили, что она составляет 6%.

В заключении заметим, что хотя в исследовании были рассмотрены выборки результатов тестов, состоящих только из 10 вопросов, метод может применяться для тестовых блоков любой длины.

## Список литературы

1. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физматгиз, 1963. – 500 с.
2. Дубров А.М., Мхитарян В.С., Трошин Л.И. Многомерные статистические методы. Учебник. М.: Финансы и статистика, 2003. – 352 с.
3. Кромер В.В. О некоторых вопросах тестовых технологий / В.В. Кромер развитие системы тестирования в России: Тезисы докладов Второй Всероссийской конференции (Москва, 23-24 ноября 2000г.) . - Ч.4. – М.: Прометей, 2000. – С. 59-61.
4. Седов М.С., Гуков Е.В., Киселев Е.Л., Куранова О.Н. Оценка соответствия результатов тестирования локомотивных бригад Октябрьской железной дороги – филиала ОАО РЖД назначенному критериальному баллу. Известия Петербургского университета путей сообщения, выпуск 3(36) – СПб.: ФГБОУ ВПО ПГУПС, 2013. – с. 171-175.
5. Seber G. A. F. Multivariate Observations. Aukhland: Wiley-Interscience, 2004, - 705 p.

### Рецензенты:

Боровских Ю.В., д.ф.-м.н., профессор кафедры «Математика и моделирование», ФГБОУ ВПО Петербургский государственный университет путей сообщения Императора Александра I, г.Санкт-Петербург.

Грибкова Н.В., д.ф.-м.н., профессор кафедры «Математика и моделирование», ФГБОУ ВПО Петербургский государственный университет путей сообщения Императора Александра I, г.Санкт-Петербург.