

Модели измерений уровня подготовленности студентов

В.И. Серых, Л.В. Гребцова, Е.И. Чернышевская

В статье получены и исследованы основные распределения вероятностей уровня подготовленности студентов.

Ключевые слова: уровень подготовленности, трудность теста, распределения вероятностей, измерительное уравнение, характеристическая функция теста.

1. Введение

Одним из основных показателей результативности образовательного процесса является уровень подготовленности студента по различным учебным дисциплинам (или комплексу дисциплин), входящих в программу подготовки специалиста.

В современной технологии обучения для оценки уровня подготовленности широко применяются различные методы тестирования.

При обработке результатов тестирования в основном применяют модель Г. Раша [2 – 7], связывающую вероятности правильного ответа одного испытуемого на одно тестовое задание с разностью между уровнем подготовленности и уровнем трудности тестового задания. Рассмотрим дихотомический случай, когда за правильный ответ ставится оценка 1, а за неправильный – 0. При этом вероятность правильного ответа, согласно модели Г. Раша, равна

$$P = \frac{1}{1 + \exp[-(y - \mu)]}, \quad (1)$$

где y – уровень подготовленности студента, μ – уровень трудности тестового задания, являющиеся независимыми случайными величинами, единицей измерения которых является логит, введённый Г. Рашем [2 – 7]. Как будет показано ниже, распределение (1) получено в предположении о логистическом распределении случайной разности $\Delta = y - \mu$ с нулевым средним и единичной дисперсией. Данное предположение является определённым ограничением, поскольку не позволяет учесть реальные распределения случайных величин y и μ , которые могут быть произвольными [8, 9], а, следовательно, распределение их разности будет необязательно логистическим в соответствии с правилом свёртки распределений разности двух независимых случайных величин [9].

В настоящей статье показано, что функция (1) является условной интегральной функцией распределения уровня подготовленности y при заданном уровне трудности теста μ . Это позволяет получить совместную плотность вероятности $f(y, \mu)$, которая является полным описанием ансамбля случайных величин (y, μ) [9]. Предложенный подход даёт возможность получить модели распределений вероятностей правильных ответов (функций успеха) при произвольных распределениях уровней подготовленности студента и трудности тестовых заданий. Показано, что понятие характеристической функции теста, которое

введено в работе [2], является частным случаем интегральной функции распределения вероятностей (ФРВ) уровня подготовленности испытуемого.

2. Функции распределения уровня подготовленности

Рассмотрим случай *одного* тестового задания с априорной плотностью вероятности трудности $f(\mu)$. Пусть имеется условная (апостериорная) плотность вероятности уровня подготовленности студента $f(y|\mu)$. Тогда совместная функция плотности вероятности определяется выражением

$$f(y, \mu) = f(\mu) f(y|\mu). \quad (2)$$

Уровни подготовленности и трудности теста являются случайными величинами и определены в интервале $(-\infty, \infty)$. В соответствии с формулой полной вероятности [9] безусловная плотность вероятности и интегральная функция распределения вероятностей (ФРВ) уровня подготовленности студента запишутся в общем виде:

$$f(y) = \int_{-\infty}^{\infty} f(\mu) f(y|\mu) d\mu, \quad (3)$$

$$F(y) = \int_{-\infty}^{\infty} f(\mu) F(y|\mu) d\mu. \quad (4)$$

Здесь $F(y|\mu)$ – условная интегральная ФРВ уровня подготовленности y .

Условная функция распределения $F(y|\mu)$ зависит от вида измерительной модели, связывающей измеряемую величину y с трудностью тестового задания μ . Из выражения (1) для модели Раша следует, что вероятность правильного ответа зависит от разности случайных величин

$$\Delta = y - \mu, \quad (5)$$

которая так же является случайной величиной с некоторой плотностью вероятности $f(\Delta)$.

Перепишем выражение (5) в виде:

$$y = \mu + \Delta. \quad (6)$$

Для аддитивной модели (6) условная плотность вероятности в (2) и (3) примет вид [6, 8, 9]:

$$f(y|\mu) = f_{\Delta}(y - \mu), \quad (7)$$

где функция в правой части есть плотность вероятности разности $\Delta = y - \mu$.

С учётом выражений (3) и (6) безусловная ФРВ уровня подготовленности

$$F(y) = \int_{-\infty}^{\infty} f(\mu) F_{\Delta}(y - \mu) d\mu. \quad (8)$$

Безусловная функция распределения (8) в общем виде определяет среднюю вероятность правильного ответа студента на тестовое задание с плотностью вероятности трудности $f(\mu)$. Выражения для средних вероятностей и анализ их свойств приводятся ниже.

Для дальнейшего изложения необходимо установить характер и причину случайности значений трудности задания и теста в целом.

Во-первых, при дихотомическом тестовом задании трудность одного тестового задания должна быть некоторой фиксированной величиной μ_{k_0} , однако фактически она может быть установлена лишь с некоторой случайной погрешностью ξ_k , т.е. $\mu_k = \mu_{k_0} + \xi_k$.

Во-вторых, для теста, содержащего k дихотомических заданий с произвольной трудностью в некотором конечном интервале $[\mu_1, \mu_m]$, плотность вероятности имеет дискретный характер:

$$f(\mu_m) = \sum_{k=1}^m p_k \delta(\mu - \mu_k), \quad (9)$$

где p_k – вероятность k -того задания, $\delta(\mu - \mu_k)$ – дельта-функция [9].

В-третьих, для политомического задания [7] с l уровнями (шагами) трудностей одного задания плотность вероятности будет аналогичной (9) при $k = \overline{1, l}$. Определим условные и безусловные ФРВ уровня подготовленности при логистическом и близком к нему нормальном распределении.

Для логистического распределения разности Δ в (5) плотность вероятности $f(\Delta)$ имеет вид:

$$f(\Delta) = \frac{\exp\left(-\frac{\Delta - \alpha}{\beta}\right)}{\beta \left[1 + \exp\left(-\frac{\Delta - \alpha}{\beta}\right)\right]^2}, \quad (10)$$

где α и β – параметры распределения. При этом математическое ожидание равно α , а дисперсия $D(\Delta) = \frac{\pi^2 \beta^2}{3}$. Для плотности вероятности (10) условная ФРВ в выражении (8) приобретает вид:

$$F(y - \mu) = \frac{1}{1 + \exp\left(-\frac{y - \mu - \alpha}{\beta}\right)}. \quad (11)$$

Из выражения (11) следует, что при $\alpha = 0$ получаем модель Бирнбаума [2 – 7] для условной вероятности правильного ответа:

$$F(y - \mu) = \frac{1}{1 + \exp\left(-\frac{y - \mu}{\beta}\right)}, \quad (12)$$

а при $\alpha = 0$ и $\beta = 1$ – модель Раша (1):

$$P = F(y - \mu) = \frac{1}{1 + \exp[-(y - \mu)]}. \quad (13)$$

Следовательно, ФРВ (11), (12) и (13) определяют вероятность правильного ответа на тестовое задание при логистическом распределении (10), являясь условными функциями.

Подставляя условную ФРВ (10) в (8) с учётом (4), запишем безусловную интегральную ФРВ уровня подготовленности испытуемого при априорной плотности вероятности трудности тестового задания $f(\mu)$:

$$F(y) = \int_{-\infty}^{\infty} f(\mu) \frac{1}{1 + \exp\left(-\frac{y - \mu - \alpha}{\beta}\right)} d\mu. \quad (14)$$

При нормальном распределении величины Δ в (5) условная плотность вероятности (7) имеет вид:

$$f(y | \mu) = \frac{1}{\sqrt{2\pi}\sigma_1^2} \exp\left[-\frac{(y - m - \mu)^2}{2\sigma_1^2}\right], \quad (15)$$

где m – математическое ожидание, σ_1 – среднее квадратичное отклонение (СКО) уровня подготовленности. Тогда безусловная интегральная ФРВ (8)

$$F(y) = \int_{-\infty}^{\infty} f(\mu) F_g\left(\frac{y - m - \mu}{\sigma_1}\right) d\mu, \quad (16)$$

где $F_g(x)$ – интеграл Лапласа.

Определим безусловные функции распределения уровня подготовленности (средние вероятности правильных ответов) для некоторых распределений вероятности трудностей тестового задания.

При равномерном распределении трудности одного задания $f(\mu)$ в интервале $[\mu_1, \mu_2]$ из выражения (14) для логистического распределения величины Δ в (5) получим ФРВ уровня подготовленности студента:

$$F(y) = \frac{1}{\mu_2 - \mu_1} \int_{\mu_1}^{\mu_2} \frac{1}{1 + \exp\left(\frac{y - \mu - \alpha}{\beta}\right)} d\mu = 1 - \frac{\beta}{\mu_2 - \mu_1} \ln \frac{1 + \exp\left(-\frac{y - \mu_2 - \alpha}{\beta}\right)}{1 + \exp\left(-\frac{y - \mu_1 - \alpha}{\beta}\right)}. \quad (17)$$

При $\alpha_i = 0$ (модель Бирнбаума) получим равенство

$$F(y) = 1 - \frac{\beta}{\mu_2 - \mu_1} \ln \frac{1 + \exp\left(-\frac{y - \mu_2}{\beta}\right)}{1 + \exp\left(-\frac{y - \mu_1}{\beta}\right)}. \quad (18)$$

При $\alpha = 0, \beta = 1$ (модель Раша) из (17) получим

$$F(y) = 1 - \frac{\beta}{\mu_2 - \mu_1} \ln \frac{1 + \exp\left[-(y - \mu_2)\right]}{1 + \exp\left[-(y - \mu_1)\right]}. \quad (19)$$

При нормальном распределении $f(\mu)$ трудности теста и нормальном распределении величины Δ в (5) из выражения (16) получим

$$F(y) = F_g \left(\frac{y - m - \bar{\mu}}{\sqrt{1 + h^2}} \right), \quad (20)$$

где $h = \frac{\sigma_2}{\sigma_1}$, $\bar{\mu}$, σ_2 – математическое ожидание и СКО трудности теста соответственно, σ – СКО уровня подготовленности μ .

Дадим оценку уровня подготовленности испытуемого по среднему первичному баллу. Вначале рассмотрим случай дискретных распределений трудностей тестовых заданий.

В работе [2] введено понятие *характеристической функции* теста (ХФТ), которая определяется как среднее значение (математическое ожидание) первичного балла $\bar{b}(\Theta)$, полученного студентом за выполнение теста, состоящего из m заданий (дискретное распределение трудностей тестовых заданий):

$$\bar{b}(\Theta) = \frac{1}{m} \sum_{k=1}^m F_{\Delta}(\Theta - \mu_k), \quad (21)$$

где $F_{\Delta}(\Theta - \mu_k)$ – функция Раша (1), где $y = \Theta$ – уровень подготовленности студента (в логитах [2]).

Примечание 1. Средний первичный балл $\bar{b}$ определяется как отношение числа правильно выполненных заданий к общему числу заданий в тесте [2].

Примечание 2. Приведённое в [2] определение (21) ХФТ никак не связано с классическим понятием характеристической функции случайной величины, связанной с плотностью вероятности этой величины парой преобразования Фурье [9].

Область определения функции (21) включает весь диапазон уровней подготовленности испытуемых, $\Theta \in (-\infty, \infty)$, её значениями являются соответствующие первичные баллы. В [2] утверждается, что знание ХФТ равносильно знанию точного соответствия между процентными первичными баллами и окончательными оценками уровня подготовленности испытуемых в логитах.

Покажем, что определение (21) характеристической функции теста есть не что иное, как безусловная функция распределения уровня подготовленности испытуемого. Для теста, содержащего k дихотомических заданий с произвольной трудностью в некотором конечном интервале $[\mu_1, \mu_2]$, плотность вероятности имеет дискретный характер (9) и определяется равенством (9).

Подставляя плотность вероятности (9) в выражение (8) для интегральной функции распределения уровня подготовленности и используя фильтрующее свойство δ -функции, получим равенство

$$F(y) = \int_{-\infty}^{\infty} \sum_{k=1}^m p_k \delta(\mu - \mu_k) F_{\Delta}(y - \mu) d\mu = \sum_{k=1}^m p_k F_{\Delta}(y - \mu_k). \quad (22)$$

Сделав замену в (22) $y = \Theta$, при равновероятном распределении трудностей заданий теста $p_k = \frac{1}{m}$ получим равенство

$$F(\Theta) = \frac{1}{m} \sum_{k=1}^m F_{\Delta}(\Theta - \mu_k). \quad (23)$$

Из сравнения выражений (23) и (21) следует их полное совпадение в частном случае равномерного дискретного распределения трудностей тестовых заданий при использовании модели Раша.

Таким образом, выражение (21), трактуемое в [2] как определение характеристической функции теста, есть безусловная функция распределения (23) уровня подготовленности испытуемого, что и требовалось показать.

Общим выражением для ФРВ уровня подготовленности испытуемого является выражение (22), с учётом которого средний первичный балл для дискретного распределения трудностей (9) определяется равенством

$$\bar{b}(\Theta) = \sum_{k=1}^m p_k F_{\Delta}(\Theta - \mu_k). \quad (24)$$

Теперь рассмотрим случай непрерывных распределений трудностей тестовых заданий. При равномерном распределении трудностей заданий в тесте средний первичный балл определяется для обобщённого логистического распределения (17) равенством

$$\bar{b}(\Theta) = F(\Theta) = 1 - \frac{\beta}{\mu_2 - \mu_1} \ln \frac{1 + \exp\left(-\frac{\Theta - \mu_2 - \alpha}{\beta}\right)}{1 + \exp\left(-\frac{\Theta - \mu_1 - \alpha}{\beta}\right)}. \quad (25)$$

Для частного случая модели Раша в работе [2] получено выражение

$$\bar{b}(\Theta) = \frac{1}{\mu_2 - \mu_1} \int_{\mu_1}^{\mu_2} \frac{1}{1 + e^{-(\Theta - \mu)}} d\mu = \frac{1}{\mu_2 - \mu_1} \ln \frac{e^{\Theta} + e^{\mu_1}}{e^{\Theta} + e^{\mu_2}} + 1. \quad (26)$$

Разделив числитель и знаменатель под знаком логарифма в (26) на e^{Θ} , получим равенство

$$\bar{b}(\Theta) = 1 - \frac{1}{\mu_2 - \mu_1} \ln \frac{1 + \exp[-(\Theta - \mu_2)]}{1 + \exp[-(\Theta - \mu_1)]}, \quad (27)$$

которое полностью совпадает с правой частью выражения (19) для безусловной функции распределения $F(y)$ при $y = \Theta$ при использовании модели Раша (частный случай для логистического распределения (17)). Выражения (21) и (26) справедливы лишь для частного случая модели Раша и равновероятном распределении трудностей заданий в тесте.

В итоге можно отметить, что введённое в работе [2] понятие характеристической функции теста не имеет смысла, поскольку она (ХФТ) является частным случаем интегральной функции распределения уровня подготовленности испытуемого.

Для оценки уровня подготовленности студента $\hat{\Theta}$ по среднему первичному баллу $\bar{b}(\hat{\Theta})$ в случае непрерывных распределений трудностей тестовых заданий необходимо использовать безусловные распределения уровня подготовленности (8), (17), (19) и (20).

Примечание 3. Для дискретных распределений трудностей используются те же распределения с учётом выражения (22).

В общем виде из выражения (8) при $y = \Theta$ получим равенство

$$\bar{b}(\Theta) = F(\Theta) = \int_{-\infty}^{\infty} f(\mu) F_{\Delta} d\mu. \quad (28)$$

Тогда оценка $\hat{\Theta}$ уровня подготовленности определяется из уравнения

$$F(\hat{\Theta}) = \bar{b}, \quad (29)$$

т.е. является $\bar{b}$ - квантилем безусловных ФРВ (16), (17), (20) уровней подготовленности испытуемого и определяется через функцию, обратную ФРВ, выражением

$$\hat{\Theta}_{\bar{b}} = F^{-1}(\bar{b}). \quad (30)$$

Точность равенства (30) зависит от уровня значимости критерия согласия опытного распределения баллов с теоретическим [9, 10]. Точность оценки среднего балла в (25) зависит от числа тестовых заданий.

На рис. 1 в качестве примера представлены графики зависимостей среднего первичного балла (25) от уровня подготовленности студента: $b_1(\Theta)$ – для обобщённой логистической модели (17) при $\alpha = 2, \beta = 0.5$; $b_2(\Theta)$ – для модели Бирнбаума (18) при $\alpha = 0, \beta = 3$; $b_3(\Theta)$ – для модели Раша (19) при $\alpha = 0, \beta = 1$. Распределение трудностей заданий $f(\mu)$ – равномерное в симметричном интервале $[-1, 1]$. Обобщённая логистическая модель (17) имеет три параметра α, β, μ , т.е. имеет больше степеней свободы (три) по сравнению с моделями Бирнбаума (два параметра – β, μ) и Раша (один параметр μ).

Следует отметить, что увеличение числа параметров ФРВ позволяет более точно проводить аппроксимацию (сглаживание) эмпирических распределений уровня подготовленности студентов.

График $b_1(\Theta)$ учитывает изменение среднего значения α , от которой зависит пересечение ФРВ с уровнем 0.5 и дисперсии распределения (17) β . Величина β определяет крутизну наклона кривой $b(\Theta)$. Коэффициент дискриминации равен $d = \frac{1}{\beta}$. Коэффициент дискриминации влияет на различие близких по уровню подготовленности студентов [2]. Графики $b_2(\Theta)$ и $b_3(\Theta)$ для моделей Бирнбаума (18) и Раша (19) пересекаются в точке (0, 0.5) и отличаются углом наклона ФРВ.

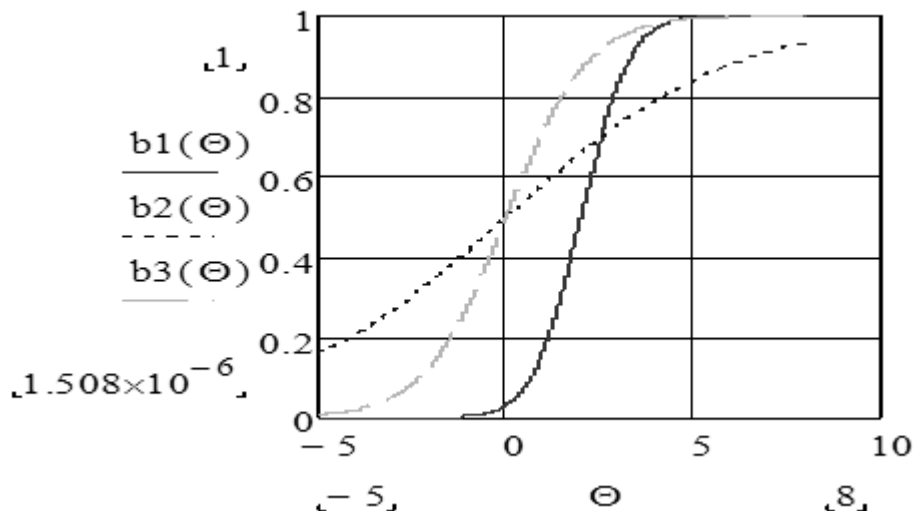


Рис. 1. Графики зависимостей среднего первичного балла (25) от уровня подготовленности студента: $b_1(\Theta)$ – для обобщённой логистической модели (17) при $\alpha = 2, \beta = 0.5$; $b_2(\Theta)$ – для модели Бирнбаума (18) при $\alpha = 0, \beta = 3$; $b_3(\Theta)$ – для модели Раша (19) при $\alpha = 0, \beta = 1$

На рис. 2 представлены графики средних значений первичного балла для обобщённой логистической модели (17) при $\alpha = 0$ для равномерного распределения трудностей тестовых

заданий в интервалах: $[3, 0]$ – кривая $b_{11}(\Theta)$; $[-1, 1]$ – кривая $b_{12}(\Theta)$; $[0, -3]$ – кривая $b_{13}(\Theta)$.

Из представленных зависимостей следует, что при симметричных относительно нуля интервалах распределения трудностей квантиль $\hat{\Theta}_{0.5} = 0$ (кривая $b_{12}(\Theta)$). При отрицательных или положительных интервалах распределения трудностей задания кривые уходят соответственно влево или вправо от нуля.

Значения 0.5-квантилей определяют точки максимальной крутизны (коэффициент дискриминации). Для любых интервалов равномерного распределения трудностей тестовых заданий $[\mu_1, \mu_2]$ и произвольных значений α для обобщённой логистической модели (19) выражение для 0.5-квантилей имеет вид

$$\hat{\Theta}_{0.5} = \frac{\mu_2 + \mu_1}{2} + \alpha. \quad (31)$$

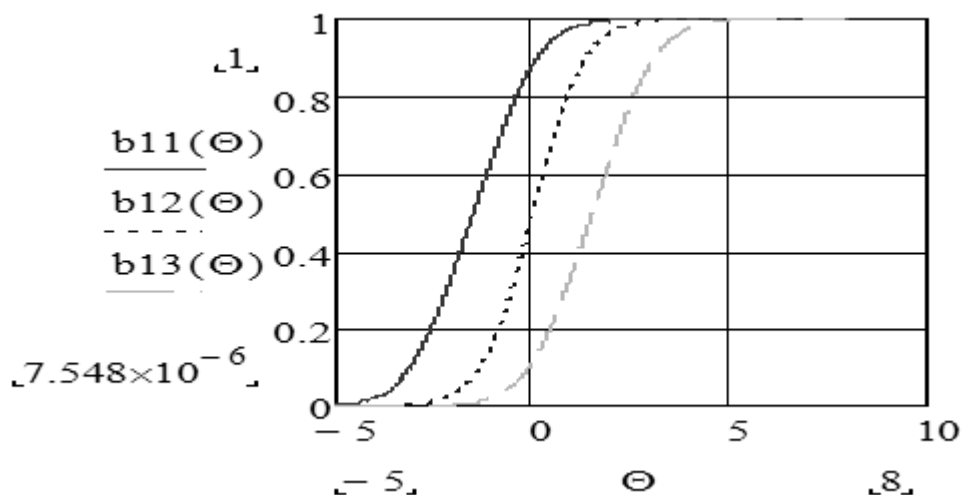


Рис.2. Графики средних значений первичного балла для обобщённой логистической модели (17) при $\alpha = 0$ для равномерного распределения трудностей тестовых заданий в интервалах: $[3, 0]$ – кривая $b_{11}(\Theta)$; $[-1, 1]$ – кривая $b_{12}(\Theta)$; $[0, -3]$ – кривая $b_{13}(\Theta)$

Выражение (31) является весьма важным для определения начала координат, а также диапазона уровней трудности тестовых заданий с учётом начального среднего уровня подготовленности в выражении (17).

Произведём сравнение дискретных и непрерывных распределений при оценках уровня подготовленности. Рассмотрим дискретное распределение (25) при использовании обобщённой логистической модели (11)

$$\bar{b}(\Theta) = \sum_{k=1}^m p_k F(\Theta - \mu_k). \quad (32)$$

Здесь

$$F(\Theta - \mu_k) = \left[1 + \exp \left(- \frac{\Theta - \alpha - (\mu_1 - k\Delta)}{\beta} \right) \right]^{-1}, \quad (33)$$

где $\Delta = \frac{\mu_2 - \mu_1}{m}$ – интервал трудностей между соседними заданиями теста.

На рис. 3 представлены графики для средних первичных баллов, построенных для непрерывного распределения (19) (кривая $b(\Theta)$) и для дискретного распределения (32) трудностей заданий в интервале $[-3, 3]$ (кривые $b_{10}(\Theta)$ и $b_5(\Theta)$ для десяти и пяти заданий в тесте соответственно).

Из рис. 3 следует, что для стандартного интервала трудностей $[-3, 3]$ для достаточно точного совпадения кривых необходимо, чтобы в тесте было не менее десяти заданий. При меньшем количестве тестовых заданий необходимо использовать дискретные распределения (24), (32), (33).

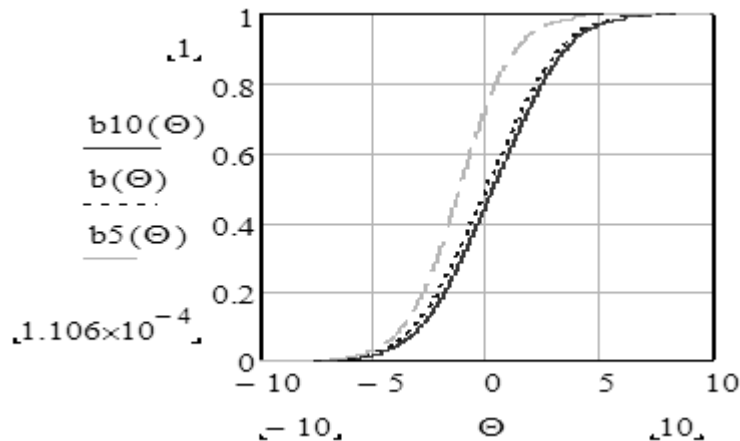


Рис. 3. Графики для средних первичных баллов, построенных для непрерывного распределения (19) (кривая $b(\Theta)$) и для дискретного распределения (32) трудностей заданий в интервале $[-3, 3]$ (кривые $b_{10}(\Theta)$ и $b_5(\Theta)$ для десяти и пяти заданий в тесте соответственно)

3. Выводы

Дано теоретическое обоснование применения условных и безусловных функций распределения для измерения уровня подготовленности студентов.

Получены выражения для оценки уровня подготовленности испытуемого по среднему первичному баллу с применением предложенных моделей для функций распределения вероятностей правильных ответов.

Литература

1. Галлагер Р. Теория информации и надёжная связь. - М.: Сов. радио, 1974. - 720 с.
2. Хлебников В.А. Характеристическая функция теста и её существенные параметры в модели Раша. Программные продукты и системы. - №4.- 2005. С. 21- 25.
3. Олейник Н.М. Тест как инструмент измерения уровня знаний и трудности заданий в современной технологии обучения. Учебное пособие по спецкурсу, - Донецк: Донецкий государственный университет, 1991. – 168 с.
4. Нейман Ю.М., Хлебников В.А. Введение в теорию моделирования и параметризации педагогических тестов.- М.: Прометей, 2000. -169 с.
5. Нейман Ю.М., Хлебников В.А. Педагогическое тестирование как измерение. – М.: Центр тестирования МО РФ, 2002. – 67 с.
6. Серых В.И. Измерения в образовательных процессах. // Актуальные проблемы электронного приборостроения: Тр. Международной науч.- техн. Конф. – Новосибирск, 2010. – Т. 3. – С. 92 -94.

7. Карданова Е.Ю., Карданов Р.С. О некоторых свойствах характеристической и информационной функций политомического тестового задания.// Вестник НовГУ. 2010.- № 55.- С.20 – 24.
8. Серых В.И., Пальчун Ю.А., Квиткова И.Г. Некоторые вопросы метрологического обеспечения продукции.// Метрология.2010 -№ 9. – с. 35 – 43.
9. Левин Б.Р. Теоретические основы статической радиотехники. - М.: Сов. радио, 1965. -Кн. I. - 752 с.
10. Р 50.1.037 – 2002. Прикладная статистика. Правила проверки согласия опытного распределения с теоретическим.

Статья поступила в редакцию 09.06.2011г.

Серых Владимир Ильич

к.т.н., доцент НГТУ, тел. 226-77-25

Гребцова Лариса Васильевна

начальник ОМК СибГУТИ, тел. 269-82-12, e-mail: smk@sibsutis.ru

Чернышевская Елена Ильинична

к.э.н., доцент СибГУТИ, тел.269-82-83, e-mail: eliche@sibsutis.ru

Models of Students' Qualification Level Measurements

V. I. Serykh, L. V. Grebtsova, E. I. Chernyshevskaya

In the article the basic distributions of probabilities of students' qualification level are obtained and investigated.

Keywords: qualification level, the difficulty of the test, distributions of probabilities, the measuring equation, characteristic function of the test.