МИНИМАКСНОЕ ОЦЕНИВАНИЕ И ПРОГНОЗИРОВАНИЕ ДЛЯ БЕТА-ИЕРАРХИЧЕСКИХ МОДЕЛЕЙ ДАННЫХ

М. А. Пашкевич

Белорусский государственный университет
г. Минск, Беларусь
pma@omegasoftware.com

Предлагаются методы робастного оценивания и прогнозирования для бета-иерархических моделей группированных бинарных данных с искаженными результатами наблюдений. Для известных интервалов вероятностей аддитивных стохастических искажений доказано, что оценки параметров моделей, оптимальные в смысле минимакса смещения, вычисляются при средних значениях уровней искажений. Доказано, что прогноз, оптимальный в смысле минимакса риска, достигается на правых границах интервалов вероятностей искажений.

Ключевые слова: группированные бинарные данные, оценивание, прогнозирование, бета-иерархические модели, искажения, робастность.

ВВЕДЕНИЕ

Для описания стохастических свойств группированных бинарных данных традиционно применяются бета-иерархические модели, которые, в отличие от классической биномиальной модели, позволяют учесть межгрупповую корреляцию и внутригрупповую зависимость результатов наблюдений [1]. В основе бета-иерархического подхода лежит бета-биномиальное распределение [2], которое позволяет получить явные выражения для Байесовской прогнозирующей функции. Это обусловливает практическую значимость рассматриваемого класса моделей в медицине, экономике, технике и в других областях [3]. При этом наиболее широко используются бета-биномиальная (ББМ) и бета-логистическая (БЛМ) модели. Для оценивания параметров ББМ и БЛМ традиционно используются метод моментов и метод максимального правдоподобия [4, 5].

Однако искажения в данных существенно влияют на точность статистических выводов, полученных на основе бета-иерархических моделей. Проведенные численные эксперименты показали, что даже при небольших уровнях искажений классические методы могут приводить к большим ошибкам в прогнозе. Поэтому актуальна задача разработки новых, устойчивых к искажениям, методов статистического оценивания и прогнозирования для бета-иерархических моделей.

В предыдущих работах [5–6] автором были получены выражения, позволяющие оценить смещение оценок параметров ББМ и БЛМ, а также вычислить увеличение риска прогнозирования при уровнях искажений, известных с точностью до значений. В данной работе предлагаются робастные методы оценивания и прогнозирования для ББМ и БЛМ для случая, когда уровни аддитивных стохастических искажений известны с точностью до интервала. Доказано, что оценки параметров моделей, оптимальные в смысле минимакса смещения, вычисляются при средних значениях уровней искажений. Доказано также, что прогноз, оптимальный в смысле минимакса риска, достигается при максимальных допустимых уровнях искажений.

ПОСТАНОВКА ЗАДАЧИ

Пусть результаты наблюдений описываются набором k бинарных векторов-строк $B = (B_1, B_2, \ldots, B_k)$, $B_i \in \{0,1\}^{n_i}$, где $B_i = (B_{i1}, B_{i2}, \ldots, B_{in_i})$ – результаты серии испытаний над i-м объектом, причем $B_{ij} = 1$, если в испытании j для объекта i случайное событие A имело место, и $B_{ij} = 0$ в противном случае. Объекту номер i в испытании номер j поставлен в соответствие некоторый m-вектор факторов $Z_i \in R^m$, который описывает свойства объекта. При этом предполагается, что размеры серий испытаний $n_1, n_2, \ldots, n_k$ малы, и объекты обладают свойством «слабой неоднородности» [1].

Для описания стохастических свойств рассматриваемых данных используется бета-иерархическая модель, основанная на следующих предположениях.

П$_1$. Для i-го объекта вероятность успеха p_i постоянна в течение всей серии испытаний.

П$_2$. Вероятность успеха p_i является случайной величиной, имеющее бета распределение с параметрами α_i^0, β_i^0, причем $p_1, p_2, \ldots, p_k$ независимы в совокупности.

П$_3$. Параметры бета распределений α_i, β_i связаны с факторами Z_i выражениями $\alpha_i = f_\alpha(Z_i)$, $\beta_i = f_\beta(Z_i)$, где $f_\alpha(.), f_\beta(.)$ – некоторые функции.

В данной работе рассматриваются две наиболее широко используемые модели из описанного класса: бета-биномиальная модель (ББМ) и бета-логистическая модель (БЛМ), которые определяются следующими дополнительными предположениями:

ББМ: $f_\alpha(Z_i) = \alpha^0$, $f_\beta(Z_i) = \beta^0$, $n_i = n$;

параметры модели: $n \in N$, $\alpha^0, \beta^0 \in R$.

БЛМ: $f_\alpha(Z_i) = \exp(Z_i^T a^0)$, $f_\beta(Z_i) = \exp(Z_i^T b^0)$;

параметры модели: $\{n_i\} \in N$, $a^0, b^0 \in R^m$.

Заметим, что размеры серий испытаний n для ББМ и $n_1, n_2, \ldots, n_k$ для БЛМ предполагаются известными.

Пусть далее данные B подвержены аддитивным стохастическим искажениям, и наблюдается искаженная матрица $\tilde{B}$:

$$\tilde{B}_{ij} = B_{ij} \oplus \eta_{ij}, \tag{1}$$

где $\{\eta_{ij}\}$ – независимые случайные бинарные величины, а $\oplus$ – операция сложения по модулю два, причем $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, n_i$. При этом для каждого i, j имеет место следующая зависимость случайной величины η_{ij} от наблюдений B_{ij}:

$$\begin{gathered} P\{\eta_{ij} = 1 \mid B_{ij} = 0\} = \varepsilon_0, \ P\{\eta_{ij} = 1 \mid B_{ij} = 1\} = \varepsilon_1, \\ \varepsilon_0 \in [\varepsilon_0^{\min}, \varepsilon_0^{\max}], \quad \varepsilon_1 \in [\varepsilon_1^{\min}, \varepsilon_1^{\max}], \end{gathered} \tag{2}$$

где $\varepsilon_0^{\min}, \varepsilon_0^{\max}, \varepsilon_1^{\min}, \varepsilon_1^{\max}$ – известные границы интервалов для уровней искажений, точные значения ε_0, ε_1 неизвестны. Задача заключается в робастном оценивании параметров ББМ и БЛМ и прогнозировании вероятностей $\{p_i\}$ на их основе в случае аддитивных стохастическим искажений (1), (2) с фиксированными уровнями ε_0, ε_1, известными с точностью до заданных интервалов $[\varepsilon_0^{\min}, \varepsilon_0^{\max}]$, $[\varepsilon_1^{\min}, \varepsilon_1^{\max}]$.

РОБАСТНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ МОДЕЛЕЙ

Рассмотрим задачу построения b-робастных оценок параметров бета-иерархической модели, обеспечивающих минимальное смещение при «наихудших» значениях уровней искажений из заданных интервалов. При этом предлагаемый подход применим сначала к оценкам параметров ББМ, построенных по методу моментов (ММ-оценок). В работе [5] автором было доказано, что для известных уровней искажений ε_0, ε_1 и при истинных значениях параметров модели α^0, β^0 справедливы следующие асимптотические разложения для смещения ММ-оценки параметров ББМ:

$$\Delta\alpha(\varepsilon_0,\varepsilon_1)=(\alpha_0+2\beta_0+1)\cdot\varepsilon_0+\alpha_0(\alpha_0+1)/\beta_0\cdot\varepsilon_1+o(\varepsilon_0,\varepsilon_1), \qquad (3)$$

$$\Delta\beta(\varepsilon_0,\varepsilon_1)=\beta_0+\beta_0(\beta_0+1)/\alpha_0\cdot\varepsilon_0+(2\alpha_0+\beta_0+1)\cdot\varepsilon_1+o(\varepsilon_0,\varepsilon_1). \qquad (4)$$

Обозначим вектор искомых параметров через $\theta=(\alpha,\beta)$, а вектор возмущений через $\varepsilon^0=(\varepsilon_0^0,\varepsilon_1^0)^T$. Тогда, используя асимптотические разложения (3), (4), смещение ММ-оценки $\hat{\theta}$ можно представить как $\Delta\hat{\theta}_E=Q(\theta^0)\cdot\varepsilon+1_m\cdot o(\|\varepsilon\|)$, где Q – (2×2)-матрица из соответствующих коэффициентов, 1_m – m-вектор, составленный из единиц, $\|.\|$ – евклидова норма. Рассмотрим далее класс оценок вида $\hat{\theta}_c(\varepsilon)=\hat{\theta}-Q(\hat{\theta})\cdot\varepsilon$, где ε – некоторое значение из заданного интервала, которые будем называть оценками с компенсированным смещением. Тогда b-робастную оценку с компенсированным смещением можно определить как $\hat{\theta}_c^b(\varepsilon^{\min},\varepsilon^{\max})=\hat{\theta}_c(\varepsilon^*)$, где ε^* – решение задачи оптимизации

$$\max_{\varepsilon^0\in\mathcal{E}}\left\|E\{\hat{\theta}^c(\varepsilon)-\theta^0\mid\varepsilon^0\}\right\|\to\min_{\varepsilon\in\mathcal{E}}, \qquad (5)$$

причем θ^0 – неизвестное истинное значение для θ, и $\mathcal{E}=[\varepsilon_0^{\min},\varepsilon_0^{\max}]\times[\varepsilon_1^{\min},\varepsilon_1^{\max}]$. Указанная оценка может быть построена с использованием следующей теоремы.

Теорема 1. Если группированные бинарные данные, описываемые бета-иерархической моделью, подвержены аддитивным стохастическим искажениям (1), (2) с уровнями, известными с точностью до интервалов $\varepsilon_r^0\in[\varepsilon_r^{min},\varepsilon_r^{max}]$, $r=0,1$, то b-робастная оценка с компенсированным смещением вычисляется при средних уровнях искажений: $\hat{\theta}_c^b(\varepsilon^{\min},\varepsilon^{\max})=\hat{\theta}_c\big((\varepsilon^{\min}+\varepsilon^{\max})/2\big)$.

Доказательство. Предположим, что матрица Q^TQ невырожденная, тогда задача оптимизации (5) сводится к эквивалентной

$$\max_{\varepsilon^0\in\mathcal{E}}\|\varepsilon-\varepsilon^0\|_V\to\min_{\varepsilon\in\mathcal{E}},$$

где $\|x\|_V=x^T\cdot V\cdot x$ – векторная норма по матрице $V=Q^TQ$. В результате необходимо найти такую точку в прямоугольнике $\mathcal{E}$, для которой максимальное расстояние до точки границ минимально. Очевидно, что это точка $(\varepsilon^{\min}+\varepsilon^{\max})/2$. Пусть далее $rank(Q^TQ)=1$, тогда одна из переменных ε_0, ε_1 исключается из рассмотрения, и задача сводится к поиску точки на отрезке, минимально удаленной от его границ. Легко показать, что случай нулевого ранга матрицы Q^TQ невозможен. Теорема доказана.

Асимптотические разложения, аналогичные выражениям (3), (4) были получены также для оценок максимального правдоподобия параметров ББМ и БЛМ [6, 7]. Это позволяет использовать результат теоремы 1 при построении на их основе b-робастных оценок с компенсированным смещением.

РОБАСТНОЕ ПРОГНОЗИРОВАНИЕ

Как было показано ранее в работе [7], оптимальный в смысле минимума среднеквадратичной ошибки Байесовский прогноз для бета-иерархической модели с искажениями (1), (2) в случае ε_0, ε_1, известных с точностью до значений, и заданных параметров модели $\{n_i,\alpha_i^0,\beta_i^0\}$ определяется как

$$\tilde{p}^i_\varepsilon(s;\varepsilon_0,\varepsilon_1)=\sum\nolimits_{l=0}^{n}\omega^i_{sl}(\varepsilon_0,\varepsilon_1)\cdot(\alpha^i_0+l)\big/(\alpha^i_0+\beta^i_0+n_i),\qquad(6)$$

где весовые коэффициенты имеют вид

$$\omega^i_{sl}(\varepsilon_0,\varepsilon_1)=C^l_{n_i}\,w^i_{sl}(\varepsilon_0,\varepsilon_1)\,\alpha_0^{i\,[l+]}\beta_0^{i\,[(n_i-l)+]}\Big/\sum\nolimits_{j=0}^{n_i}\big(C^j_{n_i}\,w^i_{sj}(\varepsilon_0,\varepsilon_1)\,\alpha_0^{i\,[j+]}\beta_0^{i\,[(n_i-j)+]}\big),$$

$$w^i_{sj}=\sum\nolimits_{l=\max(j,s)}^{\min(n_i,j+s)}C_j^{l-s}C_{n_i-j}^{l-j}\varepsilon_0^{l-j}(1-\varepsilon_0)^{n_i-l}\varepsilon_1^{l-s}(1-\varepsilon_1)^{j+s-l},$$

причем $y^{[z+]}=\prod\nolimits_{r=0}^{z-1}(y+r)$, $y\in R$, $z\in N$, а $s=\sum\nolimits_{j=1}^{n_i}B_{ij}$ – наблюдаемое число успехов для объекта i. Рассмотрим теперь задачу построения минимаксного прогноза, обеспечивающего минимальное значение среднеквадратичной ошибки при «наихудших» значениях уровней искажений из заданных интервалов $\varepsilon_0\in[\varepsilon_0^{\min},\varepsilon_0^{\max}]$, $\varepsilon_1\in[\varepsilon_1^{\min},\varepsilon_1^{\max}]$.

Решение этой задачи будем искать в классе байесовских прогнозов $\tilde{p}^i_\varepsilon(s;\varepsilon)$ вида (6), параметром которых является вектор возмущающих воздействий $\varepsilon=(\varepsilon_0,\varepsilon_1)^T$. Тогда искомое значение параметра прогноза ε^* определяется из следующей задачи оптимизации

$$\max_{\varepsilon^0\in\mathcal{E}} r^2(\tilde{p}^i_\varepsilon(s;\varepsilon)\,|\,\varepsilon^0)\;\to\;\min_{\varepsilon\in\mathcal{E}},\qquad(7)$$

где $r^2(\tilde{p}^i_\varepsilon(s;\varepsilon)\,|\,\varepsilon^0)$ – среднеквадратичная ошибка (риск) прогноза $\tilde{p}^i_\varepsilon(s;\varepsilon)$ при условии, что истинное значение уровней искажений равно ε^0, а $\mathcal{E}=[\varepsilon_0^{\min},\varepsilon_0^{\max}]\times[\varepsilon_1^{\min},\varepsilon_1^{\max}]$. Решение задачи (7) приведено в доказанной ниже теореме. Для доказательства этой теоремы полезны следующие леммы.

Л е м м а 1. Риск некоторого произвольного прогноза $\hat{p}^i(s)$ выражается через байесовский прогноз $\tilde{p}^i_\varepsilon(s;\varepsilon^0)$ и ряд распределения $\pi^{\varepsilon,i}_s(\varepsilon^0)$ как

$$r^2(\hat{p}^i)=\alpha_0^{i\,[2+]}\big/(\alpha^i_0+\beta^i_0)^{[2+]}+\sum\nolimits_{l=0}^{n_i}\big(\hat{p}^{i\,2}(s)-2\hat{p}^i(l)\cdot\tilde{p}^i_\varepsilon(l;\varepsilon^0)\big)\cdot\pi^{\varepsilon,i}_s(\varepsilon^0),$$

где $\{n_i,\alpha_i^0,\beta_i^0\}$ – истинные значения параметров модели и

$$\pi^{\varepsilon,i}_s(a,b,\varepsilon_0,\varepsilon_1)=\sum\nolimits_{j=0}^{n_i}w^i_{sj}(\varepsilon_0,\varepsilon_1)\cdot\pi^{0,i}_j,\quad \pi^{0,i}_j(a,b)=C^j_{n_i}\,\frac{B(\alpha_i^0+j,\beta_i^0+n_i-j)}{B(\alpha_i^0,\beta_i^0)}.$$

Доказательство. Воспользуемся соотношением для риска прогнозирования $E\{(p-\hat{p}^i)^2\}=E\{p^2\}-2E\{p\cdot\hat{p}^i\}+E\{(\hat{p}^i)^2\}$, первое слагаемое которого вычисляется как $E\{p^2\}=\alpha_0^{i\,[2+]}\big/(\alpha^i_0+\beta^i_0)^{[2+]}$, второе слагаемое находится по формуле полного математического ожидания, преобразование которой приводит к выражению $E\{p\cdot\hat{p}^i\}=$ $=\sum\nolimits_{l=0}^{n_i}\hat{p}^i(l)\cdot\tilde{p}^i_\varepsilon(l;\varepsilon_0,\varepsilon_1)\cdot\pi^{\varepsilon,i}_l(\varepsilon_0,\varepsilon_1)$, а третье слагаемое вычисляется по определению:

$E\{(\tilde{p}^i)^2\} = \sum_{l=0}^{n_i} \tilde{p}^i(l))^2 \pi_i^{\varepsilon,i}(\varepsilon_0, \varepsilon_1)$. В результате после подстановки получаем утверждение леммы.

Л е м м а 2. Математическое ожидание функции $f(\xi) = (\alpha + \xi - 1)^{-1}$ бета-биномиальной случайной величины ξ с параметрами n, α, β вычисляется как

$$E\{(\alpha + \xi - 1)^{-1}\} = \frac{\alpha + \beta - 1}{(\alpha - 1)(\alpha + \beta + n - 1)}.$$

Доказательство основано на свойствах бета-биномиального распределения [2].

Л е м м а 3. Для прогноза $\tilde{p}_\varepsilon^i(s;\varepsilon^*)$ справедливо следующее асимптотическое разложение риска по истинным значениям уровней искажений

$$r^2(\tilde{p}_\varepsilon^i(s;\varepsilon^*)|\varepsilon^0) = r_0^2 + \frac{n_i \beta_0^i}{(\alpha_0^i + \beta_0^i)(\alpha_0^i + \beta_0^i + n_i)^2} \cdot \varepsilon_0^0 + \frac{n_i \alpha_0^i}{(\alpha_0^i + \beta_0^i)(\alpha_0^i + \beta_0^i + n_i)^2} \cdot \varepsilon_1^0 + o(\varepsilon^0, \varepsilon^*).$$

Доказательство. Можно показать [7], что верны асимптотические разложения

$$\tilde{p}_\varepsilon^i(s \mid \varepsilon^*) = \tilde{p}_0^i(s) \cdot \left(1 + d_0^i(s)\varepsilon_0^* + d_1^i(s)\varepsilon_1^*\right) + o(\varepsilon_0^*, \varepsilon_1^*),$$

$$\tilde{p}_\varepsilon^i(s \mid \varepsilon^0) = \tilde{p}_0^i(s) \cdot \left(1 + d_0^i(s)\varepsilon_0^0 + d_1^i(s)\varepsilon_1^0\right) + o(\varepsilon_0^0, \varepsilon_1^0),$$

$$\pi_s^{\varepsilon,i} = \pi_s^{0,i} \cdot \left(1 + c_0^i(s) \cdot \varepsilon_0 + c_1^i(s) \cdot \varepsilon_1\right) + o(\varepsilon_0, \varepsilon_1),$$

где коэффициенты определяются как

$$d_0^i(s) = -s(\beta_0^i + n_i - s) \big/ \left((\alpha_0^i + s)(\alpha_0^i + s - 1)\right), \quad d_1^i(s) = (n_i - s) \big/ (\beta_0^i + n_i - s - 1).$$

$$c_0^i(s) = s(\beta_0^i + n_i - s) \big/ (\alpha_0^i + s - 1) - (n_i - s), \quad c_1^i(s) = (n_i - s)(\alpha_0^i + s) \big/ (\beta_0^i + n_i - s - 1) - s.$$

Подставим эти разложения в выражение для риска (лемма 1) и выделим коэффициенты при первых степенях $\varepsilon_0^0, \varepsilon_1^0, \varepsilon_0^*, \varepsilon_1^*$. В результате получим, что коэффициенты при $\varepsilon_0^*, \varepsilon_1^*$ равны нулю. Для коэффициента при ε_0^0 получим следующее выражение

$$\mu_0^i = \sum_{s=0}^{n_i} \left(\frac{\alpha_0^i + s}{\alpha_0^i + \beta_0^i + n_i} \right)^2 \cdot C_{n_i}^s \frac{B(\alpha_0^i + s, \beta_0^i + n_i - s)}{B(\alpha_0^i, \beta_0^i)} \cdot \left(\frac{2s(\beta_0^i + n_i - s)}{(\alpha_0^i + s)(\alpha_0^i + s - 1)} - s\frac{\beta_0^i + n_i - s}{\alpha_0^i + s - 1} + (n_i - s) \right).$$

Заметим, что данная сумма может быть рассмотрена как математическое ожидание некоторой функции бета-биномиальной случайной величины ξ с параметрами n, α, β:

$$\mu_0^i = (\alpha_0^i + \beta_0^i + n_i)^{-2} \cdot E\{h(\xi) \cdot (\alpha_0^i + \xi) \big/ (\alpha_0^i + \xi - 1)\},$$

где $h(\xi) = (\alpha_0^i + \beta_0^i + 1) \cdot \xi^2 + \left({\alpha_0^i}^2 + (-n_i + \beta_0^i - 1)\alpha_0^i - 2\beta_0^i - n_i\right) \cdot \xi + n_i \alpha_0^i (1 - \alpha_0^i)$. Пусть далее $\eta = \alpha_0 + \xi - 1$, тогда можно показать, что

$$\mu_0^i = (\alpha_0^i + \beta_0^i + n_i)^{-2} \cdot E\left\{h(\xi) + (v_2 \cdot \eta + v_1 + v_0 \eta^{-1})\right\}, \tag{8}$$

где $v_2 = (\alpha_0^i + \beta_0^i + 1)^2$, $v_1 = {\alpha_0^i}^2 + (\beta_0^i - n_i - 1)\alpha_0^i - (2\beta_0^i + n_i)$, $v_0 = {\alpha_0^i}^2 + (\beta_0^i + n_i - 2) + (1 - n_i - \beta_0^i)$.

Используя затем выражения для моментов бета-биномиального распределения и результат леммы 2, имеем

$$E\{\xi\} = \frac{n_i\alpha_0^i}{\alpha_0^i + \beta_0^i}, \quad E\{\xi^2\} = \frac{n_i\alpha_0^i}{\alpha_0^i + \beta_0^i} + \frac{n_i^{[2-]} \cdot \alpha_0^{i[2+]}}{(\alpha_0^i + \beta_0^i)^{[2+]}}, \quad E\{\eta^{-1}\} = \frac{\alpha_0^i + \beta_0^i - 1}{(\alpha_0^i - 1)(\alpha_0^i + \beta_0^i + n_i - 1)}.$$

В результате, после вычисления математического ожидания в (8) и упрощения, придем к выражению для коэффициента при ε_0, приведенному в формулировке леммы. Аналогичным образом выводится выражение для коэффициента при ε_1.

Доказанные выше леммы лежат в основе доказательства следующей теоремы.

Теорема 2. Если группированные бинарные данные, описываемые бета-иерархической моделью, подвержены аддитивным стохастическим искажениям (1), (2) с уровнями, известными с точностью до интервалов $\varepsilon_r^0 \in [\varepsilon_r^{\min}, \varepsilon_r^{\max}]$, $r = 0,1$, то при достаточно малых $\varepsilon_0^{\max}$, $\varepsilon_1^{\max}$ оптимальный в смысле минимакса риска прогноз $\tilde{p}_\varepsilon^i(s;\varepsilon^*)$ вычисляется на верхней границе уровней искажений $\varepsilon^* = \varepsilon^{\max}$.

Доказательство. Как следует из леммы 3, риск прогнозирования может быть представлен в виде $r^2(\tilde{p}_\varepsilon^i(s;\varepsilon^*) | \varepsilon^0) = r_0^2 + \mu_0^i \cdot \varepsilon_0^0 + \mu_1^i \cdot \varepsilon_1^0 + o(.)$, где $\mu_0^i, \mu_1^i > 0$. Поэтому при фиксированном значении ε^* максимум $r^2(.)$ достигается при $\varepsilon^0 = \varepsilon^{\max}$. С другой стороны, при фиксированном значении ε^0 минимум $r^2(.)$ достигается при $\varepsilon^* = \varepsilon^0$ (это следует из свойств байесовского прогноза). Поэтому решением задачи оптимизации (7) с минимаксным критерием является $\varepsilon^* = \varepsilon^{\max}$.

Полученный результат позволяет построить робастный прогноз для бета-иерархической модели, обеспечивающий минимальное значение риска при искажениях, известных с точностью до интервала.

ЗАКЛЮЧЕНИЕ

Полученные в данной работе результаты обобщают методы робастного оценивания и прогнозирования для бета-иерархических моделей группированных бинарных данных на случай вероятностей искажений, известных с точностью до интервала. При этом доказано, что оптимальные в смысле минимакса смещения оценки параметров вычисляются на середине интервалов, а прогноз, оптимальный в смысле минимакса риска, достигается на правых границах этих интервалов.

ЛИТЕРАТУРА

1. *Diggle P., Heagerty P., Liang K., Zeger S.* Analysis of Longitudinal Data. Oxford, 2002.

2. *Jonson N., Kotz S., Kemp A.* Univariate Discrete Distributions. Wiley-Interscience. New York, 1996.

3. *Wilcox R.* A review of the beta-binomial model and its extensions // Journ. of Educational Statistics. 1981. Vol. 6. P. 3–32.

4. *Tripathi R., Gupta R., Gurland J.* Estimation of parameters in the beta binomial model // Ann. Inst. Statist. Math. 1994. Vol. 46. P. 317–331.

5. *Харин Ю. С., Пашкевич М. А.* Статистическое оценивание бета-биномиального распределения при искажениях бинарных наблюдений // Весці НАН Беларусі. Сер. фіз.-мат. навук. 2003. № 1. С. 11–17.

6. *Харин Ю. С., Пашкевич М. А.* Робастность оценок максимального правдоподобия для бета-логистической модели коррелированных бинарных данных // Обозрение прикладной и промышленной математики. 2003. Т. 10. Вып. 3. С. 246–247.

7. *Pashkevich M., Dolgui A.* Robust Modeling of Consumer Behavior. Applied Optimization. Kluwer Academic Publishers. 2004. Vol. 90. P. 55–70.