

СТАТИСТИЧЕСКИЕ МЕТОДЫ И ТЕОРИЯ ВЕРОЯТНОСТЕЙ

ОПТИМАЛЬНЫЙ ДОКЛАД ПРОГНОЗА ЭКСПЕРТА*

© 1995 г. Конаков В.Д.

(Москва)

Рассматривается задача, в которой вы и группа других экспертов должны сообщить лицу, принимающему решение (ЛПР), прогнозные распределения для некоторой случайной величины X , значения которой будут наблюдаться в дальнейшем. Предполагается, что ЛПР приписывает каждому эксперту априорный вес и эти веса затем пересчитываются в апостериорные в зависимости от того, насколько хорошо было предсказано значение X . В этой ситуации вы будете стараться максимизировать свой апостериорный вес путем выбора подходящего распределения для вашего доклада. Изучаются оптимальные стратегии доклада, а также подход к получению честного доклада эксперта.

1. ВВЕДЕНИЕ

Настоящая работа продолжает исследования, проводившиеся в рамках совместного российско-канадского проекта (SBIT-project). Интерес автора к проблеме статистической обработки экспертных мнений возник в значительной степени благодаря многочисленным обсуждениям этих вопросов с проф. С. Уорнером, в частности, заметное влияние на выбор автором направлений исследований оказала работа [1], в которой путем экспертного опроса с последующей статистической обработкой доводов "за" и "против" решался вопрос: строить или нет второй международный аэропорт для Торонто до 2000 г.

2. ПОСТАНОВКА ЗАДАЧИ

Рассматривается следующая задача [2]. Субъект, называемый в дальнейшем "я", и $(m - 1)$ других субъектов образуют совет экспертов (советников). Эксперты докладывают (каждый в отдельности) свое прогнозное распределение для некоторой случайной величины X , значения которой будут наблюдаться в будущем. Предполагается, что X принимает ровно k различных значений: $x_1, \dots, x_k$, $k \geq 2$, а доклад i -го эксперта имеет вид

$$r_i(x) = \begin{cases} p_j^i, & \text{если } x = x_j, \quad j = 1, \dots, k, \\ 0, & \text{если } x \neq x_j, \quad j = 1, \dots, k. \end{cases}$$

Эти m докладов поступают к ЛПР (лицу, принимающему решение), который

* Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект 94-06-19624).

комбинирует их так, чтобы на основании этих докладов сделать выбор в некоторой конкретной задаче принятия решения. Правило комбинирования докладов включает в себя приписывание неотрицательных (априорных) весов $\alpha_1, \dots, \alpha_m$ ($\sum_{i=1}^m \alpha_i = 1$) мнениям разных экспертов и, после того как X наблюдали, эти веса пересчитываются ЛПР в зависимости от того, насколько хорошо наблюденное значение X было предсказано экспертом. Если эти пересчитанные (апостериорные) веса будут использоваться в дальнейшем и больший вес означает лучшую позицию эксперта, то естественное его желание — максимизировать свой апостериорный вес (рейтинг), оптимально подбирая свой доклад, который не всегда совпадает с истинным мнением эксперта.

Пусть $\beta_i(x)$ обозначает апостериорный вес, приписанный эксперту с номером i после того, как наблюдалось значение X , равное x . Естественная гипотеза пропорциональности

$$\beta_i(x) \propto \alpha_i r_i(x)$$

вместе с условием нормировки ($\sum_{i=1}^m \beta_i(x) = 1$) приводит к формуле

$$\beta_i(x) = \frac{\alpha_i r_i(x)}{\sum_{j=1}^m \alpha_j r_j(x)}.$$

Мы рассматриваем ситуацию, когда эксперт с номером 1 ("я") должен подобрать свой доклад оптимальным образом, т.е. так, чтобы максимизировать свой апостериорный вес. Нетрудно видеть, что данная задача сводится к случаю двух экспертов, если рассматривать остальных $(m - 1)$ экспертов как одного "обобщенного", докладывающего прогноз

$$s(x) = \sum_{j=2}^m \alpha_j r_j(x) / \sum_{j=2}^m \alpha_j.$$

Обозначая доклад, априорный и апостериорный веса 1-го эксперта через $r(x)$, α и $\beta(x)$, соответственно, имеем

$$\beta(x) = \frac{\alpha r(x)}{\alpha r(x) + (1 - \alpha) s(x)}. \quad (1)$$

Пусть S — симплекс всех возможных распределений случайной величины X , т.е. S содержит все векторы $\gamma = (\gamma_1, \dots, \gamma_k)$ такие, что $\gamma_i \geq 0$ и $\sum_{i=1}^k \gamma_i = 1$, где $\gamma_i = P(X = x_i)$.

Обозначим через $\pi = (\pi_1, \dots, \pi_k)$ честный прогноз для X 1-го эксперта (под честным прогнозом мы понимаем прогноз, отражающий истинное мнение эксперта), через $r = (r_1, \dots, r_k)$ — доложенный прогноз 1-го эксперта и через $s = (s_1, \dots, s_k)$ — прогноз другого эксперта. Все три вектора π , r и s должны принадлежать S . Если я докладываю r , а другой эксперт — s и наблюденное значение X равно x_i , то из (1) следует, что мой апостериорный вес β будет равен

$$\beta = \alpha r_i / (\alpha r_i + (1 - \alpha) s_i).$$

Мы будем предполагать, что $0 < \alpha < 1$ и что π и s лежат внутри S ; вектор r может лежать и на границе S . Возможны два принципиально разных случая: 1) к моменту, когда я должен делать свой прогноз $r(x)$, мне известен прогноз $s(x)$ другого эксперта и 2) мне не известен прогноз другого эксперта к этому моменту. Мы будем рассмат-

ривать второй случай и, следуя Баярри и Де Грооту [2], примем байесовский подход, т.е. будем считать, что имеются k условных распределений $\xi^i(s) = (\xi_1, \dots, \xi_k)$ на S , $\xi^i(s)$ – условное распределение доклада s второго эксперта, при том, что X примет значение x_i , $i = 1, \dots, k$. В такой постановке β является случайной величиной, зависящей от неизвестного мне доклада второго эксперта и от значения, которое приняла случайная величина X . Мы будем выбирать свой доклад $r = (r_1, \dots, r_k)$ таким образом, чтобы максимизировать математическое ожидание апостериорного веса E_β

$$E_\beta = \sum_{j=1}^k \pi_j \int_0^1 \alpha r_j / (\alpha r_j + (1 - \alpha) s) d\xi_j(s), \quad (2)$$

где $\xi_j(s)$ – j -е маргинальное распределение $\xi^j(s)$. Оптимальное значение $r(\alpha, \xi, \pi)$, т.е. то, которое максимизирует (2), зависит от α , от набора $\xi = (\xi_1(s), \dots, \xi_k(s))$ и от π . Как было отмечено в [2], "...явное выражение для оптимального $r(\alpha, \xi, \pi)$ не может быть дано для произвольного ξ ". Другой "натяжкой" при таком подходе является то, что я могу составить представление о функциях распределения $\xi_j(s)$. Более реалистичен, на наш взгляд, другой подход, являющийся эмпирическим аналогом подхода Баярри и Де Гроота [2]. Предположим, что мы имеем выборку из n независимых наблюдений

$$\begin{aligned} s_1(x_1), \dots, s_{n_1}(x_1), \\ s_1(x_2), \dots, s_{n_2}(x_2), \\ s_1(x_k), \dots, s_{n_k}(x_k), \end{aligned} \quad (3)$$

$\sum_{i=1}^k n_i = n$. Здесь $s_1(x_i), \dots, s_{n_i}(x_i)$, $i = 1, \dots, k$ – доложенные вероятности i -го исхода в n_i экспертных опросах при условии, что i -й исход имел место (т.е. $X = x_i$). Такую информацию можно получить, если наш "обобщенный" эксперт участвовал в нескольких задачах принятия решения подобного типа. Обозначим

$$J_i(r) = \int_0^1 \alpha \pi_i r / (\alpha r + (1 - \alpha) s) d\xi_i(s), \quad i = 1, \dots, k,$$

и пусть $\hat{J}_i(r)$ – эмпирический аналог $J_i(r)$, т.е.

$$\hat{J}_i(r) = \frac{1}{n_i} \sum_{j=1}^{n_i} \pi_i \alpha r / (\alpha r + (1 - \alpha) s_j(x_i)).$$

Очевидно, что для любого положительного κ

$$P\left(\sup_{r \in S} \left| \sum_{i=1}^k J_i(r_i) - \sum_{i=1}^k \hat{J}_i(r_i) \right| > \kappa\right) \leq \sum_{i=1}^k P\left(\sup_{r \in [0,1]} |J_i(r) - \hat{J}_i(r)| > \kappa / k\right). \quad (4)$$

Предположим, что мы получили оценку

$$P\left(\sup_{r \in S} \left| \sum_{i=1}^k J_i(r_i) - \sum_{i=1}^k \hat{J}_i(r_i) \right| > \kappa\right) \leq \eta, \quad (5)$$

где $\kappa > 0$ и $\eta > 0$ малы. Тогда если $r^* \in S$ – вектор, максимизирующий сумму $\sum_{i=1}^k J_i(r_i)$, а $r^{**} \in S$ – вектор вероятностей, максимизирующий эмпирическую сумму $\sum_{i=1}^k \hat{J}_i(r_i)$, то

с вероятностью, не меньшей, чем $1 - \eta$

$$\left| \sum_{i=1}^k J_i(r_i^*) - \sum_{i=1}^k \hat{J}_i(r_i^*) \right| \leq \kappa, \quad \left| \sum_{i=1}^k J_i(r_i^{**}) - \sum_{i=1}^k \hat{J}_i(r_i^{**}) \right| \leq \kappa.$$

Следовательно,

$$\sum_{i=1}^k J_i(r_i^*) - \sum_{i=1}^k J_i(r_i^{**}) < 2\kappa, \quad \left| \sum_{i=1}^k J_i(r_i^*) - \sum_{i=1}^k \hat{J}_i(r_i^{**}) \right| < 3\kappa.$$

Итак

$$P\left(\left| \sum_{i=1}^k J_i(r_i^*) - \sum_{i=1}^k \hat{J}_i(r_i^{**}) \right| > 3\kappa\right) \leq \eta. \quad (6)$$

Последнее неравенство означает, что если мы вместо задачи максимизации теоретической суммы будем решать задачу максимизации ее эмпирического аналога, то максимумы будут близки с большой вероятностью. В реальных задачах мы, как правило, не знаем набора функций распределения $\xi = (\xi_1(s), \dots, \xi_k(s))$. Естественно предположить, что мы имеем набор данных (3). Поэтому переход к эмпирической сумме представляется нам важным для приложений. Для максимизации эмпирической суммы можно использовать стандартные методы прикладного динамического программирования (см. [3]). Для реализации этого подхода остается указать способы получения оценок типа (1). Необходимый для этого математический аппарат описан в следующем разделе.

3. КЛАССЫ МНОЖЕСТВ С ПОЛИНОМИАЛЬНОЙ РАЗЛИЧИМОСТЬЮ

В этом разделе мы приводим необходимые нам сведения о полиномиальных классах множеств и функций. Более полное изложение содержится, например, в [4, 5].

Определение 1. Пусть B — класс подмножеств некоторого пространства S . Скажем, что этот класс имеет полиномиальную различимость степени ν , если существует полином $p(\cdot)$ (степени ν) такой, что для любого N и любого множества $S_0 \subset S$, состоящего из N точек, существует самое большее $p(N)$ различных множеств вида $S_0 \cap D$, $D \in B$.

Определение 2. Скажем, что класс множеств B исчерпывает множество точек F , если любое подмножество F (включая пустое множество и само F) может быть получено как $F \cap D$ для некоторого $D \in B$.

Следующий фундаментальный результат теории классов множеств с полиномиальной различимостью принадлежит Вапнику и Червоненкису [5]. В этой работе устанавливается, что имеется всего лишь две возможности: либо для любого целого ν найдется множество $F \subset S$, состоящее из ν точек, такое, что B исчерпывает F , либо B имеет полиномиальную различимость.

Теорема А ([4, 5]). Пусть S_0 — множество в S , состоящее из N точек. Пусть существует целое $\nu \leq N$, такое, что B не исчерпывает никакого множества из ν точек в S_0 . Тогда число различных подмножеств вида $S_0 \cap D$, $D \in B$, не превосходит $C_N^0 + \dots + C_N^{\nu-1}$.

Следствие. Если класс B не исчерпывает никакого множества из N точек, то он имеет полиномиальную различимость степени не большей $\nu - 1$.

Определение 3. Графиком вещественнозначной функции f на S назовем подмножество $G_f = \{(s, t): 0 \leq t \leq f(s) \text{ либо } f(s) \leq t \leq 0\}$ пространства $S \times R$.

Определение 4. Мы будем говорить, что класс функций F имеет полиномиальную различимость степени ν , если графики функций из F имеют полиномиальную различимость степени ν .

Сформулируем теперь необходимые нам утверждения теории равномерного уклонения эмпирических оценок от средних в общей постановке.

Пусть $Q(z, \alpha) > 0$ – параметрическое семейство положительных функций, удовлетворяющих следующим условиям: 1) при любом фиксированном значении параметра $\alpha^* \in \Lambda$ функции $Q(z, \alpha)$ измеримы по z ; 2) множество функций $Q(z, \alpha)$, $\alpha \in \Lambda$, имеет полиномиальную различимость степени ν .

Тогда справедливы следующие утверждения о скорости равномерной сходимости эмпирических средних

$$I_3(\alpha) = \frac{1}{l} \sum_{i=1}^l Q(z_i, \alpha),$$

построенных по выборке $z_1, \dots, z_l$ к их математическим ожиданиям

$$I(\alpha) = \int Q(z, \alpha) P(z) dz.$$

Теорема В ([5]). Пусть выполнены условия 1), 2) и $Q(z, \alpha) \leq r$ для всех $\alpha \in \Lambda$, z . Тогда при $l > \max(2/\kappa, \nu - 1)$

$$P(\sup_{\alpha} |(\alpha) - I_3(\alpha)| > \tau \kappa) < 9 \frac{(2l)^{\nu-1}}{(\nu-1)!} e^{-\frac{\tau^2 l}{4}}.$$

Теорема В позволяет утверждать, что максимумы по α математического ожидания $I(\alpha)$ и его эмпирического аналога $I_3(\alpha)$ близки с большой вероятностью.

В заключение приведем два вспомогательных утверждения, доказательство которых содержится в [4].

Лемма 1 ([4, с. 18]). Если ϵ и D – классы множеств, имеющие полиномиальную различимость, то каждый из классов:

- 1) $\{D^c: D \in D\}$,
- 2) $\{C \cup D: C \in \epsilon, D \in D\}$,
- 3) $\{C \cap D: C \in \epsilon, D \in D\}$

также имеет полиномиальную различимость.

Лемма 2 ([4, с. 20]). Пусть Y – конечномерное векторное пространство вещественнозначных функций на S . Тогда класс множеств вида $\{g \geq 0\}$, $g \in Y$, имеет полиномиальную различимость степени, не превосходящей размерности Y .

4. ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Рассмотрим семейство функций $Q_i(z, r)$ на отрезке $z \in [0, 1]$, зависящих от параметра $r \in [0, 1]$

$$Q_i(z, r) = \left\{ \frac{\alpha \pi_i r}{\alpha r + (1 - \alpha) z}, \quad r \in [0, 1] \right\}.$$

Очевидно $Q_i(z, r) \leq 1$ и для применения теоремы В разд. 3 достаточно проверить, что множество функций $Q_i(z, r)$ имеет полиномиальную различимость конечной степени ν .

Лемма 3. Графики функций семейства $Q_i(z, r)$ не исчерпывают никакого множества, состоящего из двух точек.

Доказательство. Рассмотрим две точки (z_1, t_1) и (z_2, t_2) в (z, t) -плоскости. Из монотонности функций из $Q_i(z, r)$ следует, что если $\min(z_1, z_2) = 0$, то графики функций

из $Q_i(z, r)$ не исчерпывают такие двухточечные множества. Пусть $\min(z_1, z_2) > 0$ и существуют две точки r^* и r^{**} такие, что $(z_1, t_1) \in G_{f_r^*}$, $(z_2, t_2) \in \bar{G}_{f_r^*}$ ($\bar{A}$ означает дополнение A) и $(z_1, t_1) \in \bar{G}_{f_r^{**}}$, $(z_2, t_2) \in G_{f_r^{**}}$. (Здесь $f_r^* = \frac{\alpha \pi_i r^*}{\alpha r^* + (1-\alpha) z}$, $f_r^{**} = \frac{\alpha \pi_i r^{**}}{\alpha r^{**} + (1-\alpha) z}$). Это означает, что $\frac{\alpha r^*}{\alpha r^* + (1-\alpha) z_1} \geq \frac{t_1}{\pi_i}$, $\frac{\alpha r^{**}}{\alpha r^{**} + (1-\alpha) z_1} < \frac{t_1}{\pi_i}$, $\frac{\alpha r^*}{\alpha r^* + (1-\alpha) z_2} < \frac{t_2}{\pi_i}$, $\frac{\alpha r^{**}}{\alpha r^{**} + (1-\alpha) z_2} \geq \frac{t_2}{\pi_i}$.

Как следствие этих неравенств получим

$$\frac{\alpha r^*}{\alpha r^* + (1-\alpha) z_2} < \frac{t_2}{\pi_i} \leq \frac{\alpha r^{**}}{\alpha r^{**} + (1-\alpha) z_2},$$

$$\frac{\alpha r^{**}}{\alpha r^{**} + (1-\alpha) z_1} < \frac{t_1}{\pi_i} \leq \frac{\alpha r^*}{\alpha r^* + (1-\alpha) z_1}.$$

Следовательно

$$\alpha(1-\alpha) z_2 (r^* - r^{**}) < 0, \quad \alpha(1-\alpha) z_1 (r^{**} - r^*) < 0. \quad (7)$$

Так как $\min(z_1, z_2) > 0$, то из первого неравенства (7) вытекает, что $r^* < r^{**}$, а из второго, что $r^{**} < r^*$. Полученное противоречие завершает доказательство леммы.

Применяя теперь теорему В разд. 3, получим, что при $n_i > 2/\kappa$, $\kappa > 0$, справедлива оценка

$$P(\sup_{r \in [0,1]} |J_i(r) - \hat{J}_i(r)| > \kappa) < 18 n_i e^{-\frac{\kappa^2 n_i}{4}}. \quad (8)$$

Далее, из (4) и (8) следует, что при $n_{\min} = \min_{i=1, \dots, k} n_i > 2/\kappa$

$$P(\sup_{r \in S} |\sum_{i=1}^k J_i(r_i) - \sum_{i=1}^k \hat{J}_i(r_i)| > \kappa) \leq 18 \sum_{i=1}^k n_i \exp\left(-\frac{\kappa^2 n_i}{4k^2}\right),$$

а из (6), что при $n_{\min} > 2/\kappa$

$$P(|\sum_{i=1}^k J_i(r_i^*) - \sum_{i=1}^k \hat{J}_i(r_i^{**})| > 3\kappa) \leq 18 \sum_{i=1}^k n_i \exp\left(-\frac{\kappa^2 n_i}{4k^2}\right). \quad (9)$$

Подытожим наши рассуждения в виде отдельного утверждения. Пусть $M = \max_{r \in S} E\beta$ – максимальное значение математического ожидания апостериорного веса эксперта (2), а $\hat{M} = \max_{r \in S} \hat{E}\beta$ – максимальное значение $\hat{E}\beta$ – эмпирического аналога $E\beta$, получающегося заменой функций распределения $\xi_i(s)$ на эмпирические функции распределения $\xi_j^{(n)}(s)$, построенные по таблице данных (3).

Теорема 1. Для любого $\kappa > 0$ при $n_{\min} > 2/\kappa$ справедливо неравенство

$$P(|M - \hat{M}| > 3\kappa) \leq 18 \sum_{i=1}^k n_i \exp\left(-\frac{\kappa^2 n_i}{4k^2}\right). \quad (10)$$

Если есть основания полагать, что $\xi_1(s) = \dots = \xi_k(s) = \xi(s)$, то тогда

$$E\beta = \int_0^1 \sum_{i=1}^k \pi_i \alpha r_i / (\alpha r_i + (1-\alpha)s) d\xi(s)$$

и в этом случае можно получить лучшую оценку, чем (10). Рассмотрим семейство функций

$$Q(z, r) = \sum_{i=1}^k \pi_i \alpha r_i / (\alpha r_i + (1-\alpha)z), \quad z \in [0, 1], \quad r \in S.$$

Если $f_r \in Q(z, r)$, то

$$G_{f_r} = \left\{ (z, t): 0 \leq t \leq \sum_{i=1}^k \pi_i \alpha r_i / (\alpha r_i + (1-\alpha)z) \right\} =$$

$$= \{z: A_1 z^k + A_2 z^{k-1} + \dots + A_k z + A_{k+1} \geq 0\},$$

где коэффициенты $A_i, i = 1, \dots, k+1$, зависят от t, α, π и r (и эта зависимость может быть легко выписана в явном виде). В этих формулах для коэффициентов A_i , нужно помнить, что $t \geq 0, a, r \in S$. Из леммы 2 разд. 3 следует, что семейство $Q(z, r)$ имеет полиномиальную различимость конечной степени h , не превосходящей $(k+1)$. Аналогично (10) получим при $h > \max\left(\frac{2}{k}, h-1\right)$

$$P(|M - \hat{M}| > 3\kappa) \leq 9 \frac{(2n)^{h-1}}{(h-1)!} \exp\left(-\frac{\kappa^2 n}{4}\right), \quad (11)$$

$n = n_1 + n_2 + \dots + n_k$. Неравенства (10) и (11) позволяют рассматривать вместо исходной задачи на максимум задачу максимизации эмпирической суммы. Программа максимизации эмпирической суммы была написана В. Кушнером (ИПУ РАН) и основана на методах динамического программирования [3]. Она использовалась при проведении модельных экспериментов. В этих экспериментах предполагалось, что $\xi_1(s) = \xi_2(s) = \dots = \xi_k(s) = \xi(s)$ и

$$\xi(s) = \begin{cases} s^{2p+1} & \text{для } p = 0, 1, 2, \dots, \\ s^p & \text{для } p = 1/2, 1/3, 1/4, \dots \end{cases}$$

Параметр p отражал степень опытности эксперта-оппонента; $p = 0$ соответствует равномерному распределению, при возрастании $p, p = 1, 2, \dots$, функция распределения $\xi(s)$ "приближается" к единичному скачку в точке 1, что соответствует идеальному эксперту-прогнозисту, убывание $p, p = 1/2, 1/3, \dots$ — уменьшению компетентности эксперта. Далее в расчетах моделировалась таблица данных (3) и проводилась максимизация эмпирической суммы.

5. О ПРОБЛЕМЕ ПОЛУЧЕНИЯ ЧЕСТНОГО ПРОГНОЗА

До сих пор мы рассматривали проблему нахождения вектора $r(\alpha, \xi, \pi)$, максимизирующего $E\beta$ или $\hat{E}\beta$. При этом не обязательно $r(\alpha, \xi, \pi) = \pi$, т.е. доложенный прогноз совпадал с истинным мнением эксперта π . Некоторые частные случаи, когда такое совпадение происходит, рассмотрены в [2]. Там же предложен и способ получения от эксперта его честного прогноза. Предлагалось максимизировать не $E\beta$, а $E\phi(\beta)$, где функция $\phi(\beta)$ подобрана так, что $E\phi(\beta)$ имеет единственный максимум при $r = \pi$. Такие функции ϕ авторы называют строго правильными (точное определение в

[2, с. 221]), ими же получено полное описание всех строго правильных функций, введение которых позволяло, с одной стороны, избежать катастрофических ситуаций вроде падения рейтинга до нуля и, с другой стороны, побуждало эксперта давать честный прогноз. Возможен и другой подход – введение специальной системы выплат (наград) за то, насколько хорошо предсказано значение X . Рассмотрим простейший пример. Пусть случайная величина X принимает значения 0 и 1, честный прогноз эксперта – (π_1, π_2) , а доложенный прогноз – (Q_1, Q_2) . После того, как реализовалась случайная величина X , ЛПР платит: $1 - Q_2^2$, если $X = 1$ и $1 - Q_1^2$, если $X = 0$. Эксперт максимизирует среднюю зарплату

$$\max_{0 \leq Q \leq 1} \{ \pi_1(1 - Q_2^2) + \pi_2(1 - Q_1^2) \}.$$

Дифференцируя по Q_1 , легко получим, что максимум достигается при $Q_1 = \pi_1$, $Q_2 = \pi_2$, т.е. выгоднее докладывать свое честное мнение. Этот пример нетрудно обобщить на случай, когда X принимает любое конечное число значений. Рассмотрим для простоты случай четырех значений: 1, 2, 3, 4. Пусть честный прогноз $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$, а доложенный прогноз – $Q = (Q_1, Q_2, Q_3, Q_4)$. Назначим следующую систему выплат

$$1 - (Q_3 + Q_4)^2 + \left(1 - \left(\frac{Q_2}{Q_1 + Q_2} \right)^2 \right), \quad \text{если } X = 1,$$

$$1 - (Q_3 + Q_4)^2 + \left(1 - \left(\frac{Q_1}{Q_1 + Q_2} \right)^2 \right), \quad \text{если } X = 2,$$

$$1 - (Q_1 + Q_2)^2 + \left(1 - \left(\frac{Q_4}{Q_3 + Q_4} \right)^2 \right), \quad \text{если } X = 3,$$

$$1 - (Q_1 + Q_2)^2 + \left(1 - \left(\frac{Q_3}{Q_3 + Q_4} \right)^2 \right), \quad \text{если } X = 4.$$

Нетрудно показать, что максимум средней зарплаты эксперта снова достигается, когда $Q = \pi$. Случай трех состояний сводится к случаю четырех состояний, если положить $\pi_4 = Q_4 = 0$ и т.д.

В заключение укажем на некоторые обобщения и открытые проблемы. Пусть $\varphi(Q)$ – убывающая функция, $\varphi(0) = 1$, $\varphi(1) = 0$. Для простоты ограничимся случаем двух состояний. Рассмотрим систему выплат: $\varphi(1 - Q_1)$, если $X = 1$, $\varphi(1 - Q_2)$, если $X = 0$ (в приведенном выше примере $\varphi(Q) = 1 - Q^2$). Назовем функцию $\varphi(\cdot)$ допустимой, если максимум средней зарплаты эксперта достигается при $\pi = Q$. С точки зрения ЛПР желательно добиться честного прогноза при наименьших издержках, т.е. использовать такую $\varphi(\cdot)$, чтобы величина $E_\pi(\varphi) = \pi_1\varphi(\pi_2) + \pi_2\varphi(\pi_1)$ была минимальной. Скажем, что φ_1 лучше φ_2 (с точки зрения ЛПР), если $\forall (\pi_1, \pi_2) E_\pi(\varphi_1) \leq E_\pi(\varphi_2)$ и хотя бы для одного (π_1, π_2) имеет место строгое неравенство.

Пример. Пусть $\varphi_1(x) = 1 - x^2$, $x \in [0, 1]$,

$$\varphi_2(x) = \begin{cases} 1 - \frac{8}{3}x^3, & \text{если } x \in [0, \frac{1}{2}], \\ -4x^2 + \frac{8}{3}x^3 + \frac{4}{3}, & \text{если } x \in [\frac{1}{2}, 1]. \end{cases}$$

Простой подсчет показывает, что $E_{\pi}(\varphi_1) = 1 - \pi_1 + \pi_1^2$, $\pi_1 \in [0, 1]$, а

$$E_{\pi}(\varphi_2) = \begin{cases} 1 - \pi_1 + \frac{4}{3} \pi_1^3, & \text{если } \pi_1 \in [0, 1/2], \\ \frac{4}{3} - 3\pi_1 + 4\pi_1^2 - \frac{4}{3} \pi_1^3, & \text{если } \pi_1 \in [1/2, 1]. \end{cases}$$

Отсюда сразу следует, что φ_2 лучше φ_1 , ибо $E_{\pi}(\varphi_2) < E_{\pi}(\varphi_1)$ всегда, кроме случаев $\pi_1 = 0$ или $\pi_1 = 1$ и задача сводится к тому, чтобы найти наилучшую $\varphi(\cdot)$ в классе допустимых функций.

Автор посвящает настоящую работу памяти проф. Стэнли Уорнера, безвременно ушедшего от нас летом 1992 г.

ЛИТЕРАТУРА

1. Warner S. Balanced Information: The Pickering Airport Experiment // The Review of Economics & Statistics. 1981. V. LXIII. № 2.
2. Bayarri M.J., De Groot M.H. Optimal Reporting of Predictions // JASA. 1989. V. 84. № 405.
3. Bellman R., Dreyfus S. Applied Dynamic programming. Princeton: Princeton Univ. Press, 1962.
4. Pollard D. Convergence of Stochastic Processes. Springer Series in Statistics. 1984. Springer-Verlag.
5. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. М.: Наука, 1979.

Поступила в редакцию
2 IX 1994