Лекция 20

ИСПОЛЬЗОВАНИЕ МЕТОДОВ СТАТИСТИЧЕСКОЙ ТЕОРИИ ПРИНЯТИЯ РЕШЕНИЙ В ТЕХНИЧЕСКОЙ ДИАГНОСТИКЕ

При использовании методов статистической теории принятия решений правило постановки диагноза определяют из условия минимизации риска (см. ниже) или вероятностей ошибочных решений при диагностировании. Выбор того, какие из указанных величин минимизируются, зависит от наличия априорной информации для постановки диагноза.

Рассмотрим случай дихотомии. Будем считать, что решение о состоянии диагностируемого объекта принимается на основе анализа диагностических параметров в n-мерном диагностическом пространстве, состояние объекта в котором определяется вектором диагностических параметров $\mathbf{x} = \left(x_1, x_2, ..., x_n\right)$.

Правило принятия решения о состоянии диагностируемого объекта сводится к определению в диагностическом пространстве *граничной поверхности*, разделяющей его на две области S_0 и S_1, такие, что если вектор $\mathbf{x}$, полученный при обследовании объекта, принадлежит S_0 $\left(\mathbf{x} \in S_0\right)$, то ставится диагноз D_0. Если $\mathbf{x} \in S_1$, принимают, что объект находится в состоянии D_1.

Разбиение диагностического пространства на области, соответствующие возможным состояниям диагностируемого объекта, и определение положения граничной поверхности проводится на основе априорных данных о распределении диагностических параметров при различных состояниях объекта и потерях от ошибочно принятых решений при постанове диагноза.

Обозначим через $w\left(\mathbf{x}/D_i\right) = w\left(x_1, x_2, ..., x_n/D_i\right)$, $\left(i = 0, 1\right)$ условные плотности вероятности распределения вектора параметров для состояний D_0 и D_1. Вероятности правильно поставленных диагнозов при разбиении пространства на области S_0 и S_1 определяются соотношениями

$$P_0 = P\left(D_0\right) P\left(\mathbf{x} \in S_0/D_0\right) = P\left(D_0\right) \int\limits_{S_0} w\left(\mathbf{x}/D_0\right) d\mathbf{x};$$

$$P_1 = P(D_1)P(\mathbf{x} \in S_1/D_1) = P(D_1)\int_{S_1} w(\mathbf{x}/D_1)d\mathbf{x}, \qquad \text{(Л20.1)}$$

где $P(\mathbf{x} \in S_i/D_i)$, $(i = 0, 1)$ — условные вероятности принадлежности вектора $\mathbf{x}$ области S_i, когда объект находится в состоянии D_i; $P(D_i)$ — априорная вероятность диагноза D_i; $d\mathbf{x} = dx_1\ dx_2 \ldots dx_n$. Вероятности ошибок 1-го и 2-го рода вычисляются по формулам:

$$P_F = P(D_0)P(\mathbf{x} \in S_1/D_0) = P(D_0)\int_{S_1} w(\mathbf{x}/D_0)d\mathbf{x};$$

$$P_M = P(D_1)P(\mathbf{x} \in S_0/D_1) = P(D_1)\int_{S_0} w(\mathbf{x}/D_1)d\mathbf{x}, \qquad \text{(Л20.2)}$$

где $P(\mathbf{x} \in S_i/D_j)$, $(i, j = 0, 1;\ i \neq j)$ — условные вероятности принадлежности вектора $\mathbf{x}$ области S_i, когда объект находится в состоянии D_j.

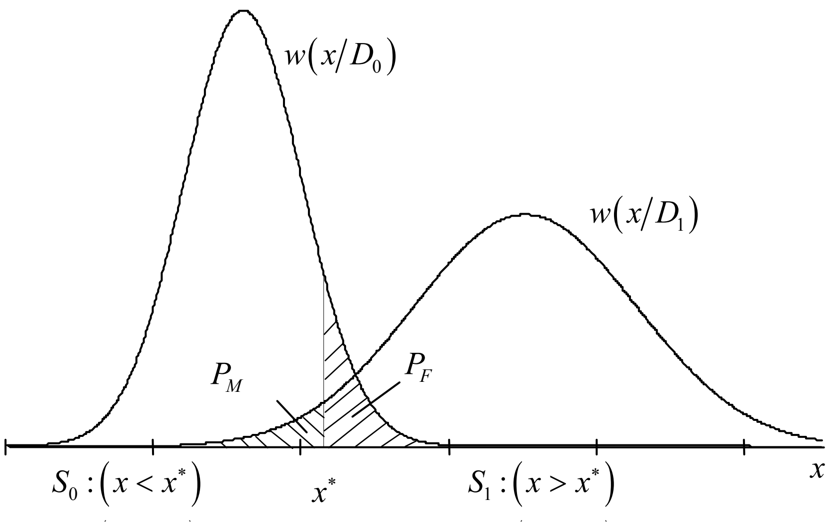

Рис. Л20.1. Условные плотности вероятности распределения параметра x в состояниях D_0 и D_1; x^* — граничное значение параметра; площади заштрихованных областей пропорциональны ошибкам 1-го и 2-го рода

На рис. Л20.1 показаны плотности вероятности распределения одного диагностического параметра x для двух различных состояний объекта. Область S_0, включающая значения $x \le x^*$, отвечает диагнозу D_0, область S_1 $(x > x^*)$ — диагнозу D_1;

x^* — значение диагностического параметра, разделяющего области S_0 и S_1. Площади различным образом заштрихованных областей пропорциональны ошибкам 1-го и 2-го рода.

1. Метод минимального риска

Если известны априорные вероятности $P(D_0)$ и $P(D_1)$ и стоимости потерь от неправильно принятых решений о состоянии диагностируемого объекта и выигрыши при правильной постановке диагноза — Π_{ij} $(i,\ j=0,1)$ (табл. Л19.3), можно вычислить средний риск или математическое ожидание потерь при диагностике. Он складывается из потерь от ошибочных решений и выигрышей от правильных диагнозов:

$$R = \Pi_{01}P_{\mathrm{F}} + \Pi_{10}P_{\mathrm{M}} + \Pi_{00}P_0 + \Pi_{11}P_1, \qquad \text{(Л20.3)}$$

или с учетом (Л20.1) и (Л20.2)

$$\begin{aligned} R = {} & \Pi_{01}P(D_0)\int\limits_{S_1} w(\mathbf{x}/D_0)\,d\mathbf{x} + \Pi_{10}P(D_1)\int\limits_{S_0} w(\mathbf{x}/D_1)\,d\mathbf{x} + \\ & + \Pi_{00}P(D_0)\int\limits_{S_0} w(\mathbf{x}/D_0)\,d\mathbf{x} + \Pi_{11}P(D_1)\int\limits_{S_1} w(\mathbf{x}/D_1)\,d\mathbf{x}. \end{aligned} \qquad \text{(Л20.4)}$$

Разбиение диагностического пространства на области S_0 и S_1 проводится из условия, чтобы средний риск R был минимален. С учетом очевидных соотношений, следующих из условия нормировки

$$\int\limits_{S_1} w(\mathbf{x}/D_0)\,d\mathbf{x} = 1 - \int\limits_{S_0} w(\mathbf{x}/D_0)\,d\mathbf{x};$$

$$\int\limits_{S_1} w(\mathbf{x}/D_1)\,d\mathbf{x} = 1 - \int\limits_{S_0} w(\mathbf{x}/D_1)\,d\mathbf{x},$$

средний риск (Л20.4) можно представить в виде

$$\begin{aligned} R = {} & \Pi_{01}P(D_0) + \Pi_{11}P(D_1) - \int\limits_{S_0} \big\{ [\Pi_{01} - \Pi_{00}] P(D_0) w(\mathbf{x}/D_0) - \\ & - [\Pi_{10} - \Pi_{11}] P(D_1) w(\mathbf{x}/D_1) \big\}\, d\mathbf{x}. \end{aligned}$$

Чтобы риск был минимален, необходимо чтобы значение интеграла в этом выражении было максимально. Для этого необходимо, чтобы S_0 — область диагноза D_0 — включала все значения (точки) $\mathbf{x}$, для которых подынтегральное выражение было положительным, то есть значения $\mathbf{x}$, для которых выполняется равенство

$$(\Pi_{01} - \Pi_{00}) P(D_0) w(\mathbf{x}/D_0) - (\Pi_{10} - \Pi_{11}) P(D_1) w(\mathbf{x}/D_1) \geq 0.$$

Отсюда область S_0 содержит все значения $\mathbf{x}$, для которых отношение правдоподобия $\Lambda(\mathbf{x})$ превышает пороговое значение Λ_0:

$$\Lambda(\mathbf{x}) = \frac{w(\mathbf{x}/D_0)}{w(\mathbf{x}/D_1)} \geq \frac{\Pi_{10} - \Pi_{11}}{\Pi_{01} - \Pi_{00}} \cdot \frac{P(D_1)}{P(D_0)} = \Lambda_0 . \qquad \text{(Л20.5)}$$

Уравнение

$$\Lambda(\mathbf{x}) = \Lambda_0 \qquad \text{(Л20.6)}$$

определяет граничную поверхность между областями S_0 и S_1 диагностического пространства. Правило постановки диагноза можно записать в виде сложного неравенства:

$$\Lambda(\mathbf{x}) \underset{D_1}{\overset{D_0}{\gtrless}} \Lambda_0 , \qquad \text{(Л20.7)}$$

то есть, если выполняется верхнее неравенство, принимается решение, что объект находится в состоянии D_0, если нижнее — в состоянии D_1. Ошибки 1-го и 2-го рода при использовании этого правила вычисляются по формулам (Л20.2) с учетом разбиения диагностического пространства на области S_0 и S_1.

Вероятность ошибок диагностирования можно вычислить также следующим образом. Будем рассматривать отношение правдоподобия $\Lambda(\mathbf{x})$ в левой части неравенства (Л20.5) как случайную функцию случайного вектора диагностических параметров $\mathbf{x}$.

Так как плотности вероятности распределения $\mathbf{x}$ для состояний D_0 и D_1 различны, различаются и плотности вероятности распределения отношения правдоподобия $\Lambda(\mathbf{x})$ в этих состояниях, то есть можно говорить об условных плотностях вероятности распределения отношения правдоподобия — $w\big(\Lambda(\mathbf{x})/D_0\big)$ и $w\big(\Lambda(\mathbf{x})/D_1\big)$. Введение этих функций позволяет перейти от многомерных плотностей вероятности распределения диагностических параметров к одномерным.

Действительно, из правила постановки диагноза (Л20.5) и (Л20.7) следует, что событие, состоящее в том, что выполняется неравенство $\Lambda(\mathbf{x}) < \Lambda_0$ при условии, что диагностируемый объект находится в состоянии D_0, и событие — $\mathbf{x} \in S_1$, если объект находится в состоянии D_0, эквивалентны. Следовательно, вероятности этих событий равны:

$$P\big(\{\Lambda(\mathbf{x}) < \Lambda_0\}/D_0\big) = P(\mathbf{x} \in S_1/D_0) .$$

Можно также утверждать, например, что событие, состоящее в том, что выполняется неравенство $\Lambda(\mathbf{x}) > \Lambda_0$ при условии, что диагностируемый объект находится в состоянии D_1, и событие — $\mathbf{x} \in S_0$ при нахождении объекта в состоянии D_1, также эквивалентны, поэтому

$$P\left(\left\{\Lambda(\mathbf{x}) > \Lambda_0\right\}/D_1\right) = P\left(\mathbf{x} \in S_0/D_1\right).$$

Нетрудно составить таблицу следующих эквивалентных событий:

Таблица Л20.1. **Эквивалентные события и равные вероятности**

Состояние объекта	Эквивалентные события	Равные вероятности
D_0	$(\mathbf{x} \in S_0)$ и $\{\Lambda(\mathbf{x}) > \Lambda_0\}$	$P(\mathbf{x} \in S_0/D_0) = P\left(\{\Lambda(\mathbf{x}) > \Lambda_0\}/D_0\right) = \int\limits_{\Lambda_0}^{+\infty} w(\Lambda(\mathbf{x})/D_0)\,d\Lambda$
D_1	$(\mathbf{x} \in S_1)$ и $\{\Lambda(\mathbf{x}) < \Lambda_0\}$	$P(\mathbf{x} \in S_1/D_1) = P\left(\{\Lambda(\mathbf{x}) < \Lambda_0\}/D_1\right) = \int\limits_{-\infty}^{\Lambda_0} w(\Lambda(\mathbf{x})/D_1)\,d\Lambda$
D_0	$(\mathbf{x} \in S_1)$ и $\{\Lambda(\mathbf{x}) < \Lambda_0\}$	$P(\mathbf{x} \in S_1/D_0) = P\left(\{\Lambda(\mathbf{x}) < \Lambda_0\}/D_0\right) = \int\limits_{-\infty}^{\Lambda_0} w(\Lambda(\mathbf{x})/D_0)\,d\Lambda$
D_1	$(\mathbf{x} \in S_0)$ и $\{\Lambda(\mathbf{x}) > \Lambda_0\}$	$P(\mathbf{x} \in S_0/D_1) = P\left(\{\Lambda(\mathbf{x}) > \Lambda_0\}/D_1\right) = \int\limits_{\Lambda_0}^{+\infty} w(\Lambda(\mathbf{x})/D_1)\,d\Lambda$

Поэтому, вероятности правильно и ошибочно поставленных диагнозов после подстановки соответствующих выражений в (Л20.1) и (Л20.2) можно представить в виде:

$$P_0 = P(D_0)P(\mathbf{x} \in S_0/D_0) = P(D_0)P\left(\{\Lambda(\mathbf{x}) > \Lambda_0\}/D_0\right) = P(D_0)\int\limits_{\Lambda_0}^{+\infty} w(\Lambda(\mathbf{x})/D_0)\,d\Lambda;$$

$$P_1 = P(D_1)P(\mathbf{x} \in S_1/D_1) = P(D_1)P\left(\{\Lambda(\mathbf{x}) < \Lambda_0\}/D_1\right) = P(D_1)\int\limits_{-\infty}^{\Lambda_0} w(\Lambda(\mathbf{x})/D_1)\,d\Lambda;$$

(Л20.8)

$$P_F = P(D_0)P(\mathbf{x} \in S_1/D_0) = P(D_0)P\left(\{\Lambda(\mathbf{x}) < \Lambda_0\}/D_0\right) = P(D_0)\int\limits_{-\infty}^{\Lambda_0} w(\Lambda(\mathbf{x})/D_0)\,d\Lambda;$$

$$P_M = P(D_1)P(\mathbf{x} \in S_0/D_1) = P(D_1)P\left(\{\Lambda(\mathbf{x}) > \Lambda_0\}/D_1\right) = P(D_1)\int\limits_{\Lambda_0}^{+\infty} w(\Lambda(\mathbf{x})/D_1)\,d\Lambda.$$

(Л20.9)

Подставив пары соотношений (Л20.8) и (Л20.9) в (Л20.3), получим выражение для среднего риска

$$R = \Pi_{01} P\left(D_0\right) \int_{-\infty}^{\Lambda_0} w\left(\Lambda(\mathbf{x}) / D_0\right) d\Lambda + \Pi_{10} P(D_1) \int_{\Lambda_0}^{+\infty} w\left(\Lambda(\mathbf{x}) / D_1\right) d\Lambda +$$

$$+ \Pi_{00} P\left(D_0\right) \int_{\Lambda_0}^{+\infty} w\left(\Lambda(\mathbf{x}) / D_0\right) d\Lambda + \Pi_{11} P(D_1) \int_{-\infty}^{\Lambda_0} w\left(\Lambda(\mathbf{x}) / D_1\right) d\Lambda.$$

Таким образом, используя в качестве диагностического параметра отношение правдоподобия $\Lambda\left(\mathbf{x}\right)$, то есть при переходе от многопараметровой к однопараметровой диагностике, области S_0, и соответственно диагнозу D_0, в исходном диагностическом пространстве параметров $\mathbf{x} = \left(x_1, x_2, ..., x_n\right)$ в пространстве параметра $\Lambda\left(\mathbf{x}\right)$ соответствует диапазон $\left(\Lambda_0; +\infty\right)$, а области S_1 и диагнозу D_1 — диапазон $\left(-\infty; \Lambda_0\right)$.

Методы расчета условных плотностей вероятности отношения правдоподобия $w\left(\Lambda\left(\mathbf{x}\right)/D_0\right)$ и $w\left(\Lambda\left(\mathbf{x}\right)/D_1\right)$ для некоторых основных видов распределений $w\left(\mathbf{x}/D_i\right)$, $\left(i = 0,1\right)$ изложены в [3].

2. Метод минимального числа ошибочных решений (метод Зигерта–Котельникова, критерий идеального наблюдателя)

Этот метод применяется, если неизвестны стоимости потерь и выигрышей при постановке диагноза. Правило постановки диагноза находится из условия минимума доли ошибочных решений. Вероятность таких решений определяется соотношением

$$P_{\text{ош}} = P\left(D_0\right) \int_{S_1} w\left(\mathbf{x}/D_0\right) d\mathbf{x} + P\left(D_1\right) \int_{S_0} w\left(\mathbf{x}/D_1\right) d\mathbf{x}.$$

Можно показать [5], что она минимальна, если область S_0 диагноза D_0 содержит значения $\mathbf{x}$, для которых

$$\Lambda\left(\mathbf{x}\right) > \frac{P\left(D_1\right)}{P\left(D_0\right)} = \Lambda_0 . \qquad \text{(Л20.10)}$$

При указанном пороговом значении Λ_0 соотношение (Л20.6) определяет граничную поверхность в диагностическом пространстве, а (Л20.7) — правило постановки диагноза. Таким образом, метод минимального числа ошибочных решений является частным случаем метода минимального риска при условии

$$\Pi_{10} - \Pi_{11} = \Pi_{01} - \Pi_{00} .$$

Кроме того, он совпадает с методом Байеса (см. (Л18.11) лекцию 18). Действительно, из (Л20.10) следует

$$\frac{P(D_0)}{P(D_1)}\Lambda(\mathbf{x})=\frac{P(D_0)\,w(\mathbf{x}/D_0)}{P(D_1)\,w(\mathbf{x}/D_1)}=\frac{P(D_0/\mathbf{x})}{P(D_1/\mathbf{x})}>1$$

или

$$P(D_0/\mathbf{x})>P(D_1/\mathbf{x}),$$

что является условием постановки диагноза D_0 по методу Байеса.

Недостаток метода минимального числа ошибочных решений — игнорирование различия последствий ошибок диагностирования. Поскольку потери от пропуска сигнала, как правило, превышают потери от ложной тревоги, применение критерия идеального наблюдателя может привести к экономически не обоснованным решениям. Поэтому этот метод рекомендуется применять, если известно, что потери от ошибочных решений примерно одинаковы.

3. Метод максимального правдоподобия

Этот метод применяется, если неизвестны стоимости потерь и выигрышей при постановке диагноза, а также априорные вероятности состояния диагностируемого объекта. Согласно этому методу в область S_0 диагноза D_0 включаются значения **x**, для которых

$$P(\mathbf{x}/D_0)>P(\mathbf{x}/D_1),$$

то есть те $\mathbf{x}$, для которых априорная плотность вероятности диагноза D_0 превышает априорную плотность вероятности диагноза D_1. Метод максимального правдоподобия по сути является частным случаем метода минимального риска при $\Lambda_0 = 1$. При этом условии (Л20.6) определяет граничную поверхность, а (Л20.7) — правило постановки диагноза.