III ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК 1
LIETUVOS MATEMATIKOS RINKINYS

1963

ПРИМЕНЕНИЕ ПОНЯТИЯ ЭНТРОПИИ В ПРОБЛЕМАХ ОБНАРУЖЕНИЯ СИГНАЛА НА ФОНЕ ШУМА*

ДОБРУШИН Р. Л., ПИНСКЕР М. С., ШИРЯЕВ А. Н.

1. Как хорошо известно (см. [1]), задача обнаружения сигнала на фоне шума является частным случаем следующей более общей задачи: на отрезке времени $[0, T]$, называемом временем наблюдения, задана траектория случайного процесса $\xi(t)$ для $0 \leqslant t \leqslant T$ (дальнейшие рассмотрения одинаково применимы как в том случае, когда t принимает лишь целые значения (процесс с дискретным временем), так и в том случае, когда t принимает все действительные значения (процесс с непрерывным временем)). Априори имеются две конкурирующие гипотезы о статистической природе процесса $\xi(t)$: гипотеза I состоит в том, что наблюдаемый процесс $\xi(t)$ характеризуется системой плотностей $p_{\mathrm{I}}\left(x(t_1), \ldots, x(t_n)\right)$ для совместных распределений $\xi(t_1), \ldots, \xi(t_n)$, а по гипотезе II процесс характеризуется системой плотностей $p_{\mathrm{II}}\left(x(t_2), \ldots, x(t_n)\right)$. (Если процессы принимают конечное число значений, то во всех дальнейших построениях надо заменить плотности на вероятности, а интегралы на суммы.)

Как известно, наилучший метод различения гипотез основан на отношении правдоподобия

$$l_T\left(x(t)\right) = \log \frac{p_{\mathrm{I}}\left(x(1), \ldots, x(T)\right)}{p_{\mathrm{II}}\left(x(1), \ldots, x(T)\right)}, \tag{1}$$

в случае дискретного времени и

$$l_T\left(x(t)\right) = \lim_{n\to\infty} \log \frac{p_{\mathrm{I}}\left(x(t_1), \ldots, x(t_n)\right)}{p_{\mathrm{II}}\left(x(t_1), \ldots, x(t_n)\right)}, \tag{2}$$

где предел берется по последовательности умельчающихся разбиений $0 \leqslant t_1 \leqslant \ldots \leqslant t_n = T$, $\max\limits_i (t_{i+1} - t_i) \to 0$, в случае непрерывного времени, и

---

* Эта статья написана в 1959 году и предназначалась для опубликования в техническом журнале, что и продиктовало стиль ее изложения. Затем, учтя справедливые замечания рецензентов о недостаточной практической обоснованности рекомендаций, авторы решили доработать статью, уточнив выводы и доказательства. Теперь ясно, что, ввиду изменения научных интересов авторов, этим намерениям уже не суждено осуществиться. Так как материал статьи можно рассматривать как постановку новых математических задач (некоторые из них исследовались в дипломной работе студента МГУ Чжоу-Ай-Жуна), то авторы решились, по просьбе редакции журнала, опубликовать статью в первоначальном виде.

состоит в том, что при некотором l_0 считается, что верна гипотеза 1, если

$$l_T\left(\xi(t)\right) \geqslant l_0,$$

и гипотеза II, если это не так. Степень различимости гипотез I и II при использовании оптимального метода различения описывается двумя вероятностями: вероятностью F решить, что верна гипотеза 1, когда на самом деле имеет место II и вероятностью G решить, что верна гипотеза II, когда на самом деле имеет место гипотеза I (в задачах радиолокации гипотеза I состоит в том, что наблюдаемый процесс является суммой полезного сигнала и шума, и гипотеза II состоит в том, что наблюдаемый процесс является чистым шумом, вероятность F называют вероятностью ложной тревоги, а G – вероятностью пропуска сигнала). При заданном T между вероятностями F и G имеется, при использовании оптимального метода приема, жесткая функциональная связь, которую можно записать в виде

$$F = F(G,\ T) \tag{3}$$

или, что эквивалентно

$$G = G(F,\ T). \tag{4}$$

Иногда оказывается удобным разрешить уравнения (3) и (4) относительно T, положив

$$T = T(F,\ G). \tag{5}$$

Более подробно в радиолокационных терминах, можно сказать, что $F(G, T)$ – наименьшее значение вероятности ложной тревоги, которую можно достичь при заданных вероятности пропуска G и времени наблюдения T, что $G(F, T)$ наименьшая вероятность пропуска, которую можно достичь при заданных вероятностях ложной тревоги и времени наблюдения и что $T = T(F, G)$ – наименьшее время наблюдения, за которое можно различить гипотезы I и II с заданными вероятностями ложной тревоги F и пропуска G. Зная явное или приближенное выражение для одной из функций (3), (4), (5), конечно, легко найти остальные.

Задача явного вычисления указанных функций является трудной даже в простейших примерах. Ввиду этого кажутся важными попытки дать приближенные асимптотические формулы для функций (3), (4), (5). Одна из целей этой работы – указать такие формулы, верные при некоторых дополнительных условиях (в случае большого времени наблюдения T). Существенное преимущество приводимых формул состоит в том, что они дают также единую характеристику степени различимости гипотез, являющуюся числом, а не функцией, что, как видно из дальнейшего, позволяет сравнивать возможности обнаружения в разных условиях.

Назовем энтропией $\mathbf{H}_{\mathrm{I/II}}$ гипотезы I относительно гипотезы II за время T

$$\mathbf{H}_{\mathrm{I/II}}(T) = \int_{-\infty}^{\infty} \cdots \int \log \frac{p_{\mathrm{I}}\left(x(1),\ \ldots,\ x(T)\right)}{p_{\mathrm{II}}\left(x(1),\ \ldots,\ x(T)\right)} \cdot p_{\mathrm{I}}\left(x(1),\ \ldots,\ x(T)\right) dx(1) \ldots dx(T), \tag{6}$$

если время дискретно, и

$$H_{\mathrm{I}/\mathrm{II}}(T) = \lim_{n\to\infty} \int_{-\infty}^{\infty} \cdots \int \log \frac{p_{\mathrm{I}}\left(x(t_1), \ldots, x(t_n)\right)}{p_{\mathrm{II}}\left(x(t_1), \ldots, x(t_n)\right)} \times$$
$$\times p_{\mathrm{I}}\left(x(t_1), \ldots, x(t_n)\right) dx(t_1) \ldots dx(t_n), \tag{7}$$

если время непрерывно, где предел берется в том же смысле, что и 2. Средней энтропией гипотезы I относительно гипотезы II назовем

$$\overline{\mathbf{H}}_{\mathrm{I}/\mathrm{II}} = \lim_{T\to\infty} \frac{1}{T} \mathbf{H}_{\mathrm{I}/\mathrm{II}}(T). \tag{8}$$

Предел (8) обычно существует, если процесс $\xi(t)$ является при каждой из гипотез стационарным или если его статистические закономерности меняются периодическим образом, а также и в других рассматриваемых ниже случаях. Более того, сходимость в (8) является обычно довольно быстрой, а именно в широком классе случаев разность $\mathbf{H}_{\mathrm{I}/\mathrm{II}}(T) - \overline{\mathbf{H}}_{\mathrm{I}/\mathrm{II}} \cdot T$ является равномерно ограниченной по T. Тогда можно утверждать, что если время наблюдения T велико $\left(\text{точнее } \sqrt{\frac{T}{\Delta\tau}} \gg 1, \text{ где } \Delta\tau \text{ — типичное время корреляции процесса*}\right)$, а вероятность ошибки F не слишком мала $\left(\text{точнее } -\log F \ll \frac{T}{\Delta\tau}\right)$, то

$$G(F, T) \approx e^{-\overline{\mathbf{H}}_{\mathrm{I}/\mathrm{II}} \cdot T} \tag{9}$$

$\left(\text{точнее } \lim_{T\to\infty} \frac{\log G(F, T)}{T} = -\overline{\mathbf{H}}_{\mathrm{I}/\mathrm{II}}\right)$. Более точные формулировки условий верности соотношения (9) и идея его вывода дается в приложении 1, но во всяком случае эти условия достаточно широки, чтобы равенство (9) было применимо во всех реальных ситуациях.

Аналогичным образом определяется энтропия $\mathbf{H}_{\mathrm{II}/\mathrm{I}}$, и тогда $\left(\text{при } \sqrt{\frac{T}{\Delta\tau}} \gg 1, -\log G \ll \frac{T}{\Delta\tau}\right)$

$$F(G, T) \approx e^{-\overline{\mathbf{H}}_{\mathrm{II}/\mathrm{I}} \cdot T}. \tag{10}$$

Практическая ценность равенств (9) и (10) снижается тем, что они применимы лишь, если одна из вероятностей ошибок много больше другой, что реально лишь в тех ситуациях, когда один из видов ошибок является много более опасным.

Однако можно указать другую асимптотическую формулу, применяемую в тех случаях, когда гипотезы I и II близки друг к другу, т. е. в радиолокационных терминах, когда рассматривается задача обнаружения слабого сигнала (именно в этом случае особенно актуально накопление за большое время наблюдения T). Положим для этого

$$\mathbf{H}(T) = \mathbf{H}_{\mathrm{I}/\mathrm{II}}(T) + \mathbf{H}_{\mathrm{II}/\mathrm{I}}(T), \quad \overline{\mathbf{H}} = \overline{\mathbf{H}}_{\mathrm{I}/\mathrm{II}} + \overline{\mathbf{H}}_{\mathrm{II}/\mathrm{I}} = \lim_{T\to\infty} \frac{1}{T} H(T).$$

---

* Под временем корреляции $\Delta\tau$ (см. [1] § 2) мы подразумеваем, как это обычно делается в технической литературе, отрезок времени столь большой, что можно пренебречь зависимостью величин $\xi(t)$ и $\xi(s)$ при $t-s > \Delta\tau$. Точное математическое определение числа $\Delta\tau$ дать нельзя, однако в большинстве прикладных примеров вопрос о значении $\Delta\tau$, с точностью до порядка, сомнений не вызывает.

Величина $H(T)$ была введена в математическую статистику Каллбеком (см. [2]) и названа им информационным расстоянием, $\overline{\mathbf{H}}$ мы будем называть средним информационным расстоянием.

Предположим теперь, что гипотеза I состоит в том, что процесс $\xi_{\mathrm{I}}(t)$ („шум") является гауссовским процессом с

$$\mathbf{M}\xi_{\mathrm{I}}(t)=0.$$

Конкурирующая гипотеза II состоит в том, что

$$\xi_{\mathrm{II}}(t)=s(t)+\xi_{\mathrm{I}}(t)$$

(детерминированный сигнал $s(t)$ на фоне гауссовского шума).

В рассматриваемом случае логарифм отношения правдоподобия $l_T\big(\xi(t)\big)$ будет линейным функционалом от $\xi(t)$ и, следовательно, при обеих гипотезах $l_T\big(\xi(t)\big)$ будет иметь нормальное распределение. Это замечание позволяет показать, что минимально необходимое время $T=T(F, G)$ для различения гипотез с заданными ошибками F и G является наименьшим корнем уравнения

$$\mathbf{H}(T)=(\alpha_F+\alpha_G)^2, \tag{12}$$

где обозначено

$$\frac{1}{\sqrt{2\pi}}\int\limits_{\alpha_Q}^{\infty} e^{-x^2/2}\,dx=Q.$$

Поэтому можно ожидать при весьма общих условиях, обеспечивающих асимптотическую (при $T\to\infty$) нормальность $l_T\big(\xi(t)\big)$, получить для $T(F, G)$ формулу, аналогичную (12). Действительно, имеет место следующая асимптотическая формула:

$$T(F,\ G)\approx\frac{(\alpha_F-\alpha_G)^2}{\overline{\mathbf{H}}}, \tag{13}$$

каковы бы ни были вероятности ошибок F и G (мы исключаем из рассмотрения неинтересный случай, когда как F, так G близки к $1/2$).

Эта формула применима, если информационное расстояние $\mathbf{H}(\Delta\tau)\approx\overline{\mathbf{H}}\Delta\tau$ мало (точнее $\sqrt{\mathbf{H}(\Delta\tau)}\ll 1$). Из предположения $\sqrt{\mathbf{H}(\Delta\tau)}\ll 1$ и (13) следует, что $\frac{T}{\Delta\tau}\gg 1$, т. е. что время наблюдения велико. Более точные формулировки утверждения (13) и идея его доказательства дается в приложении 2.

Подчеркнем, однако, что хотя соотношения (1) и (13) показывают, что энтропия является весьма важной характеристикой в задаче различения двух гипотез при большом времени наблюдения, всё же нельзя преувеличивать универсальность этой характеристики. Так, например, если предположить, что вероятности ошибок F и G имеют один и тот же порядок, а $H(\Delta\tau)$ имеет порядок 1 (т. е. гипотезы не близки), то хотя $T/\Delta\tau$ будет большим, но здесь не применимы ни формула (9), ни формула (13). Более того, для случая, когда при обеих гипотезах $\xi(t_i)$ образуют последовательность независимых одинаково распределенных величин, в указанной обстановке исследованы асимптотические соотношения между F, G и T (см., например, [3]), причем результат нельзя выразить через энтропию и ответ выглядит сжолнее.

2. Укажем теперь частные случаи, где могут быть даны более простые выражения для энтропий и средних энтропий.

Пример 1. Пусть при каждой из гипотез рассматриваемый процесс представляет собой последовательность независимых одинаково распределенных величин с плотностями $p_{\mathrm{I}}(x)$ и $p_{\mathrm{II}}(x)$ соответственно. Тогда

$$\mathbf{H}_{\mathrm{I/II}}(T) = T \int_{-\infty}^{\infty} \log \frac{p_{\mathrm{I}}(x)}{p_{\mathrm{II}}(x)} p_{\mathrm{I}}(x)\, dx;$$

$$H(T) = T \int_{-\infty}^{\infty} \log \frac{p_{\mathrm{I}}(x)}{p_{\mathrm{II}}(x)} \left[p_{\mathrm{I}}(x) - p_{\mathrm{II}}(x) \right] dx. \tag{14}$$

Для этого случая асимптотические равенства (9) и (13) получены соответственно Мурье [4] и Айвазяном [5]*.

Пример 2. Пусть при каждой из гипотез рассматриваемый процесс является стационарным марковским процессом с дискретным временем, заданным плотностью $p_{\mathrm{I}}(x)$ одномерного распределения и условной плотностью вероятностей перехода $p_{\mathrm{I}}(y \mid x)$ (аналогично $p_{\mathrm{II}}(x)$ и $p_{\mathrm{II}}(y \mid x)$). Тогда

$$\begin{aligned} \overline{\mathbf{H}}_{\mathrm{I/II}} &= \int_{-\infty}^{\infty} \cdots \int \log \frac{p_{\mathrm{I}}(y \mid x)}{p_{\mathrm{II}}(y \mid x)} p_{\mathrm{I}}(y \mid x) p_{\mathrm{I}}(x)\, dx\, dy, \\ \overline{\mathbf{H}} &= \int_{-\infty}^{\infty} \cdots \int \log \frac{p_{\mathrm{I}}(y \mid x)}{p_{\mathrm{II}}(y \mid x)} \left[p_{\mathrm{I}}(y \mid x) p_{\mathrm{I}}(x) - p_{\mathrm{II}}(y \mid x) p_{\mathrm{II}}(x) \right] dx\, dy. \end{aligned} \tag{15}$$

Пример 3. (Детерминированный сигнал на фоне гауссовского шума.) Пусть при гипотезе I рассматриваемый процесс $\xi_{\mathrm{I}}(t)$ является гауссовским стационарным процессом с непрерывным временем и нулевым средним значением, задаваемый своей спектральной плотностью $f(\lambda)$, а при гипотезе II процесс $\xi_{\mathrm{II}}(t)$ является суммой гауссовского стационарного процесса со средним нуль и спектральной плотностью $f(\lambda)$ и детерминированной функцией

$$s(t) = s_0 + \sum_{i=1}^{\infty} a_i \cos(\lambda_i t + \varphi_i), \tag{16}$$

где λ_i и φ_i — некоторые константы. Здесь

$$\overline{\mathbf{H}} = 2\overline{\mathbf{H}}_{\mathrm{I/II}} = 2\overline{\mathbf{H}}_{\mathrm{II/I}} = \frac{s_0^2}{2\pi f(0)} + \frac{1}{4\pi} \sum_{i=1}^{\infty} \frac{a_i^2}{f(\lambda_i)}. \tag{17}$$

В случае дискретного времени формулы (17) будут также верны, если только $-\pi < \lambda_i < \pi$, $i = 1, \ldots, \infty$.

Вывод этой и нижеследующих формул почти очевиден, когда время дискретно и $f(\lambda) = \mathrm{const}$**.

В самом деле, в этом случае $\xi(t)$ образуют при обеих гипотезах последовательность независимых случайных величин, и вычисление можно провес-

* По поводу формулы (13) см. также работу [14], результаты которой используются в [5].

** Строго говоря, для верности этих формул нужно еще предположить, что процесс является „достаточно хорошим": например, все эти формулы верны, если $f(\lambda)$ рациональная функция.

ти по формуле (14). В общем случае для получения (17) достаточно полосу частот разбить на большое число узких полосок, считая $f(\lambda)$ в каждой полосе постоянной; полная энтропия будет равна сумме энтропий по всем полосам, поскольку в гауссовском случае процессы, соответствующие разным полосам, независимы.

Пример 4. Обобщая пример 3, предположим, что при гипотезе I (аналогично II) рассматриваемый процесс является суммой гауссовского стационарного процесса со спектральной плотностью $f_{\mathrm{I}}(\lambda)$ $\left(\text{соответственно } f_{\mathrm{II}}(\lambda)\right)$ и средним нуль и детерминированной функции

$$s_{\mathrm{I}}(t)=s_0^{\mathrm{I}}+\sum_{i=1}^{\infty} a_i^{\mathrm{I}} \cos(\lambda_i t+\varphi_i), \tag{18}$$

и соответственно

$$s_{\mathrm{II}}(t)=s_0^{\mathrm{II}}+\sum_{i=1}^{\infty} a_i^{\mathrm{II}} \cos(\lambda_i t+\varphi_i). \tag{18}$$

Тогда

$$\overline{\mathbf{H}}_{\mathrm{I/II}}=\frac{1}{2\pi}\int_0^{\infty}\left(\frac{f_{\mathrm{I}}(\lambda)}{f_{\mathrm{II}}(\lambda)}-1-\log\frac{f_{\mathrm{I}}(\lambda)}{f_{\mathrm{II}}(\lambda)}\right)d\lambda+\frac{(s_0^{\mathrm{I}}-s_0^{\mathrm{II}})^2}{2\pi f_{\mathrm{II}}(0)}+\frac{1}{4\pi}\sum_{i=1}^{\infty}\frac{(a_i^{\mathrm{I}}-a_i^{\mathrm{II}})^2}{f_{\mathrm{II}}(\lambda_i)},$$

$$\overline{\mathbf{H}}=\frac{1}{2\pi}\int_0^{\infty}\frac{f_{\mathrm{I}}(\lambda)-f_{\mathrm{II}}(\lambda)}{f_{\mathrm{I}}(\lambda)f_{\mathrm{II}}(\lambda)}d\lambda+\frac{1}{2\pi}\frac{(s_0^{\mathrm{I}}-s_0^{\mathrm{II}})^2\left(f_{\mathrm{I}}(0)+f_{\mathrm{II}}(0)\right)}{f_{\mathrm{I}}(0)f_{\mathrm{II}}(0)}+ \tag{19}$$

$$+\frac{1}{4\pi}\sum_{i=1}^{\infty}\frac{(a_i^{\mathrm{I}}-a_i^{\mathrm{II}})^2\left(f_{\mathrm{I}}(\lambda_i)+f_{\mathrm{II}}(\lambda_i)\right)}{f_{\mathrm{I}}(\lambda_i)f_{\mathrm{II}}(\lambda_i)}.$$

Вывод этой формулы проводится совсем аналогично выводу формулы (17).

Формулы (19) останутся верными, если предположение о том, что φ_i являются фиксированными числами заменить на предположение о том, что все или часть этих величин являются случайными величинами (т. е. если предположить, что фазы случайны*). Переход к случаю дискретного времени делается заменой предела интегрирования $+\infty$ на π.

3. Рассмотрим теперь задачу об оптимальном выборе формы и статистики зондирующего сигнала в проблемах обнаружения объекта и радиолокации**. Наиболее общая математическая постановка этой задачи состоит в следующем. Имеются две гипотезы I и II о состоянии канала связи, по которому распространяется сигнал (например, гипотеза I состоит в том, что отражающий объект имеется, а гипотеза II в том, что он отсутствует). Канал связи понимается здесь в смысле Шеннона, т. е. предполагается, что каждой функции $s(t)$ (зондирующему сигналу) соответствует некоторый случайный процесс, задающий сигнал на входе приёмника. Таким образом, канал, соответствующий гипотезе I, задаётся системой условных плотностей $p_{\mathrm{I}}\left(x(t_1), \ldots, x(t_n) \mid s(t)\right)$ для значения сигнала на выходе при условии,

* Реальный смысл предположения о том, что φ_i являются случайными величинами, состоит в том, что получаемые выражения можно рассматривать как приближенный для случая, когда $\varphi_i(t)$ представляет собой медленно меняющийся случайный процесс.

** Мы не рассматриваем здесь более сложную задачу оценки параметров этого объекта.

что на вход подан сигнал $s(t)$. Аналогичным образом канал, соответствующий гипотезе II, задаётся системой условных плотностей $p_{II}\left(x(t_1), \ldots, x(t_n) \mid s(t)\right)$. Если задать распределение вероятностей $\tilde{P}(\cdot)$ в пространстве всевозможных траекторий сигнала на входе $s(t)$ (т. е. предположить, что $s(t)$ – фиксированный случайный процесс), то формулы

$$p_I\left(x(t_1), \ldots, x(t_n)\right) = \int p_I\left(x(t_1), \ldots, x(t_n) \mid s(t)\right) \cdot \tilde{P}\left(ds(t)\right),$$

$$p_{II}\left(x(t_1), \ldots, x(t_n)\right) = \int p_{II}\left(x(t_1), \ldots, x(t_n) \mid s(t)\right) \cdot \tilde{P}\left(ds(t)\right)$$

зададут плотности распределения для процесса на выходе, когда на вход подается сигнал $s(t)$. Таким образом, при заданном $\tilde{P}(\cdot)$ возникает рассмотренная выше задача различения двух гипотез о процессах.

При заданных вероятностях ошибок F и G естественно считать оптимальным то распределение зондирующего сигнала $s(t)$, для которого $T(F, G)$ является минимальным, так что обнаружение с заданными вероятностями ошибок может быть проведено за минимальное время.

К сожалению, в такой постановке оптимальное распределение $\tilde{P}$ зависит от F и G, а найти его явно трудно. Однако полученные выше результаты (см. (13)) показывают, что если интересоваться случаем трудно различимых близких гипотез, то естественно считать оптимальным то распределение $\tilde{P}$, для которого среднее информационное расстояние $\overline{\mathbf{H}}(\tilde{P})$, вычисляемое по формуле (11), для процессов с плотностями (20) является максимальным. Среди всех таких процессов это дает нам критерий оптимальности, не зависящий от F и G*.

Мы будем называть

$$\overline{\overline{\mathbf{H}}} = \max_{\tilde{P}} \overline{\mathbf{H}}(\tilde{P}) \tag{21}$$

различающей способностью каналов I и II (термин введен по аналогии с пропускной способностью по Шеннону (6)). Распределение $\tilde{P}_0$ такое, что

$$\overline{\overline{\mathbf{H}}} = \overline{\mathbf{H}}(\tilde{P}_0) \tag{22}$$

мы будем называть оптимальным распределением входного сигнала. Подчеркнем, что это оптимальное распределение $\tilde{P}_0$ определяется далеко не единственным образом, и в рассматриваемых ниже примерах мы указываем лишь один из вариантов этого оптимального распределения.

При заданном $\tilde{P}_0$ по формуле (22) подсчитывается различающая способность $\overline{\overline{\mathbf{H}}}$, а затем соотношение

$$T(F, G) \approx \frac{(\alpha_F + \alpha_G)^2}{\overline{\overline{\mathbf{H}}}} \tag{23}$$

позволяет вычислить наименьшее время наблюдения, за которое можно различать гипотезы I и II с канала с заданными вероятностями ошибок.

Отметим следующий важный факт: нетрудно доказать, что всегда можно выбрать оптимальное распределение входного сигнала так, чтобы оно было

* Конечно, аналогичная теория может быть развита и на основе соотношения (9), применимого при иных условиях. Тогда надо было бы считать оптимальным то распределение $\tilde{P}$, для которого $\mathbf{H}_{I/II}(\tilde{P})$ максимально.

сосредоточено на одной фиксированной функции, т. е. оптимальным оказывается посылать не случайный, а детерминированный сигнал.

Отметим также, что в случае детерминированного сигнала на фоне гауссовского „белого" шума, т. е. гауссовского процесса со спектральной плотностью $f(\lambda)=\sigma^2$ для всех λ,

$$\mathbf{H}(T)=\frac{\int\limits_0^T s^2(t)\,dt}{\sigma^2}.$$

Отсюда непосредственно следует хорошо известный факт, что в рассматриваемом случае при заданной энергии выбор формы сигнала роли не играет.

Рассмотрим некоторые примеры.

Пример 5. Предположим, что рассматриваемый канал является при любой гипотезе однородным каналом с дискретным временем без памяти, т. е. заданы условные плотности $p_{\mathrm{I}}(x\,|\,s)$ и $p_{\mathrm{II}}(x\,|\,s)$ для сигнала на выходе x при заданном сигнале s на входе в тот же момент времени, так, что

$$p_{\mathrm{I}}\Big(x(t_1),\ \ldots\ x(t_n)\,|\,s(t)\Big)=p_{\mathrm{I}}\Big(x(t_1)\,|\,s(t_1)\Big)\ \ldots\ p_{\mathrm{I}}\Big(x(t_n)\,|\,s(t_n)\Big),$$

$$p_{\mathrm{II}}\Big(x(t_1),\ \ldots\ x(t_n)\,|\,s(t)\Big)=p_{\mathrm{II}}\Big(x(t_1)\,|\,s(t_1)\Big)\ \ldots\ p_{\mathrm{II}}\Big(x(t_n)\,|\,s(t_n)\Big),$$

тогда различающая способность

$$\overline{\mathbf{H}}=\max_s\int\limits_{-\infty}^{\infty}\log\frac{p_{\mathrm{I}}(x\,|\,s)}{p_{\mathrm{II}}(x\,|\,s)}\Big[p_{\mathrm{I}}(x\,|\,s)-p_{\mathrm{II}}(x\,|\,s)\Big]dx, \tag{24}$$

а оптимальным будет сигнал $s(t)=s_0$, где s_0 — таково, что при $s=s_0$ выражение, стоящее в (24) под знаком максимума, принимает максимальное значение.

Пример 6. Предположим, что рассматриваемые каналы являются гауссовскими каналами с непрерывным временем и аддитивным шумом, действующими в ограниченной полосе частот $[\lambda_0,\ \lambda_0+w]$. Точнее, заданы спектральные плотности шума $f_{\mathrm{I}}(\lambda)$ и $f_{\mathrm{II}}(\lambda)$. Кроме того, задан линейный фильтр A_{I}, воздействующий на сигнал $s(t)$. Этот фильтр задается при помощи своей спектральной характеристики $g_{\mathrm{I}}(\lambda)$, которая обращается в нуль вне отрезка $\lambda_0\leqslant\lambda\leqslant\lambda_0+w$. Аналогичным образом задан линейный фильтр A_{II} со спектральной характеристикой $g_{\mathrm{II}}(\lambda)$. Средняя мощность $\lim\limits_{T\to\infty}\frac{1}{T}\int\limits_0^T s^2(t)\,dt$ сигнала $s(t)$ не должна превосходить заданной константы p_s^2. Тогда по определению при заданном $s(t)$ распределения $p_{\mathrm{I}}\Big(x(t_1),\ \ldots\ x(t_n)\,|\,s(t)\Big)$ $\Big($аналогично $p_{\mathrm{II}}\Big(x(t_1),\ \ldots\ x(t_n)\,|\,s(t)\Big)\Big)$ должны совпадать с распределением случайного процесса $\xi_{\mathrm{I}}(t)+A_{\mathrm{I}}s(t)$ $\Big($аналогично $\xi_{\mathrm{II}}(t)+A_{\mathrm{II}}s(t)\Big)$, где $\xi_{\mathrm{I}}(t)$ — гауссовский стационарный процесс со спектральной плотностью $f_{\mathrm{I}}(\lambda)$ $\Big($аналогично

$\xi_{\mathrm{II}}(t)-cf_{\mathrm{II}}(\lambda)\Big)$. В этом примере оптимальным будет сигнал $s(t)=$ $=\sqrt{2}\,p_s\,\cos(\bar{\lambda}t+\varphi)$*, где $\bar{\lambda}$ выбирается так, чтобы при $\lambda=\bar{\lambda}$ выражение

$$\frac{p_s^2\Big(g_{\mathrm{I}}(\lambda)-g_{\mathrm{II}}(\lambda)\Big)^2\Big(f_{\mathrm{I}}(\lambda)+f_{\mathrm{II}}(\lambda)\Big)}{f_{\mathrm{I}}(\lambda)f_{\mathrm{II}}(\lambda)} \tag{25}$$

принимало максимальное значение. Различающая способность равна при этом

$$\bar{\mathbf{H}}=\frac{1}{2\pi}\int\limits_{\lambda_0}^{\lambda_0+w}\frac{\Big(f_{\mathrm{I}}(\lambda)-f_{\mathrm{II}}(\lambda)\Big)^2}{f_{\mathrm{I}}(\lambda)f_{\mathrm{II}}(\lambda)}d\lambda+\frac{p_s^2\Big(g_{\mathrm{I}}(\bar{\lambda})-g_{\mathrm{II}}(\bar{\lambda})\Big)^2\Big(f_{\mathrm{I}}(\bar{\lambda})+f_{\mathrm{II}}(\bar{\lambda})\Big)}{4\pi f_{\mathrm{I}}(\bar{\lambda})f_{\mathrm{II}}(\bar{\lambda})}. \tag{26}$$

Пример 7. Особенно естественным с радиолокационной точки зрения является тот частный случай, когда $f(\lambda)=f_{\mathrm{I}}(\lambda)=f_{\mathrm{II}}(\lambda)$ (спектр аддитивного шума не зависит от гипотезы) и $g_{\mathrm{I}}(\lambda)\equiv 0$, $g_{\mathrm{II}}(\lambda)=g(\lambda)$ (при гипотезе I отраженный сигнал отсутствует и на вход приемника поступает чистый шум). Тогда $\bar{\lambda}$ выбирается так, чтобы при $\lambda=\bar{\lambda}$ отношение $g(\lambda)|f(\lambda)$ принимало максимальное значение, а формула (26) упрощается и принимает вид

$$\bar{\mathbf{H}}=\frac{p_s^2 g^2(\bar{\lambda})}{2\pi f(\bar{\lambda})}. \tag{27}$$

Естественным обобщением рассмотренной выше ситуации является переход к случаю, когда вероятностные параметры каналов не известны точно и могут меняться в некоторых определенных пределах (аналогичные постановки в Шенновской задаче оптимального кодирования информации см. [7]). Подобная ситуация возникает, например, из-за того, что спектральная характеристика $g(\lambda)$ фильтра A (см. выше пример 6) может меняться в зависимости от свойств отражающего объекта, а спектральная плотность шума $f(\lambda)$ может меняться в зависимости от состояния среды, где распространяется сигнал, и т. п. Можно также представить себе теоретико-игровую ситуацию, где один из игроков („наблюдатель") стремится выбрать зондирующий сигнал так, чтобы уменьшить время наблюдения и вероятности ошибок, а его „противник" стремится воспрепятствовать этому, имея в своем распоряжении некоторые параметры каналов.

Математическое описание этой ситуации является следующим. Предполагается, что задана совокупность значений Γ_1 параметра γ_1 так, что каждому γ_1 сопоставляется канал, заданный системой плотностей $p_{\mathrm{I}}^{\gamma_1}\Big(x(t_1),$ $\ldots\; x(t_n)\,|\,s(t)\Big)$ (на языке математической статистики это означает, что гипотезы I и II о канале являются сложными). Требуется по результатам наблюдения установить, верна ли гипотеза I (т. е. верно, что посланный сигнал распространялся по одному из каналов, соответствующих $\gamma_1\in\Gamma_1$) или же верна гипотеза II (т. е. верно, что посланный сигнал распространялся по одному из каналов, соответствующих $\gamma_2\in\Gamma_2$).

Теперь информационное расстояние $\bar{\mathbf{H}}(\tilde{P}, \gamma_1, \gamma_2)$ является функцией не только распределения $\tilde{P}$ сигнала на входе, как было раньше, но и значений $\gamma_1\in\Gamma_1$, $\gamma_2\in\Gamma_2$ параметров, задающих канал.

* При $\bar{\lambda}=0$ множитель $\sqrt{2}$ должен быть, конечно, заменен на 1.

Мы будем характеризовать степень различности гипотез при заданном распределении $\tilde{P}$ величиной

$$\overline{\mathbf{H}}(\tilde{P}) = \min \overline{\mathbf{H}}(\tilde{P}, \gamma_1, \gamma_2), \tag{28}$$

где минимум берется по всевозможным парам значений параметров γ_1, γ_2. Таким образом, как это всегда делается в теории решающих функций (см. [8]), если гипотезы сложные вероятности неизвестны, мы будем ориентироваться на тот наихудший случай, когда параметры γ_1, γ_2 приняли самые неблагоприятные для различения гипотез значения. Мы будем называть здесь различающей способностью

$$\overline{\mathbf{H}} = \max_{\tilde{P}} \overline{\mathbf{H}}(\tilde{P}), \tag{28}$$

где максимум берется по всевозможным распределениям сигнала на входе $\tilde{P}$. Разумность подобного определения подтверждается следующим обобщением равенства (13). Обозначим через $T(F, G, \tilde{P}, \gamma_1, \gamma_2)$ функцию $T(F, G)$ для случая, когда распределением на входе канала является распределение $\tilde{P}$ и параметры канала приняли значение γ_1 и γ_2. Тогда

$$\max_{P} \min_{\gamma_1, \gamma_2} T(F, G, \tilde{P}, \gamma_1, \gamma_2) \approx \frac{(\alpha_F + \alpha_G)^2}{\overline{\mathbf{H}}}, \tag{30}$$

где, конечно, нужно ввести дополнительные предположения о близости гипотез I и II. Таким образом, различающая способность $\overline{\mathbf{H}}$ и здесь характеризует наименьшее время, за которое можно различить гипотезы I и II при оптимальном выборе распределения зондирующего сигнала.

Отметим, однако, что при рассмотренной постановке задачи способ обнаружения основан при фиксированных γ_1, γ_2 на соответствующем отношении правдоподобия и, значит, зависит от γ_1 и γ_2. Есть некоторые соображения, подсказывающие, что по крайней мере в наиболее интересных для приложений случаях (например, в приводимых ниже примерах) асимптотика времени наблюдения не изменится, если рассматривать лишь способы обнаружения, не зависящие от γ_1 и γ_2. Этот вопрос заслуживает дальнейшего изучения.

Отметим, что в рассматриваемом теперь случае сложных гипотез оптимальное распределение $\tilde{P}_0$ уже не обязано быть сосредоточенным на некоторой фиксированной функции, т. е. соответствовать детерминированному сигналу.

Пример 8. Вернемся к ситуации, рассмотренной в примере 7, но предположим, что введены множества значений параметра $\Gamma = \Gamma_{\mathrm{I}} = \Gamma_{\mathrm{II}}$ так, что спектр шума $f_\gamma(\lambda)$ зависит от параметра γ, причем при разных $\gamma \in \Gamma$ спектральная плотность $f_\gamma(\lambda)$ принимает всевозможные значения, подчиненные лишь одному дополнительному условию

$$\int_{\lambda_0}^{\lambda_0 + w} f_\gamma(\lambda)\, d\lambda \leqslant p_n^2. \tag{31}$$

Наглядно, мы предполагаем, что спектр шума нам не известен, но зато известно, что мощность шума не может превосходить P_n^2. Тогда оптимальным будет сигнал на входе вида

$$s(t) = \sqrt{2}\, p_s \cos(\Lambda t + \varphi), \tag{33}$$

где Λ—случайная величина с плотностью распределения

$$p_\Lambda(\lambda) = \frac{g^2(\lambda)}{\int\limits_{\lambda_0}^{\lambda_0+w} g^2(\lambda)\,d\lambda},$$

а

$$\overline{\mathbf{H}} = \frac{p_s^2}{p_n^2} \int\limits_{\lambda_0}^{\lambda_0+w} g^2(\lambda)\,d\lambda. \tag{34}$$

Пример 9. В ситуации примера 7 предположим в отличие от предыдущего примера, что спектр шума $f(\lambda) \neq 0$, $\lambda_0 \leqslant \lambda \leqslant \lambda_0 + w$ фиксирован, но зато зависит от параметра $\gamma \in \Gamma$; спектральная характеристика $g_\gamma(\lambda)$ линейного фильтра, принимающая всевозможные значения, подчинена следующему условию

$$\int\limits_{\lambda_0}^{\lambda_0+w} |g_\gamma(\lambda)|^2\,d\lambda \geqslant p_f^2. \tag{35}$$

Тогда оптимальный сигнал на входе представляется в таком же виде (33), что и в примере 8, с тем только отличием, что Λ—случайная величина с плотностью распределения

$$p_\Lambda(\lambda) = \frac{f(\lambda)}{p_n^2},$$

а

$$\overline{\mathbf{H}} = \frac{p_s^2 p_f^2}{p_n^2}. \tag{36}$$

Пример 10. Предположим теперь, что не известны как спектр шума, так и спектральная характеристика фильтра, причем они принимают всевозможные значения, удовлетворяющие условиям (31), (35) и условию

$$f(\lambda) \leqslant K^2. \tag{37}$$

Тогда оптимальный сигнал на входе также представляется в виде (33), но Λ имеет равномерное распределение на интервале $(\lambda_0, \lambda_0 + w)$, а

$$\mathbf{H} = \frac{p_s^2 p_f^2}{K^2}. \tag{38}$$

Обратим внимание на то, что если $K^2 = \infty$, то формула (38) показывает, что $\overline{\mathbf{H}} = 0$.

Рассмотрение величин (28) и (29) является естественным, говоря теоретико-игровым языком, с точки зрения „наблюдателя". „Противник" же, если „наблюдатель" выбрал наилучшую стратегию, т. е. распределение $\tilde{P}$, стремится обратить в минимум величину

$$\mathfrak{N}(\gamma_1, \gamma_2) = \max_{\tilde{P}} \overline{\mathbf{H}}(\tilde{P}, \gamma_1, \gamma_2). \tag{39}$$

Обозначим

$$\mathfrak{N} = \min_{\gamma_1, \gamma_2} \mathfrak{N}(\gamma_1, \gamma_2). \tag{40}$$

Представляется интересным изучение тех игровых ситуаций, в которых существует цена игры, т. е. $\overline{\mathbf{H}} = \mathfrak{N}$ [7].

Отметим, что в рассмотренных выше примерах $\overline{\mathbf{H}} = \mathfrak{N}$.

Приложение 1.

Чтобы точно сформулировать утверждение (9), введем следующее понятие (см. [9]): мы будем говорить, что имеет место энтропийная устойчивость гипотезы I относительно гипотезы II, если

$$\lim_{T\to\infty} \frac{l_T\left(\xi(t)\right)}{T} = \overline{\mathbf{H}}_{\mathrm{I/II}}, \tag{1.П}$$

где предел понимается в смысле сходимости по вероятности относительно распределения $p_{\mathrm{I}}(\cdot,\ \ldots,\ \cdot)$. Заметим, что $\mathbf{H}_{\mathrm{I/II}}(T)/T \approx \mathbf{H}_{\mathrm{I/II}}$ является математическим ожиданием величины, стоящей под знаком предела.

Энтропийная устойчивость является весьма общим свойством случайных процессов, родственным закону больших чисел. Она может быть получена как следствие этого закона, если время дискретно и при обеих гипотезах случайные величины $\xi(x)$, $x = \ldots, -1, 0, 1, \ldots$ — независимы и одинаково распределены. Тогда

$$l_T\left(\xi(t)\right) = \sum_{i=1}^{T} \log \frac{p_{\mathrm{I}}\left(\xi(i)\right)}{p_{\mathrm{II}}\left(\xi(i)\right)}. \tag{2.П}$$

и под знаком суммы стоят здесь независимые слагаемые, математические ожидания которых равны $\mathbf{H}_{\mathrm{I/II}}$. В общем случае доказательство сложнее, однако оно проведено для достаточно широкого класса случаев (см. [9], [10]), так что можно практически считать, что она верна всегда, когда при обоих гипотезах бесконечно далекое прошлое процесса не влияет на его бесконечно далекое будущее, т. е. если $\Delta\tau < \infty$.

Можно доказать следующую теорему: Если гипотеза I энтропийно устойчива относительно гипотезы II, то при любом фиксированном

$$\lim_{T\to\infty} \frac{\log G(F,\ T)}{T} = -\overline{H}_{\mathrm{I/II}}. \tag{3.П}$$

Доказательство этого факта представляет собой уточнение следующего рассуждения (мы будем для определенности считать время дискретным). Из (1) и (1.П) видно, что

$$G(F,\ T) = \int\limits_{l_T(x(t))<l_0}\!\!\ldots\int p_{\mathrm{II}}\left(x(1),\ \ldots\ x(T)\right)dx(1)\ \ldots\ dx(T) =$$

$$= \int\limits_{l_T(x(t))<l_0}\!\!\ldots\int e^{-l_T(x(t))}\, p_{\mathrm{I}}\left(x(1),\ \ldots\ x(T)\right)dx(1)\ \ldots\ dx(T) \approx$$

$$\approx e^{-\overline{H}_{\mathrm{I/II}}\cdot T} \int\limits_{l_T(x(t))<l_0}\!\!\ldots\int p_{\mathrm{I}}\left(x(1),\ \ldots\ x(T)\right)dx(1)\ \ldots\ dx(T) = e^{-\overline{H}_{\mathrm{I/II}}\cdot T}(1-F)$$

и так как F — постоянно, то отсюда следует предельное равенство (3.П). Отмеченные в основном тексте условия для применимости приближенного равенства (9) оправдываются тем, что, как можно доказать для простейших случаев (и что, по-видимому, верно и в достаточно широкой обстановке),

$$G(F,\ T) = \exp\left\{-\mathbf{H}_{\mathrm{I/II}}\cdot T\left(1 + O\left(\sqrt{\frac{\Delta\tau}{T}}\right)\right) + O\left(-\frac{\log G}{T}\Delta\tau\right)\right\}.$$

Приложение 2.

Точная формулировка утверждения (13) состоит в следующем. Рассматривается последовательность пар гипотез $\mathrm{I}^{(n)}$ и $\mathrm{II}^{(n)}$. Тогда все рассматриваемые в статье величины зависят от индекса n. Утверждается, что если $\overline{H}^{(n)} \to 0$ при $n \to \infty$, то равномерно по F и G в области $F+G<1-\varepsilon$, где ε – фиксировано,

$$\lim_{n\to\infty} \frac{T^{(n)}(F,\,G)\,\overline{H}^{(n)}}{(\alpha_F+\alpha_G)^2} = 1.$$

Достаточно общие и математически законченные условия, которые надо наложить на $p_{\mathrm{I}}^{(n)}$ и $p_{\mathrm{II}}^{(n)}$ для того, чтобы была верна сформулированная теорема, еще не найдены. Заведомо эти условия должны быть таковы, чтобы типичное время корреляции для процессов с распределениями $p_{\mathrm{I}}^{(n)}$ и $p_{\mathrm{II}}^{(n)}$ оставалось бы ограниченным и поэтому из $\overline{\mathrm{H}}^{(n)} \to 0$ следовало бы $\overline{\mathrm{H}}^{(n)}(\Delta\tau^{(n)}) \to 0$. В настоящее время эта теорема строго доказана для случая, когда рассматриваемые процессы являются последовательностями независимых величин, Айвазяном [5] (он не отметил равномерности по G и F, которую можно вывести из асимптотического разложения Крамера). Для случая, когда рассматриваются марковские процессы с дискретным временем и равномерно ограниченные снизу коэффициентом эргодичности (см. [11]), она может быть получена, исходя из известных предельных теорем для цепей Маркова. Для случая, когда процессы гауссовские, этот факт доказывается методами Пинскера [9]. Наконец, для общего случая здесь могут быть использованы рассуждения типа [12], хотя это еще и не сделано. Все же практически можно не сомневаться в приложимости соотношения (13) ко всем важным случаям.

Мы ограничимся здесь самым общим описанием идеи доказательства этой теоремы. Для простоты будем считать время дискретным, а процессы стационарными.

Разобьем отрезок $[0,\,T]$, где $T=K\Delta\tau$, на K равных отрезков длины $\Delta\tau$ с точками деления t_i, $i=0,\,\ldots\,K$. По определению времени корреляции $\Delta\tau$ значения процесса на разных отрезках приближенно независимы и значит

$$p_{\mathrm{I}}\big(x(1),\,\ldots\,x(T)\big) \approx p_{\mathrm{I}}\big(x(1),\,\ldots,\,x(t_1)\big)\,p_{\mathrm{I}}\big(x(t_1+1),\,\ldots\,x(t_2)\big)\times\ldots\times$$
$$\times\, p_{\mathrm{I}}\big(x(t_{n-1}+1),\,\ldots,\,x(t_n)\big),$$

$$p_{\mathrm{II}}\big(x(1),\,\ldots\,x(T)\big) \approx p_{\mathrm{II}}\big(x(1),\,\ldots,\,x(t_1)\big)\,p_{\mathrm{II}}\big(x(t_1+1),\,\ldots\,x(t_2)\big)\times\ldots\times$$
$$\times\, p_{\mathrm{II}}\big(x(t_{n-1}+1),\,\ldots\,x(t_n)\big).$$

Значит

$$l_T\big(x(t)\big) \approx \sum_{i=1}^{K} \log \frac{p_{\mathrm{I}}\big(x(t_i+1),\,\ldots\,x(t_{i+1})\big)}{p_{\mathrm{II}}\big(x(t_i+1),\,\ldots\,x(t_{i+1})\big)}\,. \tag{A}$$

Таким образом, $l_T\big(x(t)\big)$ будет при большом T суммой большого числа независимых величин и ее распределением при большом T можно считать гауссовским при обеих гипотезах. Нужно лишь сосчитать асимптотически математические ожидания $M_{\rm I}(T)$, $M_{\rm II}(T)$ и дисперсии $D_{\rm I}(T)$, $D_{\rm II}(T)$ величины $l_T(\cdot)$ при обеих гипотезах. Легко видеть из формулы (A), что

$$M_{\rm I}(T)=K\int\limits_{-\infty}^{\infty}\dots\int\log\frac{p_{\rm I}\left(x(0),\dots x(\Delta\tau)\right)}{p_{\rm II}\left(x(0),\dots x(\Delta\tau)\right)}\,p_{\rm I}\left(x(0),\ \dots\ x(\Delta\tau)\right)\times$$
$$\times\,dx(0)\ \dots\ dx(\Delta\tau),$$

$$D_{\rm I}(T)=K\int\limits_{-\infty}^{\infty}\dots\int\left[\log\frac{p_{\rm I}\left(x(0),\dots x(\Delta\tau)\right)}{p_{\rm II}\left(x(0),\dots x(\Delta\tau)\right)}\right]^2 p_{\rm I}\left(x(0),\ \dots\ x(\Delta\tau)\right)\times$$
$$\times\,dx(0)\ \dots\ dx(\Delta\tau)-\left(M_{\rm I}(T)\right)^2,$$

$$M_{\rm II}(T)=K\int\limits_{-\infty}^{\infty}\dots\int\log\frac{p_{\rm I}\left(x(0),\dots x(\Delta\tau)\right)}{p_{\rm II}\left(x(0),\dots x(\Delta\tau)\right)}\,p_{\rm II}\left(x(0),\ \dots\ x(\Delta\tau)\right)\times$$
$$\times\,dx(0)\ \dots\ dx(\Delta\tau),$$

$$D_{\rm II}(T)=K\int\limits_{-\infty}^{\infty}\dots\int\left[\log\frac{p_{\rm I}\left(x(0),\dots x(\Delta\tau)\right)}{p_{\rm II}\left(x(0),\dots x(\Delta\tau)\right)}\right]^2 p_{\rm II}\left(x(0),\ \dots\ x(\Delta\tau)\right)\times$$
$$\times\,dx(0)\ \dots\ dx(\Delta\tau)-\left(M_{\rm II}(T)\right)^2.$$

Из предположения $H^{(n)}(\Delta\tau)\to 0$ следует, что мы можем считать $H(\Delta\tau)$ малым. Но это означает, что

$$\frac{p_{\rm I}\left(x(0),\ \dots,\ x(\Delta\tau)\right)}{p_{\rm II}\left(x(0),\ \dots,\ x(\Delta\tau)\right)}$$

близко к 1 (ведь см. [2] равенство нулю энтропии эквивалентно совпадению плотностей). Поэтому, разлагая соответствующую функцию в ряд Тейлора, получаем, что (мы обозначаем для краткости все аргументы одной буквой)

$$M_{\rm I}(T)\approx -K\int\log\frac{p_{\rm II}(x)}{p_{\rm I}(x)}\,p_{\rm I}(x)\,dx\approx -K\int\left(p_{\rm II}(x)-p_{\rm I}(x)\right)dx-$$
$$-\frac{K}{2}\int\frac{\left(p_{\rm II}(x)-p_{\rm I}(x)\right)^2}{p_{\rm I}(x)}\,dx=\frac{K}{2}\int\frac{\left(p_{\rm II}(x)-p_{\rm I}(x)\right)^2}{p_{\rm I}(x)}\,dx,$$

$$M_{\rm II}(T)\approx\frac{K}{2}\int\frac{\left(p_{\rm I}(x)-p_{\rm II}(x)\right)^2}{p_{\rm II}(x)}\,dx,$$

$$D_{\rm I}(T)\approx K\int\frac{\left(p_{\rm I}(x)-p_{\rm II}(x)\right)^2}{p_{\rm I}(x)}\,dx,$$

$$D_{\rm II}(T)\approx\int\frac{\left(p_{\rm I}(x)-p_{\rm II}(x)\right)^2}{p_{\rm II}(x)}\,dx,$$

или учитывая, что по определению $M_{\mathrm{I}}(T)=H_{\mathrm{I/II}}(T)$, $M_{\mathrm{II/I}}(T)=-H_{\mathrm{II/I}}(T)$ и что $p_{\mathrm{I}}(x)\approx p_{\mathrm{II}}(x)$, обнаруживаем, что

$$\begin{aligned} M_{\mathrm{I}}(T)&\approx -M_{\mathrm{II}}(T)\approx\frac{H(T)}{2}\approx\frac{\overline{H}T}{2}, \\ D_{\mathrm{I}}(T)&\approx D_{\mathrm{II}}(T)\approx H(T)\approx\overline{H}T. \end{aligned} \qquad (\text{Б})$$

Таким образом, мы можем считать, что логарифм отношения правдоподобия распределен, при обеих гипотезах, нормально с параметрами, заданными равенствами (Б). Поэтому из определения чисел F и G и уровня l_0 следует, что

$$l_0=-\frac{\overline{H}T}{2}+\alpha_F\sqrt{\overline{H}T}, \qquad l_0=\frac{\overline{H}T}{2}-\alpha_G\sqrt{\overline{H}T}.$$

Приравнивая это равенство, выводим утверждение (13).

Поступила в редакцию
29.VI.1962

СПИСОК ЛИТЕРАТУРЫ

1. Вайнштейн Л. А., Зубаков В. Д., Выделение сигналов на фоне случайных помех, Сов. Радио, Москва, 1960.
2. Kullback S., Information theary and statistics, N.—Y., Willey, 1959.
3. Санов И. Н., О вероятности больших отклонений случайных величин, Матем. сб., 42, № 1, П-44, 1957.
4. Mourier E., Etude du chaix entre deux lais de prabability, C. R. Ac. Sci. Paris, 223, 19, 712—714, 1946.
5. Айвазян С. А., Сравнение оптимальных свойств критериев Неймана—Пирсона и Вальда, Теор. вер. и её примен., IV, № 1, 86—93, 1959.
6. Петров А. А., Проверка статистических гипотез о типе распределения по малым выборкам, Теор вер. и её примен., 1, № 2, 248—271.
7. Добрушин Р. Л., Оптимальная передача информации по каналу с неизвестными параметрами, Рад. и эл., IV, № 12, 1951—1956, 1959.
8. Блекуэлл Д., Гиршик М. А., Теория игр и статистических решений, ИЛ, Москва, 1958.
9. Пинскер М. С., Информация и информационная устойчивость случайных величин и процессов. Проблемы передачи информации, вып. 7, Изд. АН, 1960.
10. Feinstein A., Faundations of information theory, N.—Y., Mc Craw-Hill, 1958.
11. Добрушин Р. Л., Центральная предельная теорема для неоднородных цепей Маркова, Теор. вер. и её примен., 1, № 1, 72—89, № 4, 365—426, 1956.
12. Волконский В. А., Розанов Ю. А., Некоторые предельные теоремы для случайных функций, Теор. вер. и её примен., 4, № 02, 186—207, 1959.

ENTROPIJOS SĄVOKOS PRITAIKYMAS SIGNALO SURADIMO TRIUKŠMO FONE PROBLEMOSE

R. DOBRUŠINAS, M. PINSKERIS, A. ŠIRIAJEVAS

Reziumė

Šiame darbe randamos asimptotinės išraiškos signalo suradimo triukšmo fone uždavinio paklaidų tikimybėms.

Įvedamos suradimo kanalo ir jo skiriamosios galios sąvokos. Šios sąvokos leidžia suformuluoti zonduojančio kanalo optimalios formos uždavinį. Eilei konkrečių atvejų šis uždavinys išsprendžiamas.

APPLICATION OF THE NOTION OF ENTROPY IN THE PROBLEMS OF DETECTING A SIGNAL IN NOISE

R. L. DOBRUSHIN, M. S. PINSKER, A. N. SHIRIAJEV (MOSCOW)

Summary

In this work asymptotic expressions for error probability in the problem of detecting a signal in noise are found.

The concept of detecting canal and its distinctive capacity are introduced, there concepts permit to raise the problem of choosing the optimal form of the sounding canal, this problem is solved for some concrete situations.