Threshold Decoding for Disjunctive Group Testing

D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.

Shchukin V.Yu.

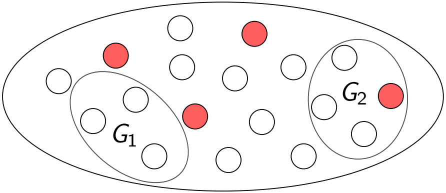
Lomonosov Moscow State University, Faculty of Mechanics and Mathematics

ACCT 2016

Group Testing Problem

Given a pool of t elements. $\mathcal{S}_{un} \subset [t]$ – a set of defective elements.

Group test: $G \subset [t]$. The test result: $y = \begin{cases} 1, & |G \cap \mathcal{S}_{un}| > 0, \\ 0, & |G \cap \mathcal{S}_{un}| = 0. \end{cases}$



$y_1 = 0$
$y_2 = 1$

Conventional Problem

Assumption: the number of defects $\leqslant s$.
Problem: to find all defective elements.

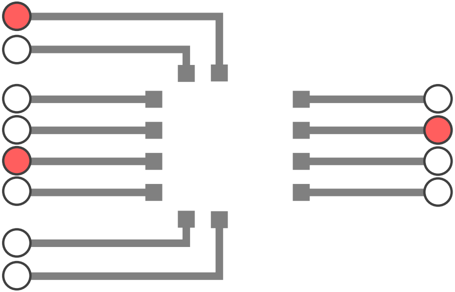
Threshold Problem

Problem: to decide whether $|\mathcal{S}_{un}| \leqslant s$ or $|\mathcal{S}_{un}| > s$.

Motivation of Threshold Problem

[1976, Malyutov]

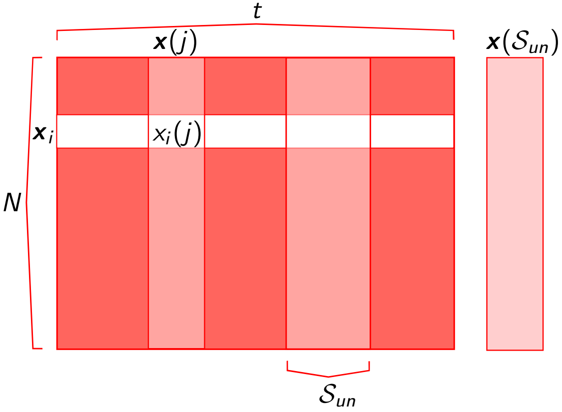
Given a complex electronic circuit of size t.
The circuit is not working properly if $> s$ elements are defective.
Such a circuit should be changed by a new similar circuit.



Testing time should be small.
We consider only nonadaptive algorithms.

Binary Code

N tests $\longleftrightarrow$ binary $(N \times t)$-matrix X (code X with codewords $\boldsymbol{x}(1), \ldots, \boldsymbol{x}(t)$)



$$x_i(j) = 1$$
$$\Longleftrightarrow$$
j-th element is in i-th test

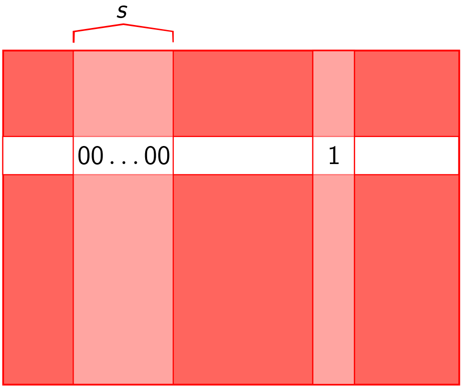
Result column:

$$\boldsymbol{x}(\mathcal{S}_{un}) = \bigvee_{j \in \mathcal{S}_{un}} \boldsymbol{x}(j)$$

Conventional Decoding Algorithm

We say that a binary column $\mathbf{u}$ covers a binary column $\mathbf{v}$ if $\mathbf{u}_i \geqslant \mathbf{v}_i, \forall\, i \in [N]$.

Definition. A code X is called a disjunctive s-code if the disjunctive sum of any s codewords of X does not cover any other codeword.



To find all $\leqslant s$ defects using disjunctive s-code, we find all codewords, which are covered by the result column.

Necessity and Sufficiency

Conventional Problem

Assumption: $|\mathcal{S}_{un}| \leqslant s$.
Problem: to find $\mathcal{S}_{un}$.

Threshold Problem

Problem: to decide whether
$|\mathcal{S}_{un}| \leqslant s$ or $|\mathcal{S}_{un}| > s$.

Proposition 1. A code X is disjunctive s-code $\Longrightarrow$ X solves conventional problem.

Proposition 2. A code X solves conventional problem $\Longrightarrow$ X is disjunctive $(s-1)$-code.

Proposition 3. A code X is disjunctive s-code $\Longrightarrow$ X solves threshold problem.

Proposition 4. A code X solves threshold problem $\Longrightarrow$ X is disjunctive s-code.

Hypothesis Testing

The null hypothesis
$H_0 \; : \; |\mathcal{S}_{un}| \leqslant s$

The alternative
$H_1 \; : \; |\mathcal{S}_{un}| \geqslant s+1$

The probability distribution of the random collection $\mathcal{S}_{un}$, is identified by vector $\boldsymbol{p} \triangleq (p_0, p_1, \ldots, p_t)$, $p_k \geqslant 0$, $\sum_{k=0}^{t} p_k = 1$, as follows:

$$\Pr\{\mathcal{S}_{un} = \mathcal{S}\} \triangleq \frac{p_{|\mathcal{S}|}}{\binom{t}{|S|}} \quad \text{for any subset} \quad \mathcal{S} \subset [t].$$

Hypothesis Testing

The null hypothesis
$H_0 \ : \ |\mathcal{S}_{un}| \leqslant s$

The alternative
$H_1 \ : \ |\mathcal{S}_{un}| \geqslant s+1$

Suppose we have some group testing procedure (code X) and some decision rule (DR).
Introduce maximal error probability

$$\varepsilon_s(\mathrm{DR}, \boldsymbol{p}, X) \triangleq \max \left\{ \Pr\{\text{accept } H_1 \big| H_0\}, \Pr\{\text{accept } H_0 \big| H_1\} \right\}.$$

Hypothesis Testing

The null hypothesis
$$H_0 \; : \; |\mathcal{S}_{un}| \leqslant s$$

The alternative
$$H_1 \; : \; |\mathcal{S}_{un}| \geqslant s+1$$

$$\varepsilon_s(\mathrm{DR}, \boldsymbol{p}, X) \triangleq \max \left\{ \Pr\{\text{accept } H_1 \big| H_0\} , \Pr\{\text{accept } H_0 \big| H_1\} \right\}.$$

Introduce universal error probability for the DR

$$\varepsilon_s^N(\mathrm{DR}, R) \triangleq \max_{\boldsymbol{p}} \left\{ \min_X \varepsilon_s(\mathrm{DR}, \boldsymbol{p}, X) \right\},$$

where the minimum is taken over all codes X of length N and size $t = \lfloor 2^{RN} \rfloor$.

Hypothesis Testing

The null hypothesis
$H_0 \; : \; |\mathcal{S}_{un}| \leqslant s$

The alternative
$H_1 \; : \; |\mathcal{S}_{un}| \geqslant s+1$

$$\varepsilon_s(\mathrm{DR}, \boldsymbol{p}, X) \triangleq \max \left\{ \Pr\{\text{accept } H_1 \big| H_0\}, \, \Pr\{\text{accept } H_0 \big| H_1\} \right\}.$$

$$\varepsilon_s^N(\mathrm{DR}, R) \triangleq \max_{\boldsymbol{p}} \left\{ \min_{X} \varepsilon_s(\mathrm{DR}, \boldsymbol{p}, X) \right\},$$

The error exponent

$$E_s(\mathrm{DR}, R) \triangleq \overline{\lim_{N\to\infty}} \frac{-\log_2 \varepsilon_s^N(\mathrm{DR}, R)}{N},$$

identifies the asymptotic behavior of α-level of significance for the decision rule, i.e.,

$$\alpha \triangleq 2^{-N\,[E_s(\mathrm{DR},R)+o(1)]}, \quad \text{if} \quad E_s(\mathrm{DR}, R) > 0, \quad N \to \infty.$$

Disjunctive Decision Rule (Previous Results)

The disjunctive decision rule (DDR) is based on the conventional decoding algorithm:

$$\begin{cases} \text{accept } \{H_0 : |\mathcal{S}_{un}| \leqslant s\} & \text{if } \boldsymbol{x}(\mathcal{S}_{un}) \text{ covers} \leqslant s \text{ codewords of } X, \\ \text{accept } \{H_1 : |\mathcal{S}_{un}| \geqslant s+1\} & \text{if } \boldsymbol{x}(\mathcal{S}_{un}) \text{ covers} \geqslant s+1 \text{ codewords of } X. \end{cases}$$

The maximum in universal error probability for the DDR is attained at the distribution vector

$$\boldsymbol{p} = (0, 0, \ldots, 0, \underset{s\uparrow}{1}, 0, \ldots, 0)$$

[2015, D'yachkov et al]

Theorem 1. If $R \geqslant \frac{1}{s}$, then $E_s(\mathrm{DDR}, R) = 0$.

Theorem 2. If $R < \frac{\ln 2}{s}$, then $E_s(\mathrm{DDR}, R) > 0$.

Threshold Decision Rule (New Results)

For a fixed parameter τ, $0 < \tau < 1$, introduce the τ-threshold decision rule:

$$\begin{cases} \text{accept } \{H_0 \,:\, |\mathcal{S}_{un}| \leqslant s\} & \text{if } |\boldsymbol{x}(\mathcal{S}_{un})| \leqslant \lfloor \tau N \rfloor, \\ \text{accept } \{H_1 \,:\, |\mathcal{S}_{un}| \geqslant s+1\} & \text{if } |\boldsymbol{x}(\mathcal{S}_{un})| \geqslant \lfloor \tau N \rfloor + 1. \end{cases}$$

The maximum in universal error probability for the DDR is attained at the distribution vector

$$\boldsymbol{p} = (0, 0, \ldots, 0, \tfrac{1}{2}, \tfrac{1}{2}, 0, \ldots, 0)$$

$$s \Big\uparrow$$

Theorem. The error exponent for τ-threshold decision rule satisfies the inequality:

$$E_s(\tau, R) \geqslant \underline{E}_s(\tau).$$

The optimal value of error exponent is not less than

$$\underline{E}_{\text{Thr}}(s) \triangleq \max_{0<\tau<1} \underline{E}_s(\tau, R) \geqslant \frac{\log_2 e}{4s^2}(1 + o(1)), \quad s \to \infty.$$

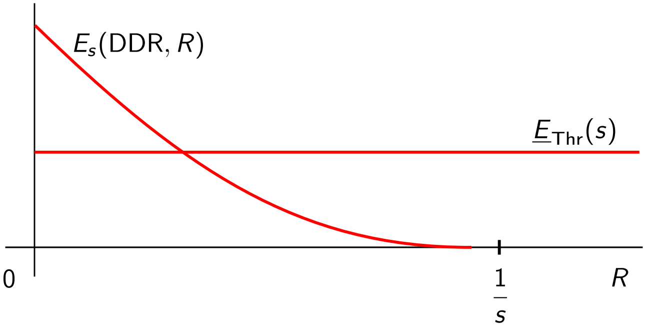
Comparison of Decision Rules (Error Exponent)

Disjunctive decision rule (DDR)

$$E_s(\mathrm{DDR}, R) = 0 \quad \text{if} \quad R \geqslant \frac{1}{s}.$$

Threshold decision rule (TDR)

$$E_s(\mathrm{TDR}, R) \geqslant \underline{E}_{\mathrm{Thr}}(s) \geqslant \frac{\log_2 e}{4s^2}.$$



Comparison of Decision Rules (Simulation)

The probability distribution vector $\boldsymbol{p}$ is such that $|\mathcal{S}_{un}|$ has binomial distribution with the mean $s + \frac{1}{2}$.

Code X of length N and size t is generated randomly from the ensemble of constant weight codes, i.e. every codeword of X has weight w.

$$s = 2, \quad t = 15$$

	Threshold decision rule			Disjunctive decision rule	
N	$\varepsilon_s(\mathrm{TDR}, \boldsymbol{p}, X)$	w	$\lfloor \tau N \rfloor$	$\varepsilon_s(\mathrm{DDR}, \boldsymbol{p}, X)$	w
5	**0.1366**	2	3	0.4780	2
8	**0.0824**	3	5	0.3610	2
10	**0.0744**	1	2	0.2390	3
12	**0.0440**	1	2	0.1220	3
14	**0.0349**	2	4	0.0537	3
15	0.0258	2	4	**0.0195**	3

Comparison of Decision Rules (Simulation)

The probability distribution vector $\boldsymbol{p}$ is such that $|\mathcal{S}_{un}|$ has binomial distribution with the mean $s + \frac{1}{2}$.

Code X of length N and size t is generated randomly from the ensemble of constant weight codes, i.e. every codeword of X has weight w.

$$s = 2, \quad t = 20$$

	Threshold decision rule			Disjunctive decision rule	
N	$\varepsilon_s(\text{TDR}, \boldsymbol{p}, X)$	w	$\lfloor \tau N \rfloor$	$\varepsilon_s(\text{DDR}, \boldsymbol{p}, X)$	w
5	**0.1398**	2	3	0.5356	2
8	**0.0897**	3	5	0.4169	2
10	**0.0897**	3	5	0.3008	3
12	**0.0580**	4	7	0.1979	3
14	**0.0429**	2	4	0.1214	4
15	**0.0324**	2	4	0.0792	4
17	**0.0246**	2	4	0.0449	4
18	**0.0139**	3	6	0.0290	5
20	0.0148	3	6	**0.0026**	5

Thank you for your attention!