

## ПОХИБКА ПРИ ОБЧИСЛЕННІ ЙМОВІРНОСТІ ПРАВИЛЬНОГО РОЗПІЗНАВАННЯ КЛАСУ ПОВІТРЯНОЇ СИТУАЦІЇ

*Розглянуто результати дослідження похибки, що враховує неточне задання щільності при обчисленні ймовірності правильного розпізнавання класу повітряної ситуації. Наведено залежності похибки від кількості вимірювань і обсягу навчальної вибірки. Дано рекомендації щодо розрахунку оптимальних вихідних параметрів для отримання мінімальної похибки і, як результат, максимальної ймовірності правильного розпізнавання.*

Використання багатоальтернативного підходу до класифікації стану повітряного корабля (ПК) на етапі польоту дозволяє значно підвищити безпеку авіаційних перевезень і зменшити ризик виникнення конфліктної ситуації. До виникнення спектру ситуацій призводить дія різних факторів, що знижують точність визначення місця розташування ПК. Крім того, чималий вплив справляє людський фактор.

При польоті по трасі використовується класифікація на п'ять повітряних станів залежно від положення літака у просторі [1]: нормальна ситуація (НС), ускладнення умов польоту (УУП), складна ситуація (СС), аварійна ситуація (АС) і катастрофічна ситуація (КС). На рис.1 наведено приклад класифікації повітряного стану ПК у вертикальній площині.

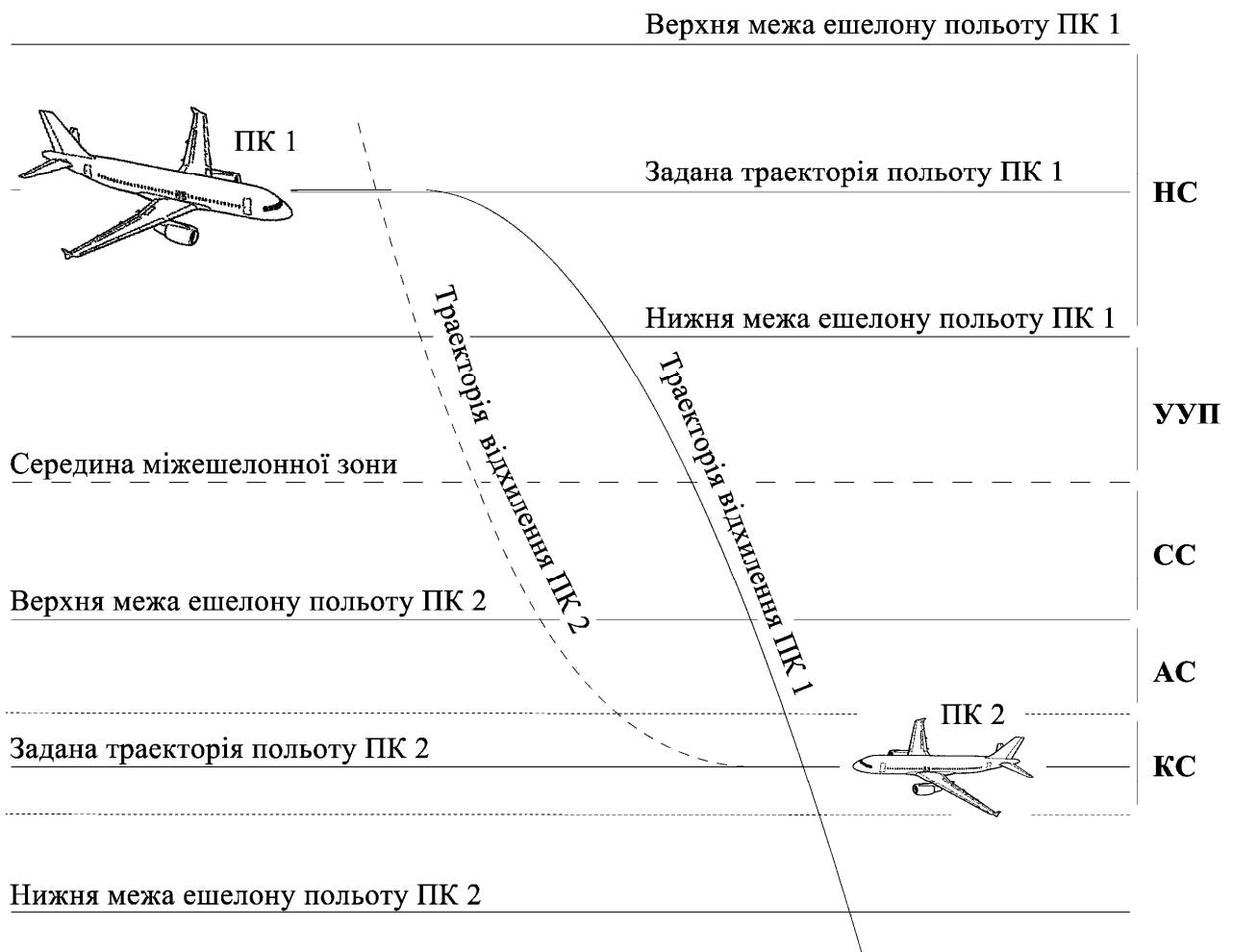


Рис.1. Класифікація стану ПК у вертикальній площині

Нехай у межах кожного класу оцінювана величина (наприклад, висота польоту) розподілена за нормальним законом, рис.2. Кожному з класів  $S_k$ , де  $k=\overline{1,5}$ , який характеризується математичним сподіванням  $m_k$  та дисперсією  $\sigma_k^2$ , відповідає своя апіорна ймовірність  $p_k$ .

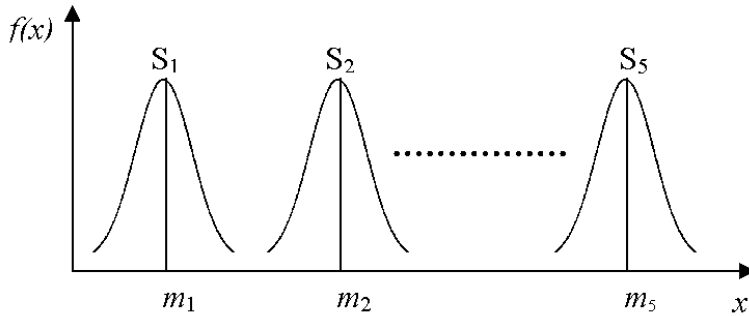


Рис. 2. П'ять класів ситуацій

вимірювань нам відоме неточне значення щільності  $\hat{\rho}(x)$ . Причиною неточності може слугувати вплив неточно відомого значення дисперсії  $\sigma_k^2$  рис.3.

Необхідно враховувати цю неточність у формулі Баєса при розрахунку ймовірності правильного вибору класу.

Після виконання вимірювань маємо  $n$  незалежних спостережень  $x_1, \dots, x_n$  параметра  $x$ , який є однією з просторових координат ПК (наприклад абсолютною висотою польоту). Кожному з вимірювань відповідає своя щільність  $\rho_1(x), \dots, \rho_5(x)$ . Враховуючи, що щільності  $\rho_1(x), \dots, \rho_5(x)$  відомі неточно, позначимо їх відомі наближені значення через  $\hat{\rho}_1(x), \dots, \hat{\rho}_5(x)$ .

Побудуємо щільності

$$\hat{\rho}_k^{(n)}(x^{(n)}) = \prod_{i=1}^n \hat{\rho}_k(x_i), \quad x^{(n)} = (x_1 \dots x_n), \quad k = \overline{1,5}.$$

Рішення щодо належності ситуації до  $k$ -го класу може бути прийнято за умови

$$p_k \hat{\rho}_k^{(n)}(x^{(n)}) = \max \{ p_j \hat{\rho}_j^{(n)}(x^{(n)}) \}, \quad j = \overline{1,5}.$$

Якщо максимум досягається на декількох класах, тоді рішення приймається на користь будь-якого з класів при відповідних  $j_1 \dots j_p$ . Наприклад, на користь  $S_{j_{\min}}$ , де  $j_{\min} := \min(j_1 \dots j_p)$ .

Якщо позначити через  $\Omega_k$  область у просторі реалізацій  $x^{(n)}$ , де рішення приймається на користь  $H_k$ , тоді ймовірність правильного розпізнавання рівна

$$P_{\Pi} = \sum_{k=1}^5 p_k \int_{\Omega_k} \rho_k^{(n)}(x^{(n)}) dx^{(n)},$$

$$P_{\Pi} = \sum_{k=1}^5 p_k \int_{\Omega_k} \hat{\rho}_k^{(n)}(x^{(n)}) dx^{(n)} + \sum_{k=1}^5 p_k \int_{\Omega_k} \left( -\frac{\hat{\rho}_k^{(n)}(x^{(n)})}{\rho_k^{(n)}(x^{(n)})} + 1 \right) \rho_k^{(n)}(x^{(n)}) dx^{(n)}, \quad (1)$$

$$P_{\Pi} = \hat{P}_{\Pi} + R_1,$$

де

$\hat{P}_{\Pi}$  - імовірність правильного розпізнавання, якби істинними були щільності  $\hat{\rho}_k$ , а не  $\rho_k$ ,

$R_1$  - похибка, що враховує неточність задання щільності розподілу.

$$\text{Маємо } P_{\Pi} \geq \hat{P}_{\Pi} - |R_1|,$$

Крім того, щільність розподілу ймовірності  $\rho(x)$  для кожного класу матиме різні залежності. Зазвичай ймовірність правильного розпізнавання класу ситуації знаходиться за максимумом апостеріорної ймовірності, що розраховується за формулою Баєса. На практиці при виконанні

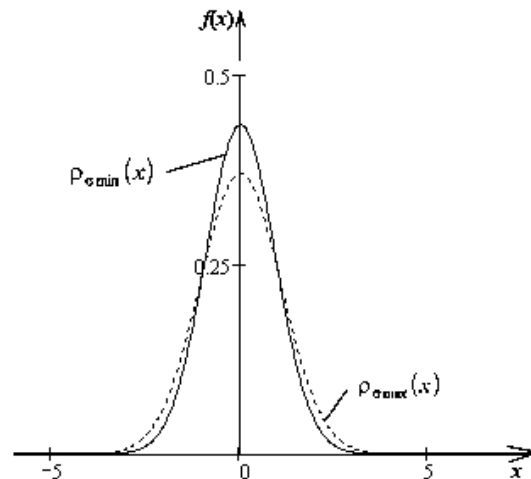
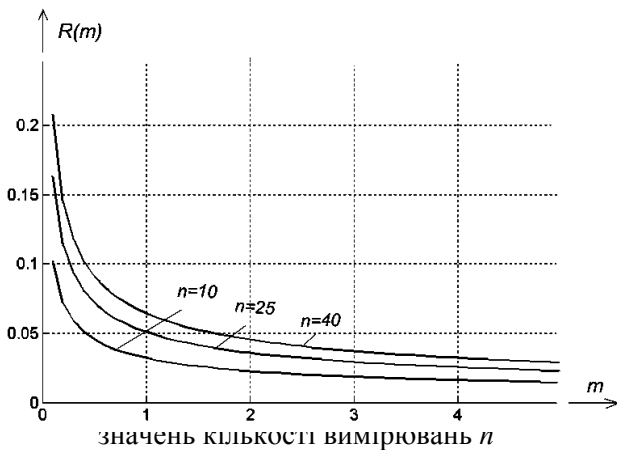


Рис.3. Неточне значення щільності

Ймовірність правильного розпізнавання класу розглядається як апостеріорна ймовірність розрахована за формулою Баєса [2]



$$\max_{1 \leq k \leq S} \hat{q}_k(x^{(n,j)}) = \frac{\max_{1 \leq k \leq S} (p_k \hat{p}_k^{(n)}(x^{(n,j)}))}{\sum_{k=1}^S p_k \hat{p}_k^{(n)}(x^{(n,j)})} \quad (2)$$

$$\hat{P}_{II} = \hat{E} \max_{1 \leq k \leq S} \hat{q}_k(x^{(n,j)}), \quad (3)$$

З (1) маємо

де  $\hat{E}$  - символ математичного сподівання за умови, що  $x^{(n,j)}$  має щільність розподілу

$$\sum_{k=1}^S p_k \hat{p}_k^{(n)}(x^{(n,j)}), \quad 1 \leq j \leq M.$$

Тоді наближене значення  $\hat{P}_{II}$  отримується

шляхом застосування методу Монте-Карло до формули Баєса(2).

Оцінка похибки  $R_1$  може бути представлена нерівністю [3]

$$|R| \leq \sum_{k=1}^S p_k \omega_\phi(\delta_1^{(k)}(n, m), \delta_2^{(k)}(n, m)) + 1 - (1 - \alpha)^5,$$

де  $\alpha$  - рівень довіри.

Функція  $\omega_\phi(\delta_1^{(k)}(n, m), \delta_2^{(k)}(n, m))$  відповідає максимуму функції  $\phi(a, b)$ , що має вигляд

$$\phi(a, b) = 2\Phi\left(-\frac{a}{b}\right) - 1 + e^{\frac{a+b^2}{2}} \left(1 - 2\Phi\left(-b - \frac{a}{b}\right)\right)$$

для параметрів, що лежать у межах:  $-\delta_1 \leq a \leq 0, 0 \leq b \leq \delta_2$ .

Граничні значення для параметрів  $a$  та  $b$  розраховуються за наступними формулами [3]

$$\delta_1 = \delta_1(n, m) = -\frac{n}{2} [h(m) + \ln(1 - h(m))], \quad \delta_2 = \delta_2(n, m) = \frac{\sqrt{n} \cdot \sqrt{3} \cdot n_{\alpha/2}}{\sqrt{m}},$$

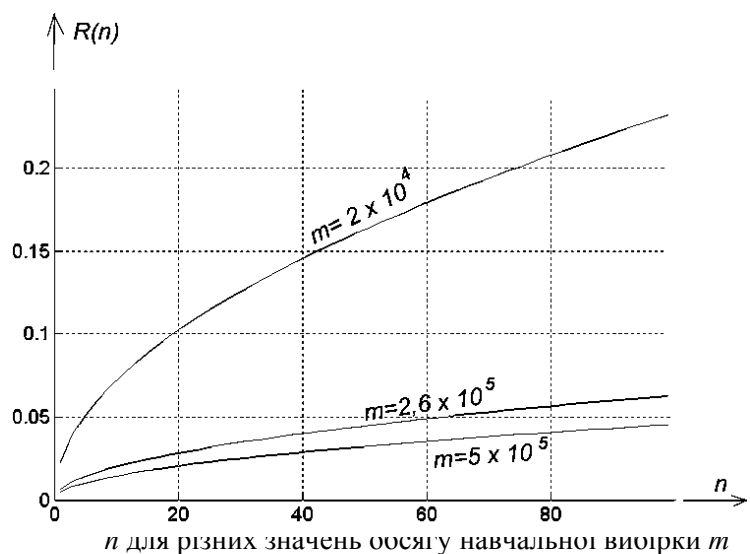
де

$$h(m) = \frac{\sqrt{2} \cdot n_{\alpha/2}}{\sqrt{m}} \quad \text{при умові, що} \\ 0 < h < 1,$$

$n_{\alpha/2}$  -  $\frac{\alpha}{2}$  квантиль нормального закону (функція є табличною і оберненою до функції нормального розподілу).

Граничні значення залежать від обсягу навчальної вибірки  $m$  і кількості вимірювань  $n$ .

На рис.4. наведено залежність похибки, що враховує неточність задання щільності розподілу від обсягу навчальної вибірки для декількох значень кількості вимірювань, а на рис. 5. представлені залежності похибки  $R$  від кількості вимірювань  $n$  для різних значень обсягу навчальної вибірки  $m$ .



З графіка залежності похибки  $R$  від кількості вимірювань  $n$  і від обсягу навчальної вибірки  $m$  (рис.6) видно, що похибка зростає зі збільшенням кількості вимірювань і зменшується зі збільшенням обсягу навчальної вибірки. Тоді для зменшення похибки необхідно виконувати якомога менше вимірювань висоти польоту (чи однієї з просторових тримірних координат літака у просторі). Але зменшення кількості вимірювань приводить до значного зменшення ймовірності правильного розпізнавання, розрахованої за правилом Баєса (3). Тому зменшити величину похибки і максимізувати оцінку ймовірності правильного розпізнавання класу ситуації стає можливим за умов зміни обсягу основної вибірки. Обсяг навчальної вибірки значно впливає на величину похибки тільки при малих значеннях  $m$ . При збільшенні обсягу навчальної вибірки величина похибки практично не змінюється. Отже можна стверджувати, що значне збільшення обсягу навчальної вибірки не впливає на величину похибки.

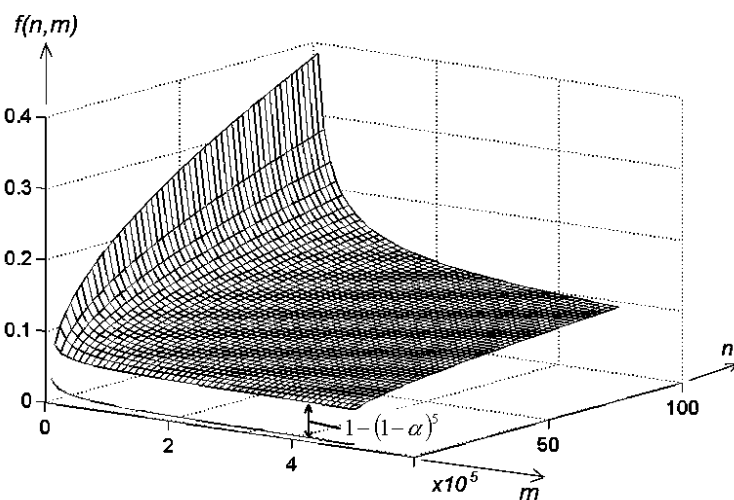


Рис. 6. Загальна похибка розрахунку ймовірності правильного розпізнавання

Для пошуку оптимальних значень кількості вимірювань і обсягу навчальної вибірки необхідно обмежити похибку максимально допустимим значенням. Наприклад, якщо обмежити похибку величиною рівною 0.1, тоді можна буде гарантувати розпізнавання класу ситуації з ймовірністю не меншою за 90%

$$R(n, m) + 1 - (1 - \alpha)^5 \leq 0.1.$$

Відповідні значення  $n$  і  $m$  знаходяться під площиною, проведеною через  $R=0.1$ . З спектру значень обсягу основної вибірки вибирається значення, що відповідатиме максимальній ймовірності правильного розпізнавання класу ситуації, розрахованої за формулою Баєса (3). Таким чином, знаходяться оптимальні значення для  $n$  та  $m$ , які гарантували б розпізнавання класу ситуації з ймовірністю не меншою за 90%.

В результаті доведено наявність похибки, що присутня при визначенні ймовірності правильного розпізнавання класу повітряної ситуації при умові, що щільності розподілу відомі неточно і наведена методика оцінки її впливу. Проведено дослідження залежності похибки  $R$  від обсягу навчальної вибірки і кількості вимірювань рис.4., рис.5., рис.6. Дано рекомендації щодо зменшення її впливу на ймовірність правильного розпізнавання класу повітряної ситуації.

### Список літератури

1. Харченко В.П., Косенко Г.Г. Многоальтернативный последовательный метод в задачах ситуационного анализа воздушной обстановки // Моделирование радиоэлектронных систем и комплексов обеспечения полётов: Сб. науч. тр. – К.:КМУГА, 1996.-С. 3-10.
2. Бабак В.П., Харченко В.П., Максимов В.О., та інші. Безпека авіації. Монографія. -К.: Техніка, 2004.- 584 с.
3. Остроумов И.В., Кукуш А.Г., Харченко В.П. Оценка вероятности правильного распознавания по правилу Байеса при неточно известной плотности распределения