

Введение в статистическую теорию распознавания образов. Байесовский подход.

Тверская Е. С.
e_tverskaya@bmstu.ru

Московский государственный технический университет им. Н. Э. Баумана
(национальный исследовательский университет)

Москва, 2025



ИУ-6
Компьютерные
системы и сети

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Пусть $\mathbf{X}$ – вектор наблюдений.

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Пусть $\mathbf{X}$ – вектор наблюдений.

Рассмотрим 2-е *гипотезы*:

- H_1 – объект принадлежит к классу ω_1 ;
- H_2 – объект принадлежит к классу ω_2 .



Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

Пусть $\mathbf{X}$ – вектор наблюдений.

Рассмотрим 2-е **гипотезы**:

- H_1 – объект принадлежит к классу ω_1 ;
- H_2 – объект принадлежит к классу ω_2 .

Постановка задачи.

Определить принадлежность вектора $\mathbf{X}$ к классу ω_1 или ω_2 .



Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

Пусть $\mathbf{X}$ – вектор наблюдений.

Рассмотрим 2-е **гипотезы**:

- H_1 – объект принадлежит к классу ω_1 ;
- H_2 – объект принадлежит к классу ω_2 .

Постановка задачи.

Определить принадлежность вектора $\mathbf{X}$ к классу ω_1 или ω_2 .

Решающее правило

$$P(\omega_1|\mathbf{X}) \geq P(\omega_2|\mathbf{X}) \implies \mathbf{X} \in \left\{ \begin{array}{l} \omega_1 \\ \omega_2 \end{array} \right.,$$

где $P(\omega_1|\mathbf{X})$ и $P(\omega_2|\mathbf{X})$ – апостериорные вероятности.



Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

Пусть $\mathbf{X}$ – вектор наблюдений.

Рассмотрим 2-е **гипотезы**:

- H_1 – объект принадлежит к классу ω_1 ;
- H_2 – объект принадлежит к классу ω_2 .

Постановка задачи.

Определить принадлежность вектора $\mathbf{X}$ к классу ω_1 или ω_2 .

Решающее правило

$$P(\omega_1|\mathbf{X}) \geq P(\omega_2|\mathbf{X}) \implies \mathbf{X} \in \left\{ \begin{array}{l} \omega_1 \\ \omega_2 \end{array} \right.,$$

где $P(\omega_1|\mathbf{X})$ и $P(\omega_2|\mathbf{X})$ – апостериорные вероятности.

По теореме Байеса $P(\omega_i|\mathbf{X}) = \frac{P(\omega_i)f(\mathbf{X}|\omega_i)}{P(\mathbf{X})}$, $i = \{1, 2\}$.

Решающее правило

$$f(\mathbf{X}|\omega_1)P(\omega_1) \geq f(\mathbf{X}|\omega_2)P(\omega_2) \implies \mathbf{X} \in \left\{ \begin{array}{l} \omega_1 \\ \omega_2 \end{array} \right.,$$

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Введем в рассмотрение величину, которая называется **отношение правдоподобия** $l(X) = f(X|\omega_1)/f(X|\omega_2)$:

$$l(X) = \frac{f(X|\omega_1)}{f(X|\omega_2)} \geq \frac{P(\omega_2)}{P(\omega_1)} \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Введем в рассмотрение величину, которая называется **отношение правдоподобия** $l(X) = f(X|\omega_1)/f(X|\omega_2)$:

$$l(X) = \frac{f(X|\omega_1)}{f(X|\omega_2)} \geq \frac{P(\omega_2)}{P(\omega_1)} \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$

Окончательно, предварительно прологарифмировав, получаем:

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Введем в рассмотрение величину, которая называется **отношение правдоподобия** $l(X) = f(X|\omega_1)/f(X|\omega_2)$:

$$l(X) = \frac{f(X|\omega_1)}{f(X|\omega_2)} \geq \frac{P(\omega_2)}{P(\omega_1)} \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$

Окончательно, предварительно прологарифмировав, получаем:

Решающее правило, минимизирующее ошибку решения.

$$-\ln l(X) = -\ln f(X|\omega_1) + \ln f(X|\omega_2) \geq \ln \left(\frac{P(\omega_1)}{P(\omega_2)} \right) \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$



Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

Введем в рассмотрение величину, которая называется **отношение правдоподобия** $l(X) = f(X|\omega_1)/f(X|\omega_2)$:

$$l(X) = \frac{f(X|\omega_1)}{f(X|\omega_2)} \geq \frac{P(\omega_2)}{P(\omega_1)} \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$

Окончательно, предварительно прологарифмировав, получаем:

Решающее правило, минимизирующее ошибку решения.

$$-\ln l(X) = -\ln f(X|\omega_1) + \ln f(X|\omega_2) \geq \ln \left(\frac{P(\omega_1)}{P(\omega_2)} \right) \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$

Введем в рассмотрение области X_1 и X_2 :

- $X \in X_1$, если $P(\omega_1|X) > P(\omega_2|X)$ и X относится к классу ω_1 ,
- $X \in X_2$, если $P(\omega_1|X) < P(\omega_2|X)$ и X относится к классу ω_2 .

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

*Вероятность ошибки решения.*

$$\begin{aligned}\varepsilon &= P\{error\} = P\{error|\omega_1\}P(\omega_1) + P\{error|\omega_2\}P(\omega_2) = \\ &= P(X \in X_2|\omega_1)P(\omega_1) + P(X \in X_1|\omega_2)P(\omega_2) = \\ &= P(\omega_1) \int_{X_2} f(X|\omega_1) dX + P(\omega_2) \int_{X_1} f(X|\omega_2) dX = \\ &= P(\omega_2) + \int_{X_2} \{P(\omega_1)f(X|\omega_1) - P(\omega_2)f(X|\omega_2)\} dX = \\ &= P(\omega_1)\varepsilon_1 + P(\omega_2)\varepsilon_2\end{aligned}$$

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

*Вероятность ошибки решения.*

$$\begin{aligned}
\varepsilon &= P\{error\} = P\{error|\omega_1\}P(\omega_1) + P\{error|\omega_2\}P(\omega_2) = \\
&= P(X \in X_2|\omega_1)P(\omega_1) + P(X \in X_1|\omega_2)P(\omega_2) = \\
&= P(\omega_1) \int_{X_2} f(X|\omega_1) dX + P(\omega_2) \int_{X_1} f(X|\omega_2) dX = \\
&= P(\omega_2) + \int_{X_2} \{P(\omega_1)f(X|\omega_1) - P(\omega_2)f(X|\omega_2)\} dX = \\
&= P(\omega_1)\varepsilon_1 + P(\omega_2)\varepsilon_2
\end{aligned}$$

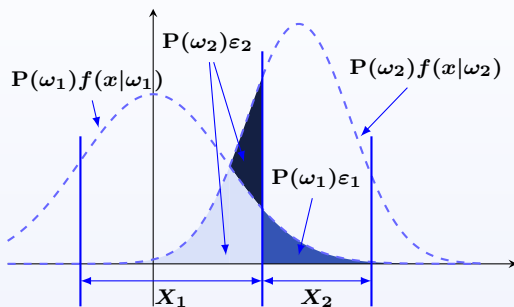
Интеграл в предпоследнем равенстве принимает наименьшее значение, когда область X_2 состоит из всех точек, где подынтегральная функция отрицательна, т. е.

$$P(\omega_1)f(X|\omega_1) < P(\omega_2)f(X|\omega_2)$$

Т. о. минимизируется вероятность принятия ошибочного решения.

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Полученный результат для **одномерного случая** иллюстрируется на рисунке. Два слагаемых в этом выражении, по существу, представляют площади, накрываемые «хвостами» функций $P(\omega_k)f(x|\omega_k)$, $k = 1, 2$. В силу произвольности выбора областей X_1 и X_2 вероятность ошибки в примере не столь мала, как могла бы быть. Видно, что смещая границу области влево, можно свести к нулю площадь темного «треугольника» и тем самым уменьшить вероятность ошибки.

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.



Два типа ошибок:

- Отклонить гипотезу H_1 , когда она верна, т. е. $X \in X_2$, в то время, когда верна гипотеза H_1 . $P(H_2|H_1)$ – **вероятность ошибки I-го типа**.
- Принять гипотезу H_1 , когда она ложная, т. е. $X \in X_1$, но гипотеза H_1 не верна. $P(H_1|H_2)$ – **вероятность ошибки II-го типа**.

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

**Два типа ошибок:**

- Отклонить гипотезу H_1 , когда она верна, т. е. $X \in X_2$, в то время, когда верна гипотеза H_1 . $P(H_2|H_1)$ – **вероятность ошибки I-го типа**.
- Принять гипотезу H_1 , когда она ложная, т. е. $X \in X_1$, но гипотеза H_1 не верна. $P(H_1|H_2)$ – **вероятность ошибки II-го типа**.

Замечание.

При вычислении вероятности ошибки решения в общем случае приходится вычислять кратные интегралы, что не всегда удобно. Поэтому иногда приходится вычислять вероятность ошибки через отношение правдоподобия, получая однократные интегралы:

$$\begin{aligned} \varepsilon &= P \left\{ l(X) < \frac{P(\omega_2)}{P(\omega_1)} \middle| \omega_1 \right\} P(\omega_1) + P \left\{ l(X) > \frac{P(\omega_2)}{P(\omega_1)} \middle| \omega_2 \right\} P(\omega_2) = \\ &= P(\omega_1) \int_0^{P(\omega_2)/P(\omega_1)} f(l|\omega_1) dl + P(\omega_2) \int_{P(\omega_2)/P(\omega_1)}^{+\infty} f(l|\omega_2) dl. \end{aligned}$$

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

**Модель двух нормальных распределений с разными ковариационными матрицами.**

Пусть X – нормальная случайная величина с вектором математического ожидания $M_i \in \mathbb{R}^2$, $i = 1, 2$ и ковариационной матрицей Σ_i , $i = 1, 2$. Тогда

$$f(X|\omega_1) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\Sigma_1|^{1/2}} \exp \left\{ -\frac{1}{2} (X - M_1)^T \Sigma_1^{-1} (X - M_1) \right\},$$
$$f(X|\omega_2) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\Sigma_2|^{1/2}} \exp \left\{ -\frac{1}{2} (X - M_2)^T \Sigma_2^{-1} (X - M_2) \right\}.$$

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

**Модель двух нормальных распределений с разными ковариационными матрицами.**

Пусть X – нормальная случайная величина с вектором математического ожидания $M_i \in \mathbb{R}^2$, $i = 1, 2$ и ковариационной матрицей Σ_i , $i = 1, 2$. Тогда

$$f(X|\omega_1) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\Sigma_1|^{1/2}} \exp \left\{ -\frac{1}{2} (X - M_1)^T \Sigma_1^{-1} (X - M_1) \right\},$$

$$f(X|\omega_2) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|\Sigma_2|^{1/2}} \exp \left\{ -\frac{1}{2} (X - M_2)^T \Sigma_2^{-1} (X - M_2) \right\}.$$

Тогда области X_1 и X_2 определяются выражением:

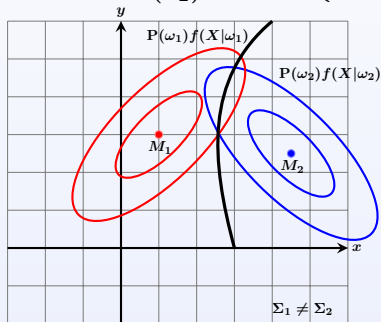
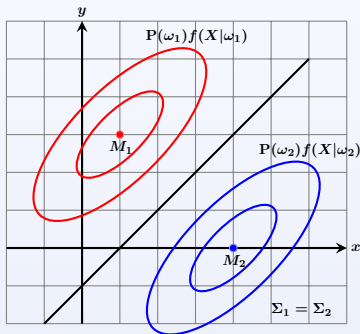
$$h(X) = \ln \frac{|\Sigma_1|}{|\Sigma_2|} + (X - M_1)^T \Sigma_1^{-1} (X - M_1) - (X - M_2)^T \Sigma_2^{-1} (X - M_2) \leq 2 \ln \frac{P(\omega_1)}{P(\omega_2)} \implies X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$

Байесовские методы.

Байесовское решающее правило, минимизирующее ошибку решения.

**Модель двух нормальных распределений с общей ковариационной матрицей (модель Фишера).**Области X_1 и X_2 определяются выражением:

$$h(X) = (M_2 - M_1)^T \Sigma^{-1} X + \frac{1}{2} (M_1^T \Sigma^{-1} M_1 - M_2^T \Sigma^{-1} M_2) \leq \ln \frac{P(\omega_1)}{P(\omega_2)} \Rightarrow X \in \begin{cases} \omega_1 \\ \omega_2 \end{cases}$$





Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

Функция потерь

$L(\theta, \hat{\theta})$ – функция потерь.

$\hat{\theta} = \hat{\theta}(\mathbf{y})$ – точечная оценка, зависящая от выборки наблюдений $\mathbf{y} = [y_1, \dots, y_n]$ (реализация случайной выборки).

θ – случайная величина, следовательно L – случайная величина.



Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

Функция потерь

$L(\theta, \hat{\theta})$ – функция потерь.

$\hat{\theta} = \hat{\theta}(y)$ – точечная оценка, зависящая от выборки наблюдений $\vec{y} = [y_1, \dots, y_n]$ (реализация случайной выборки).

θ – случайная величина, следовательно L – случайная величина.

Гипотеза ожидаемой полезности

Имеется значение θ , минимизирующее математическое ожидание функции потерь.

$$\min_{\hat{\theta}} ML(\theta, \hat{\theta}) = \min_{\hat{\theta}} \int_{R_{\theta}} L(\theta, \hat{\theta}) f(\theta|y) d\theta,$$

при этом предполагается конечность математического ожидания и существование минимума.

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



Пример №1.

Рассмотрим квадратичную функцию потерь $L = (\theta - \hat{\theta})^T C (\theta - \hat{\theta})$, где C – нестохастическая положительно-определенная симметрическая матрица.



Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

Пример №1.

Рассмотрим квадратичную функцию потерь $L = (\theta - \hat{\theta})^T C (\theta - \hat{\theta})$, где C – нестохастическая положительно-определенная симметрическая матрица.

$$\begin{aligned}
 ML &= M \left[(\theta - \hat{\theta})^T C (\theta - \hat{\theta}) \right] = \\
 &= M \left[((\theta - M\theta) - (\hat{\theta} - M\theta))^T C ((\theta - M\theta) - (\hat{\theta} - M\theta)) \right] = \\
 &= M \left[(\theta - M\theta)^T C (\theta - M\theta) \right] - M \left[(\hat{\theta} - M\theta)^T C (\theta - M\theta) \right] - \\
 &\quad - M \left[(\theta - M\theta)^T C (\hat{\theta} - M\theta) \right] + M \left[(\hat{\theta} - M\theta)^T C (\hat{\theta} - M\theta) \right] = \\
 &= \underbrace{M \left[(\theta - M\theta)^T C (\theta - M\theta) \right]}_{\text{не содержит } \hat{\theta}} + \underbrace{M \left[(\hat{\theta} - M\theta)^T C (\hat{\theta} - M\theta) \right]}_{\text{не стохастический}}.
 \end{aligned}$$



Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

Пример №1.

Рассмотрим квадратичную функцию потерь $L = (\theta - \hat{\theta})^T C (\theta - \hat{\theta})$, где C – нестохастическая положительно-определенная симметрическая матрица.

$$\begin{aligned}
 ML &= M \left[(\theta - \hat{\theta})^T C (\theta - \hat{\theta}) \right] = \\
 &= M \left[((\theta - M\theta) - (\hat{\theta} - M\theta))^T C ((\theta - M\theta) - (\hat{\theta} - M\theta)) \right] = \\
 &= M \left[(\theta - M\theta)^T C (\theta - M\theta) \right] - M \left[(\hat{\theta} - M\theta)^T C (\theta - M\theta) \right] - \\
 &\quad - M \left[(\theta - M\theta)^T C (\hat{\theta} - M\theta) \right] + M \left[(\hat{\theta} - M\theta)^T C (\hat{\theta} - M\theta) \right] = \\
 &= \underbrace{M \left[(\theta - M\theta)^T C (\theta - M\theta) \right]}_{\text{не содержит } \hat{\theta}} + \underbrace{(\hat{\theta} - M\theta)^T C (\hat{\theta} - M\theta)}_{\text{не стохастический}}.
 \end{aligned}$$

Следовательно, $\hat{\theta} = M\theta$ – оптимальная точечная оценка.



Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

Пример №2.Пусть $L = |\theta - \hat{\theta}|$ – функция потерь. $f(\theta|y)$ – ФПВ, при $a \leq \theta \leq b$, где a и b известны. Найдём точечную оценку $\hat{\theta}$ минимизирующую функцию потерь.

$$\begin{aligned}
 ML &= \int_a^b |\theta - \hat{\theta}| f(\theta|y) d\theta = \\
 &= \int_a^{\hat{\theta}} (\hat{\theta} - \theta) f(\theta|y) d\theta + \int_{\hat{\theta}}^b (\theta - \hat{\theta}) f(\theta|y) d\theta = \\
 &= \hat{\theta} F(\hat{\theta}|y) - \int_a^{\hat{\theta}} \theta f(\theta|y) d\theta + \int_{\hat{\theta}}^b \theta f(\theta|y) d\theta - \hat{\theta}(1 - F(\hat{\theta}|y)).
 \end{aligned}$$

Дифференцируем

$$\begin{aligned}
 \frac{dML}{d\hat{\theta}} &= F(\hat{\theta}|y) + \hat{\theta} f(\hat{\theta}|y) - \hat{\theta} f(\hat{\theta}|y) - \hat{\theta} f(\hat{\theta}|y) - 1 + F(\hat{\theta}|y) + \hat{\theta} f(\hat{\theta}|y) = \\
 &= F(\hat{\theta}|y) - 1 + F(\hat{\theta}|y) = 0 \iff F(\hat{\theta}|y) = \frac{1}{2}. \\
 \frac{d^2 ML}{d\hat{\theta}^2} &> 0 \text{ при } F(\hat{\theta}|y) = \frac{1}{2}
 \end{aligned}$$

Т. о. медиана является оптимальной точечной оценкой.

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

**Функция риска.**

Рассмотрим $\tilde{\theta} = \tilde{\theta}(y)$ – **оцениватель**, полученный методами теории выборочных исследований. Рассмотрим функцию, называемую **функцией риска**, связанной с оценивателем $\tilde{\theta}$:

$$r(\theta) = \int_{R_y} L(\theta, \tilde{\theta}) f(y|\theta) dy.$$

В данном случае термин «оцениватель» указывает на то, что $\tilde{\theta} = \tilde{\theta}(y)$ – случайная величина.

Будем искать оцениватель, минимизирующий средний риск, т. е.

$$Mr(\theta) = \int_{R_\theta} r(\theta) f(\theta) d\theta$$



Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

Тогда

$$\begin{aligned}
 \min_{\tilde{\theta}} Mr(\theta) &= \min_{\tilde{\theta}} \int_{R_{\theta}} \int_{R_y} f(\theta) L(\theta, \tilde{\theta}) f(y|\theta) dy d\theta = \\
 &= \min_{\tilde{\theta}} \int_{R_y} \underbrace{\left[\int_{R_{\theta}} L(\theta, \tilde{\theta}) f(\theta|y) d\theta \right]}_{\substack{\text{функция} \\ \text{апостериорного риска}}} f(y) dy
 \end{aligned}$$

Если рассматривается задача классификации, то задача оптимизация выглядит следующим образом:

$$\begin{aligned}
 \min_{\tilde{\theta}} Mr(\theta) &= \min_{\tilde{\theta}} \int_{R_y} \underbrace{\left[\sum_{\theta \in R_{\theta}} L(\theta, \tilde{\theta}) P(\theta|y) \right]}_{\substack{\text{функция} \\ \text{апостериорного риска}}} f(y) dy
 \end{aligned}$$

При условии, что $Mr(\theta)$ конечно, $\tilde{\theta}$ минимизирует апостериорный риск.

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



Общая схема задач оценивания параметра классификации.

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



Общая схема задач оценивания параметра классификации.

1. Получить (оценить, найти) распределение вероятностей имеющейся случайной величины: $F(X; \theta)$.
2. Найти априорное распределение вероятностей: $f(\theta|I_0)$. Характеризует степень неопределенности имеющейся априорной информации I_0 .
3. Задать функцию потерь $L(\theta, \hat{\theta})$. Характеризует потери, связанные с заменой истинного значения параметра θ его оценкой $\hat{\theta}$.
4. Задать (знать, получить) план испытаний, предписывающий способ получения экспериментальных данных.

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



Общая схема задач оценивания параметра классификации.

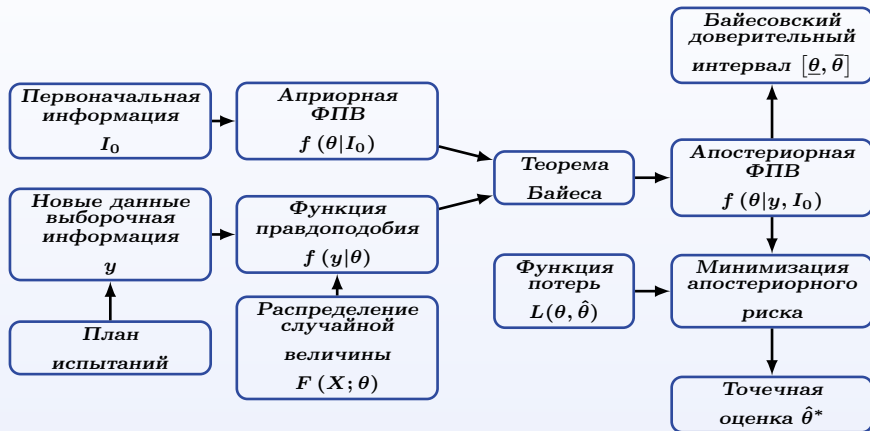
1. Получить (оценить, найти) распределение вероятностей имеющейся случайной величины: $F(X; \theta)$.
2. Найти априорное распределение вероятностей: $f(\theta|I_0)$. Характеризует степень неопределенности имеющейся априорной информации I_0 .
3. Задать функцию потерь $L(\theta, \hat{\theta})$. Характеризует потери, связанные с заменой истинного значения параметра θ его оценкой $\hat{\theta}$.
4. Задать (знать, получить) план испытаний, предписывающий способ получения экспериментальных данных.

Две формы представления оценок параметра классификации.

1. Совокупность точечной оценки $\hat{\theta}$ и среднеквадратичного отклонения $\sigma_{\hat{\theta}}$, являющегося характеристикой точности определения $\hat{\theta}$.
2. Доверительный интервал $[\underline{\theta}, \bar{\theta}]$ при заданной доверительной вероятности γ .

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.





Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.

- **Этап I.** Составление функции правдоподобия $f(y|\theta)$. Для этого используется какая-либо статистическая модель, основанная на распределении случайной величины $F(X; \theta)$ и новых данных y , полученных вследствие реализации плана испытаний.
- **Этап II.** Построение апостериорного распределения $f(\theta|y, I_0)$ с использованием формулы Байеса.

$$f(\theta|y, I_0) = \frac{f(y|\theta)f(\theta|I_0)}{\int_{R_\theta} f(y|\theta)f(\theta|I_0) d\theta}.$$

- **Этап III.** Получение байесовских оценок.

Байесовский доверительный интервал определяется условием:

$$\int_{\underline{\theta} \leq \theta \leq \bar{\theta}} f(\theta|y, I_0) d\theta = \gamma.$$

Для нахождения байесовской точечной оценки $\hat{\theta}^*$ записывают функцию апостериорного риска (в общем случае, его математическое ожидание):

$$G(\hat{\theta}) = \int_{R_\theta} L(\theta, \hat{\theta}) f(\theta|y, I_0) d\theta, \quad \hat{\theta}^* = \arg \min_{\hat{\theta} \in R_\theta} G(\hat{\theta})$$

Погрешность расчета значения $\hat{\theta}^*$ оценивается с помощью апостериорного среднеквадратичного отклонения $\sigma_{\hat{\theta}^*}$, для которого справедливо:

$$\sigma_{\hat{\theta}^*}^2 = \int_{R_\theta} \theta^2 f(\theta|y, I_0) d\theta - \hat{\theta}^{*2}.$$

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



На практике, ошибки первого и второго рода не всегда эквивалентны.

Например. При диспансеризации не поставленный диагноз при его наличии опаснее, чем ложно поставленный диагноз, при условии, что человек здоров. Тогда возникает взвешенная ошибка классификации:

$$\varepsilon = c_1 P(\omega_1) \varepsilon_1 + c_2 P(\omega_2) \varepsilon_2$$

В общем случае. Пусть за принятое решение необходимо платить штраф.

- c_{11} – штраф за решение $X \in \omega_1$, если в действительности $X \in \omega_1$.
- c_{12} – штраф за решение $X \in \omega_2$, если в действительности $X \in \omega_1$.
- c_{21} – штраф за решение $X \in \omega_1$, если в действительности $X \in \omega_2$.
- c_{22} – штраф за решение $X \in \omega_2$, если в действительности $X \in \omega_2$.

При условии $c_{12} > c_{11}$ и $c_{21} < c_{22}$.

Тогда функцию потерь можно задать с помощью наборов штрафов и записать функции апостериорного риска:

$$R(\omega_1|X) = c_{11}P(\omega_1|X) + c_{12}P(\omega_2|X),$$

$$R(\omega_2|X) = c_{21}P(\omega_1|X) + c_{22}P(\omega_2|X).$$

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



Тогда средний риск выглядит следующим образом:

$$\begin{aligned}
M[r] &= \int_{X_1 \cup X_2} \left[\sum_{i=1}^2 R(\omega_i|X) P(\omega_i|X) \right] f(X) dX = \\
&= \int_{X_1} c_{11} P(\omega_1|X) f(X) dX + \int_{X_2} c_{12} P(\omega_1|X) f(X) dX + \\
&+ \int_{X_1} c_{21} P(\omega_2|X) f(X) dX + \int_{X_2} c_{22} P(\omega_2|X) f(X) dX = \\
&= \int_{X_1} [c_{11} P(\omega_1) f(X|\omega_1) + c_{21} P(\omega_2) f(X|\omega_2)] dX + \\
&+ \int_{X_2} [c_{12} P(\omega_1) f(X|\omega_1) + c_{22} P(\omega_2) f(X|\omega_2)] dX = \\
&= c_{12} P(\omega_1) + c_{21} P(\omega_2) + \\
&+ \int_{X_1} [(c_{21} - c_{22}) P(\omega_2) f(X|\omega_2) - (c_{12} - c_{11}) P(\omega_1) f(X|\omega_1)] dX.
\end{aligned}$$

Байесовские методы.

Байесовское решающее правило, минимизирующее функцию риска.



Задача. Выбрать область X_1 для того, чтобы минимизировать средний риск.

Если подынтегральное выражение отрицательно, то средний риск можно уменьшить, отнеся X к области X_1 .

Если подынтегральное выражение положительно, то средний риск можно уменьшить, отнеся X к области X_2 .

Решающее правило (байесовский критерий, минимизирующий риск).

$$(c_{21} - c_{22})P(\omega_2)f(X|\omega_2) \leq (c_{12} - c_{11})P(\omega_1)f(X|\omega_1) \Rightarrow X \in \begin{cases} X_1 \\ X_2 \end{cases}.$$

или

$$\frac{f(X|\omega_1)}{f(X|\omega_2)} \geq \frac{(c_{21} - c_{22})P(\omega_2)}{(c_{12} - c_{11})P(\omega_1)} \Rightarrow X \in \begin{cases} X_1 \\ X_2 \end{cases}.$$

Если $c_{21} - c_{22} = c_{12} - c_{11}$, то в этом случае получаем **симметричную функцию штрафа**. И решающее правило не отличается от правила минимизирующего ошибку решения.

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Некоторые сведения о проверке гипотез.

Пусть $\mathbb{F} = \{F(x; \theta), \theta \in \Theta\}$ – класс специального функционального вида допустимых распределений наблюдаемой случайной величины ξ .

$$\theta = (\theta_1, \dots, \theta_r), \quad \Theta \subseteq \mathbb{R}^r.$$

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Некоторые сведения о проверке гипотез.

Пусть $\mathbb{F} = \{F(x; \theta), \theta \in \Theta\}$ – класс специального функционального вида допустимых распределений наблюдаемой случайной величины ξ .

$$\theta = (\theta_1, \dots, \theta_r), \quad \Theta \subseteq \mathbb{R}^r.$$

Примеры параметрических гипотез:

$$H_1 : \theta = \theta_1;$$

$$H_1 : \theta_1 = \dots = \theta_r;$$

$$H_1 : g(\theta) = \theta_1$$

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Некоторые сведения о проверке гипотез.

Пусть $\mathbb{F} = \{F(x; \theta), \theta \in \Theta\}$ – класс специального функционального вида допустимых распределений наблюдаемой случайной величины ξ .

$$\theta = (\theta_1, \dots, \theta_r), \quad \Theta \subseteq \mathbb{R}^r.$$

Примеры параметрических гипотез:

$$H_1 : \theta = \theta_1;$$

$$H_1 : \theta_1 = \dots = \theta_r;$$

$$H_1 : g(\theta) = \theta_1$$

В общем случае

$$H_1 : \theta \in \Theta_1, \quad \Theta_1 \subset \Theta;$$

$$H_2 : \theta \in \Theta_2, \quad \Theta_2 = \Theta \setminus \Theta_1;$$

H_2 – гипотеза, альтернативная гипотезе H_1 (*альтернатива*).



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Определение

В общей постановке задачи применяется некоторый **критерий** – правило, которое для каждой реализации $\mathbf{x}$ выборки $\mathbf{X}$, приводит к одному из двух решений: принять гипотезу H_1 или отклонить ее (принять альтернативу H_2).



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Определение

В общей постановке задачи применяется некоторый **критерий** – правило, которое для каждой реализации $\mathbf{x}$ выборки $\mathbf{X}$, приводит к одному из двух решений: принять гипотезу H_1 или отклонить ее (принять альтернативу H_2).

Определение

$\mathcal{X}_1$ – область принятия гипотезы H_1 ,

$\mathcal{X}_2$ – область отклонения гипотезы H_1 (принятия гипотезы H_2), принято называть **критической областью**.

Причем $\mathcal{X}_1 \cap \mathcal{X}_2 = \emptyset$ и $\mathcal{X}_1 \cup \mathcal{X}_2 = \mathcal{X}$, т. е.

$$H_1 \text{ отвергается} \iff X \in \mathcal{X}_2.$$

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Общий принцип принятия решения

Если в результате эксперимента наблюдается маловероятное событие при справедливости гипотезы H_1 , то говорят, что гипотеза H_1 не согласуется с данными. В этом случае она отвергается. В противном случае, говорят, что данные не противоречат гипотезе (согласуются с ней) и H_1 принимается.



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Общий принцип принятия решения

Если в результате эксперимента наблюдается маловероятное событие при справедливости гипотезы H_1 , то говорят, что гипотеза H_1 не согласуется с данными. В этом случае она отвергается. В противном случае, говорят, что данные не противоречат гипотезе (согласуются с ней) и H_1 принимается.

Выбор критической области

Критическая область $\mathfrak{X}_2$ должна быть выбрана таким образом, чтобы была мала вероятность

$$P(X \in \mathfrak{X}_2 | H_1),$$

– вероятность попадания значения выборки X в $\mathfrak{X}_2$ при условии, что H_1 справедлива. В этом случае можно ввести понятие **уровня значимости критерия** α . Тогда $\mathfrak{X}_{2\alpha} = \mathfrak{X}_2$

$$P\{X \in \mathfrak{X}_2 | H_1\} \leq \alpha$$



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Понятие мощности критерия

Величина

$$W(\theta) = W(\mathfrak{X}_{2\alpha}; \theta) = P_{\theta}\{X \in \mathfrak{X}_{2\alpha}\}, \quad \theta \in \Theta.$$

называется **функцией мощности критерия** $\mathfrak{X}_{2\alpha}$. Иначе, это вероятность попадания значений выборки $\mathbf{X}$ в критическую область $\mathfrak{X}_{2\alpha}$, когда истинное распределение F , которое однозначно определяется выбором истинного параметра θ .

В этом случае $W(\theta)$ при $\theta \in \Theta_1$ – вероятность ошибки I-го типа, а $1 - W(\theta)$, при $\theta \in \Theta_2$ – вероятность ошибки II-го типа.



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Понятие мощности критерия

Величина

$$W(\theta) = W(\mathfrak{X}_{2\alpha}; \theta) = P_{\theta}\{X \in \mathfrak{X}_{2\alpha}\}, \quad \theta \in \Theta.$$

называется **функцией мощности критерия** $\mathfrak{X}_{2\alpha}$. Иначе, это вероятность попадания значений выборки $\mathbf{X}$ в критическую область $\mathfrak{X}_{2\alpha}$, когда истинное распределение F , которое однозначно определяется выбором истинного параметра θ .

В этом случае $W(\theta)$ при $\theta \in \Theta_1$ – вероятность ошибки I-го типа, а $1 - W(\theta)$, при $\theta \in \Theta_2$ – вероятность ошибки II-го типа.

Основная цель!

Свести к минимуму вероятности ошибок обоих типов. НО! При заданном числе испытаний, в общем случае, нет возможности сделать одновременно обе вероятности сколь угодно малыми. Можно добиться произвольной малости вероятности одной из ошибок первого или второго типа.



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Постановка задачи.

При заданном числе испытаний n устанавливается граница для вероятности ошибки первого типа и при этом выбирается та критическая область $\mathfrak{X}_2$ для которой вероятность ошибки второго типа минимальна.

$$W(\theta) \leq \alpha, \quad \theta \in \Theta_1;$$

$$W(\theta) \rightarrow \max, \quad \theta \in \Theta_2, \quad \text{или} \quad 1 - W(\theta) \rightarrow \min, \quad \theta \in \Theta_2$$

Определение

$\mathfrak{X}_{2\alpha}$ и $\mathfrak{X}_{2\alpha}^*$ – два критерия одного и того же уровня значимости α для гипотезы H_1 .

Если

$$W(\mathfrak{X}_{2\alpha}^*; \theta) \leq W(\mathfrak{X}_{2\alpha}; \theta), \quad \theta \in \Theta_1 \quad (1)$$

$$W(\mathfrak{X}_{2\alpha}^*; \theta) \geq W(\mathfrak{X}_{2\alpha}; \theta), \quad \theta \in \Theta_2. \quad (2)$$

Строгое неравенство в (2) должно достигаться хотябы при одном значении θ . Тогда говорят, что $\mathfrak{X}_{2\alpha}^*$ **равномерно мощнее** критерия $\mathfrak{X}_{2\alpha}$. Если (1) и (2) выполняются для любого критерия $\mathfrak{X}_{2\alpha}$, то $\mathfrak{X}_{2\alpha}^*$ называют **равномерно наиболее мощным критерием** для проверки гипотезы H_1 .



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Критерий Неймана-Пирсона. Выбор из двух простых гипотез.

$$H_1 : \theta = \theta_1;$$

$$H_2 : \theta = \theta_2.$$

В этом случае допустимыми распределениями являются $F_1(x) = F_1(x; \theta_1)$ и $F_2(x) = F_2(x; \theta_2)$.

Задача! По выборке $X = (X_1, \dots, X_n)$ из некоторого распределения определить, какое из этих двух распределений истинно.

Пусть уровень значимости критерия α задан, тогда

$$W(\mathfrak{X}_{2\alpha}; \theta_2) \rightarrow \max$$

$$W(\mathfrak{X}_{2\alpha}; \theta_1) = \alpha.$$

Решение данной задачи оптимизации для нахождения критической области $\mathfrak{X}_{2\alpha}^*$ было предложено **Ю. Нейманом** и **Э. Пирсоном** в 1933 году.

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Рассмотрим случай двух абсолютно непрерывных распределений!

$F_1(x) = F_1(x; \theta_1)$ и $F_2(x) = F_2(x; \theta_2)$ – абсолютно непрерывны с ФПВ $f(x; \theta_1)$ и $f(x; \theta_2)$.

Рассмотрим
$$l(X) = \frac{f(X; \theta_2)}{f(X; \theta_1)} = \prod_{i=1}^n f_2(X_i) / \prod_{i=1}^n f_1(X_i).$$

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Рассмотрим случай двух абсолютно непрерывных распределений!

$F_1(x) = F_1(x; \theta_1)$ и $F_2(x) = F_2(x; \theta_2)$ – абсолютно непрерывны с ФПВ $f(x; \theta_1)$ и $f(x; \theta_2)$.

Рассмотрим
$$l(X) = \frac{f(X; \theta_2)}{f(X; \theta_1)} = \prod_{i=1}^n f_2(X_i) / \prod_{i=1}^n f_1(X_i).$$

Заметим, что в предыдущих рассуждениях отношение правдоподобия было введено как $l(X) = f(X; \theta_1)/f(X; \theta_2)$. Но принципиального значения это не окажет на дальнейшие рассуждения.

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Рассмотрим случай двух абсолютно непрерывных распределений!

$F_1(x) = F_1(x; \theta_1)$ и $F_2(x) = F_2(x; \theta_2)$ – абсолютно непрерывны с ФПВ $f(x; \theta_1)$ и $f(x; \theta_2)$.

Рассмотрим $l(X) = \frac{f(X; \theta_2)}{f(X; \theta_1)} = \prod_{i=1}^n f_2(X_i) / \prod_{i=1}^n f_1(X_i)$.

Заметим, что в предыдущих рассуждениях отношение правдоподобия было введено как $l(X) = f(X; \theta_1)/f(X; \theta_2)$. Но принципиального значения это не окажет на дальнейшие рассуждения.

Введем в рассмотрение функцию

$$\psi(c) = P\{l(X) \geq c | H_1\} = P_{\theta_1}\{l(X) \geq c\}.$$

- $\psi(0) = 1$;
- функция $\psi(c) \leq 1/c$ – убывающая функция и $\psi(c) \rightarrow 0$ при $c \rightarrow \infty$.

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Рассмотрим случай двух абсолютно непрерывных распределений!

$F_1(x) = F_1(x; \theta_1)$ и $F_2(x) = F_2(x; \theta_2)$ – абсолютно непрерывны с ФПВ $f(x; \theta_1)$ и $f(x; \theta_2)$.

Рассмотрим
$$l(X) = \frac{f(X; \theta_2)}{f(X; \theta_1)} = \prod_{i=1}^n f_2(X_i) / \prod_{i=1}^n f_1(X_i).$$

Заметим, что в предыдущих рассуждениях отношение правдоподобия было введено как $l(X) = f(X; \theta_1)/f(X; \theta_2)$. Но принципиального значения это не окажет на дальнейшие рассуждения.

Введем в рассмотрение функцию

$$\psi(c) = P\{l(X) \geq c | H_1\} = P_{\theta_1}\{l(X) \geq c\}.$$

- $\psi(0) = 1$;
- функция $\psi(c) \leq 1/c$ – убывающая функция и $\psi(c) \rightarrow 0$ при $c \rightarrow \infty$.

$$\begin{aligned}
 1 \geq P_{\theta_2}\{l(X) \geq c\} &= \int_{x, l(X) \geq c} f(x; \theta_2) dx \geq c \int_{x, l(X) \geq c} f(x; \theta_1) dx = \\
 &= c P_{\theta_1}\{l(X) \geq c\} = c\psi(c)
 \end{aligned}$$



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Будем полагать, что $\exists c = c_\alpha$ такое, что $\psi(c) = \alpha$.

Теорема

При сделанных предположениях наиболее мощный критерий проверки гипотезы H_1 существует и имеет вид:

$$\mathfrak{X}_{2\alpha}^* = \{x : l(x) \geq c_\alpha\}$$

Доказательство.

Пусть $\mathfrak{X}_{2\alpha}$ – произвольный критерий мощности α . Тогда

$$W(\mathfrak{X}_{2\alpha}; \theta_2) = \int_{\mathfrak{X}_{2\alpha}} f(x; \theta_2) dx = \int_{\mathfrak{X}_{2\alpha} \cap \mathfrak{X}_{2\alpha}^*} f(x; \theta_2) dx + \int_{\mathfrak{X}_{2\alpha} \cap \overline{\mathfrak{X}_{2\alpha}^*}} f(x; \theta_2) dx,$$

$$W(\mathfrak{X}_{2\alpha}^*; \theta_2) = \int_{\mathfrak{X}_{2\alpha}^*} f(x; \theta_2) dx = \int_{\mathfrak{X}_{2\alpha} \cap \mathfrak{X}_{2\alpha}^*} f(x; \theta_2) dx + \int_{\overline{\mathfrak{X}_{2\alpha}} \cap \mathfrak{X}_{2\alpha}^*} f(x; \theta_2) dx.$$



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Доказательство.

$$\begin{aligned}
 W(\mathfrak{x}_{2\alpha}; \theta_2) &= W(\mathfrak{x}_{2\alpha}^*; \theta_2) + \int_{\mathfrak{x}_{2\alpha} \bar{\mathfrak{x}}_{2\alpha}^*} f(x; \theta_2) dx - \int_{\bar{\mathfrak{x}}_{2\alpha} \mathfrak{x}_{2\alpha}^*} f(x; \theta_2) dx = \\
 &= W(\mathfrak{x}_{2\alpha}^*; \theta_2) + \left\{ \underbrace{\int_{\mathfrak{x}_{2\alpha} \bar{\mathfrak{x}}_{2\alpha}^*} l(x) f(x; \theta_1) dx}_{l(x) < c_\alpha} - \underbrace{\int_{\bar{\mathfrak{x}}_{2\alpha} \mathfrak{x}_{2\alpha}^*} l(x) f(x; \theta_1) dx}_{l(x) \geq c_\alpha} \right\} < \\
 &< W(\mathfrak{x}_{2\alpha}^*; \theta_2) + c_\alpha \left\{ \int_{\mathfrak{x}_{2\alpha} \bar{\mathfrak{x}}_{2\alpha}^*} f(x; \theta_1) dx - \int_{\bar{\mathfrak{x}}_{2\alpha} \mathfrak{x}_{2\alpha}^*} f(x; \theta_1) dx \right\}.
 \end{aligned}$$

Так как $\int_{\mathfrak{x}_{2\alpha}} f(x; \theta_1) dx = \alpha$ и $\int_{\mathfrak{x}_{2\alpha}^*} f(x; \theta_1) dx = \alpha$, то каждый из интегралов в скобках отличается от α на величину $\int_{\mathfrak{x}_{2\alpha} \bar{\mathfrak{x}}_{2\alpha}^*} f(x; \theta_1) dx$. Следовательно оба интеграла в скобках равны. И требуемое неравенство получено:

$$W(\mathfrak{x}_{2\alpha}; \theta_2) < W(\mathfrak{x}_{2\alpha}^*; \theta_2)$$





Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Вопрос о вычислении критической границы c_α в данном курсе мы поднимать не будем, ровно как и доказывать несмещенность критерия Неймана-Пирсона.

Возвращаясь к байесовскому подходу классификации «в свете» критерия Неймана-Пирсона.

В предыдущих обозначениях:

$$\varepsilon_1 = P \{H_2|H_1\} = \int_{X_2} f(X|\omega_1) dX,$$

$$\varepsilon_2 = P \{H_1|H_2\} = \int_{X_1} f(X|\omega_2) dX$$

Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



Вопрос о вычислении критической границы c_α в данном курсе мы поднимать не будем, ровно как и доказывать несмещенность критерия Неймана-Пирсона.

Возвращаясь к байесовскому подходу классификации «в свете» критерия Неймана-Пирсона.

В предыдущих обозначениях:

$$\varepsilon_1 = P \{H_2|H_1\} = \int_{X_2} f(X|\omega_1) dX,$$
$$\varepsilon_2 = P \{H_1|H_2\} = \int_{X_1} f(X|\omega_2) dX$$

При рассмотрении ошибок первого и второго типа нет разницы которую из них мы будем минимизировать, а для которой будем задавать ограничения (уровень значимости).



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Вопрос о вычислении критической границы c_α в данном курсе мы поднимать не будем, ровно как и доказывать несмещенность критерия Неймана-Пирсона.

Возвращаясь к байесовскому подходу классификации «в свете» критерия Неймана-Пирсона.

В предыдущих обозначениях:

$$\varepsilon_1 = P \{H_2|H_1\} = \int_{X_2} f(X|\omega_1) dX,$$

$$\varepsilon_2 = P \{H_1|H_2\} = \int_{X_1} f(X|\omega_2) dX$$

При рассмотрении ошибок первого и второго типа нет разницы которую из них мы будем минимизировать, а для которой будем задавать ограничения (уровень значимости).

Ненарушая общности, запишем постановку задачи

$$\varepsilon_1 \rightarrow \min,$$

$$\varepsilon_2 = \varepsilon_0.$$



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Полученная задача – классическая задача на нахождение условного экстремума. Запишем функцию Лагранжа:

$$L = \varepsilon_1 + \mu(\varepsilon_2 - \varepsilon_0) = \int_{X_2} f(X|\omega_1) dX + \mu \left(\int_{X_1} f(X|\omega_2) dX - \varepsilon_0 \right) = \\ = (1 - \mu\varepsilon_0) + \int_{X_1} \{\mu f(X|\omega_2) - f(X|\omega_1)\} dX.$$

Рассуждая как и в предыдущих пунктах, значение L можно уменьшить, путем выбора областей X_1 и X_2 следующим образом:

$$\mu f(X|\omega_2) \leq f(X|\omega_1) \iff \frac{f(X|\omega_1)}{f(X|\omega_2)} \geq \mu \text{ при } X \in \begin{cases} X_1 \\ X_2 \end{cases}.$$

Вывод. Сравнивая полученный результат с предыдущими, можно сделать вывод, что критерий Неймана-Пирсона не дает какого-то нового решающего правила, а основан на критерии отношения правдоподобия. **НО!** Критерий отношения правдоподобия является критерием, минимизирующим вероятность ошибки решения и, согласно полученным результатам, минимизирующим вероятность ошибки решения для одного класса, в то время как для другого класса вероятность ошибки остается неизменной!



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Пример. Задано двумерное нормальное распределение:

$$M_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Sigma_1 = \Sigma_2 = \Sigma = E_2.$$

Тогда, решающую границу можно записать в виде:

$$h(X) = (M_2 - M_1)^T \Sigma^{-1} X + 1/2 \left(M_1^T \Sigma^{-1} M_1 - M_2^T \Sigma^{-1} M_2 \right) \leq \ln \mu.$$

$$2x_1 \leq \ln \mu.$$

При заданной вероятности ошибки ϵ_0 порог μ можно определить решив уравнение (в общем виде решать данное не будем, рассматривать будем только частные случаи):

$$\epsilon_2 = \int_{X_1} f(X|\omega_2) dX = \epsilon_0.$$

Или используя отношение правдоподобия

$$\epsilon_2 = \int_{\mu}^{\infty} f(l|\omega_2) dl = \epsilon_0.$$

Вывод. Так как $f(l|\omega_2) \geq 0$, то вероятность ошибки второго рода монотонно убывает при увеличении порога μ .



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.

Для демонстрации результата рассмотрим подробно случай, когда пороговое значение $\mu = 2$. Тогда решающая граница

$$2x_1 = \ln 2 \Rightarrow x_1 = -0.34$$

Вычислим уровень значимости нашего критерия, то есть значение ε_2 при $\mu = 2$:

$$\begin{aligned} \varepsilon_2 &= \int_{X_1} f(X|\omega_2) dX = \frac{1}{2\pi} \iint_{X_1} \exp \left\{ -\frac{1}{2} [(x_1 - 1)^2 + x_2^2] \right\} dx_1 dx_2 = \\ &= \frac{1}{2\pi} \int_{-\infty}^{-0.34} \exp \left\{ -\frac{(x_1 - 1)^2}{2} \right\} dx_1 \int_{-\infty}^{\infty} \exp \left\{ -\frac{x_2^2}{2} \right\} dx_2. \end{aligned}$$

Второй интеграл – является интегралом Пуассона: $\int_{-\infty}^{\infty} \exp \{ -x_2^2/2 \} dx_2 = \sqrt{2\pi}$

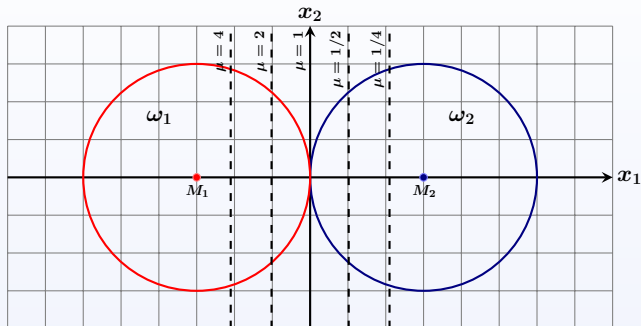
Первый интеграл можно свести к интегралу ошибок, или к интегралу Лапласа. Табличные значения этих интегралов есть в учебниках по математической статистике и теории вероятностей. Тогда

$$\varepsilon_2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-1.34} \exp \left\{ -\frac{\xi^2}{2} \right\} d\xi = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\pi}}{2} \sqrt{2} - 0.40988 \simeq 0.09$$



Байесовские методы.

Байесовское решающее правило. Критерий Неймана-Пирсона.



μ	4	2	1	1/2	1/4
ε_2	0.04	0.09	0.16	0.25	0.38