

Bayesian Decision

Jin Young Choi

Seoul National University

Outline

- Bayes Formula
 - Priori probability
 - Likelihood
 - Posterior Probability
 - Bayes Decision
- Risk Formulation
 - Conditional Risk
 - Likelihood Ratio Test
 - Zero-one Loss Function (Bayes Decision)
- Maximum Likelihood Estimation (MLE)
- Error Probability

Bayesian Decision

- Question:
 - There live two kinds of fishes in a lake: tuna or salmon.
 - If you catch a fish by fishing, is the fish likely to be tuna or salmon?

Bayesian Decision

- We have experienced that salmon has been caught in 70% and tuna in 30%.
- What is the next fish likely to be?

Bayesian Decision

- If other types of fish are irrelevant:

$$p(\omega = \omega_1) + p(\omega = \omega_2) = 1,$$

ω is random variable, ω_1 and ω_2 denote salmon and tuna.

- Probabilities reflect our prior knowledge obtained from past experience.
- **Simple Decision Rule:**
 - Make a decision without seeing the fish.
 - Decide ω_1 if $p(\omega = \omega_1) > p(\omega = \omega_2)$
 ω_2 otherwise.

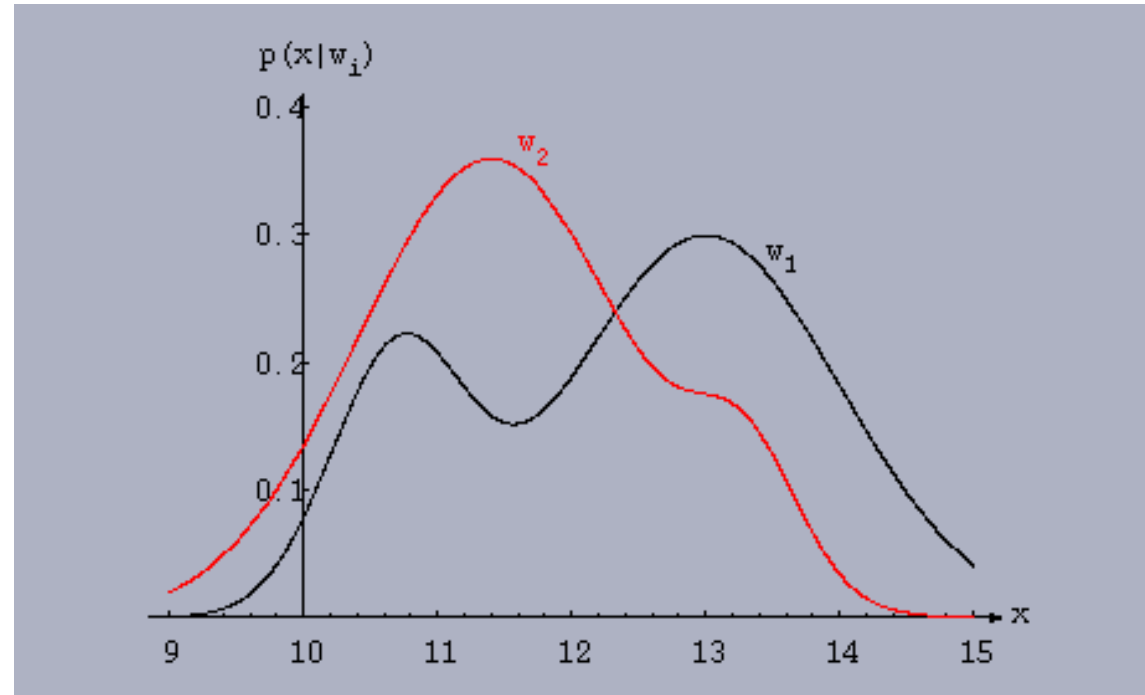
Bayesian Decision

- In general, we will have some features and more information.
- Feature: lightness measurement = x
 - Different fish yields different lightness readings (x is a random variable)

Bayesian Decision

- Define
 - $p(x|\omega_i)$ = Class Conditional Probability Density
 - The difference between $p(x|\omega_1)$ and $p(x|\omega_2)$ describes the difference in lightness between tuna and salmon.

Bayesian Decision



- Hypothetical class-conditional probability
- Density functions are normalized (area under each curve is 1.0)

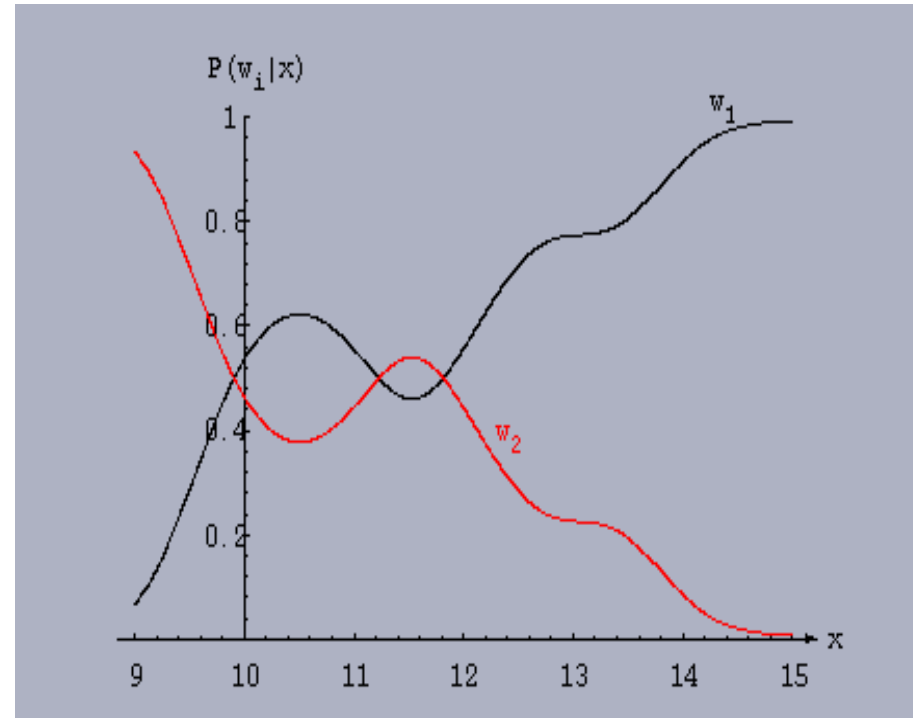
Bayesian Decision

- Suppose that we know
 - The prior probabilities $p(\omega_1)$ and $p(\omega_2)$
 - The conditional densities $p(x|\omega_1)$ and $p(x|\omega_2)$
 - Measure lightness of a fish $= x$
- What is the category of the fish with lightness of x ?
- The probability that the fish has category of ω_i is $p(\omega_i|x)$.

Bayes formula

- $p(\omega_i|x) = \frac{p(x|\omega_i)p(\omega_i)}{p(x)}$,
where $p(x) = \sum_j p(x, \omega_j) = \sum_j p(x|\omega_j)p(\omega_j)$.
- $Posterior = \frac{Likelihood * Prior}{Evidence}$
- $p(x|\omega_i)$ is called the **likelihood** of ω_i with respect to x .
 - The ω_i category for which $p(x|\omega_i)$ is large is more "likely" to be the true category
- $p(x)$ is the **evidence**
 - How frequently is a pattern with feature value x observed.
 - Scale factor that the posterior probabilities sum to 1.

Bayes formula



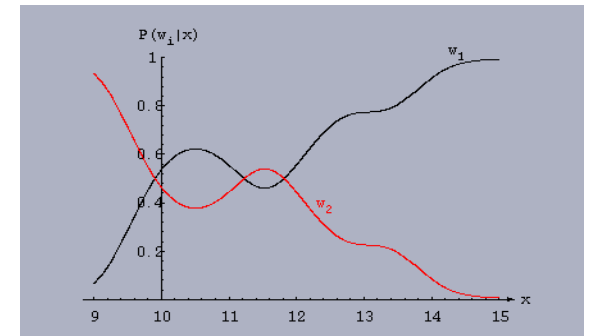
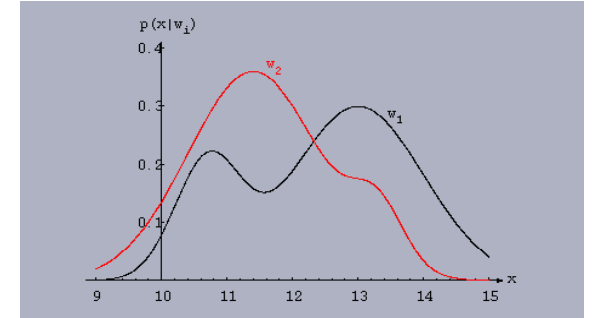
- Posterior probabilities for the particular priors $p(\omega_1) = 2/3$ and $p(\omega_2) = 1/3$. At every x the posteriors sum to 1.

Bayes Decision Rule (Minimal probability error)

- Likelihood Decision:
 - ω_1 : if $p(x|\omega_1) > p(x|\omega_2)$
 - ω_2 : otherwise
- Posteriori Decision:
 - ω_1 : if $p(x|\omega_1)p(\omega_1) > p(x|\omega_2)p(\omega_2)$
 - ω_2 : otherwise
- Decision Error Probability
 - $p(\text{error}|x) = \min(p(\omega_1|x), p(\omega_2|x))$

where the decision error is given by

$$p(\text{error}|x) = \begin{cases} p(\omega_2|x) & \text{if we decide } \omega_1 \text{ for } \omega_2 \\ p(\omega_1|x) & \text{if we decide } \omega_2 \text{ for } \omega_1 \end{cases}$$



Exercise

- 지금까지 샌디에고 만에서 잡힌 연어의 20%가 40cm 이하였고, 잡힌 농어의 30%가 40cm 이하였다. 또한 연어와 송어의 잡힌 비율은 7:3이었다. 잡힌 물고기의 크기가 40cm 이하인데, 연어와 송어 둘 중 하나로 보인다. 연어인지, 농어인지 판단해 보시오.

Exercise

- 지금까지 샌디에고 만에서 잡힌 연어의 20%가 40cm 이하였고, 잡힌 농어의 30%가 40cm 이하였다. 또한 연어와 송어의 잡힌 비율은 7:3이었다. 잡힌 물고기의 크기가 40cm 이하인데, 연어와 송어 둘 중 하나로 보인다. 연어인지, 농어인지 판단해 보시오.
- Sol.
 - ✓ (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

Exercise

- 지금까지 샌디에고 만에서 잡힌 연어의 20%가 40cm 이하였고, 잡힌 농어의 30%가 40cm 이하였다. 또한 연어와 송어의 잡힌 비율은 7:3이었다. 잡힌 물고기의 크기가 40cm 이하인데, 연어와 송어 둘 중 하나로 보인다. 연어인지, 농어인지 판단해 보시오.
- Sol.
 - ✓ (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.
 - ✓ 연어: $X = 0$, 농어: $X = 1$, 크기: Y .
 - ✓ $P(Y \leq 40cm | X = 0) = 0.2$, $P(Y \leq 40cm | X = 1) = 0.3$, $P(X = 0) = 0.7$, $P(X = 1) = 0.3$

Exercise

- 지금까지 샌디에고 만에서 잡힌 연어의 20%가 40cm 이하였고, 잡힌 농어의 30%가 40cm 이하였다. 또한 연어와 송어의 잡힌 비율은 7:3이었다. 잡힌 물고기의 크기가 40cm 이하인데, 연어와 송어 둘 중 하나로 보인다. 연어인지, 농어인지 판단해 보시오.
- Sol.
 - ✓ (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.
 - ✓ 연어: $X = 0$, 농어: $X = 1$, 크기: Y .
 - ✓ $P(Y \leq 40cm | X = 0) = 0.2$, $P(Y \leq 40cm | X = 1) = 0.3$, $P(X = 0) = 0.7$, $P(X = 1) = 0.3$
 - ✓ 질문: posteriori: $P(X = 0 | Y \leq 40cm) = ?$, $P(X = 1 | Y \leq 40cm) = ?$

$$p(\omega_i | x) = \frac{p(x | \omega_i) p(\omega_i)}{p(x)}$$

Exercise

- 지금까지 샌디에고 만에서 잡힌 연어의 20%가 40cm 이하였고, 잡힌 농어의 30%가 40cm 이하였다. 또한 연어와 송어의 잡힌 비율은 7:3이었다. 잡힌 물고기의 크기가 40cm 이하인데, 연어와 송어 둘 중 하나로 보인다. 연어인지, 농어인지 판단해 보시오.
- Sol.
 - ✓ (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.
 - ✓ 연어: $X = 0$, 농어: $X = 1$, 크기: Y .
 - ✓ $P(Y \leq 40cm | X = 0) = 0.2$, $P(Y \leq 40cm | X = 1) = 0.3$, $P(X = 0) = 0.7$, $P(X = 1) = 0.3$
 - ✓ 질문: posteriori: $P(X = 0 | Y \leq 40cm) = ?$, $P(X = 1 | Y \leq 40cm) = ?$
 - ✓
$$P(X = 0 | Y \leq 40cm) = \frac{P(Y \leq 40cm | X = 0)P(X = 0)}{P(Y \leq 40cm)} = \frac{0.2 \times 0.7}{0.2 \times 0.7 + 0.3 \times 0.4} = 0.54$$

Exercise

- 지금까지 샌디에고 만에서 잡힌 연어의 20%가 40cm 이하였고, 잡힌 농어의 30%가 40cm 이하였다. 또한 연어와 송어의 잡힌 비율은 7:3이었다. 잡힌 물고기의 크기가 40cm 이하인데, 연어와 송어 둘 중 하나로 보인다. 연어인지, 농어인지 판단해 보시오.
- Sol.
 - ✓ (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.
 - ✓ 연어: $X = 0$, 농어: $X = 1$, 크기: Y .
 - ✓ $P(Y \leq 40cm | X = 0) = 0.2$, $P(Y \leq 40cm | X = 1) = 0.3$, $P(X = 0) = 0.7$, $P(X = 1) = 0.3$
 - ✓ 질문: posteriori: $P(X = 0 | Y \leq 40cm) = ?$, $P(X = 1 | Y \leq 40cm) = ?$
 - ✓
$$P(X = 0 | Y \leq 40cm) = \frac{P(Y \leq 40cm | X=0)P(X=0)}{P(Y \leq 40cm)} = \frac{0.2 \times 0.7}{0.2 \times 0.7 + 0.3 \times 0.3} = 0.54$$
 - ✓
$$P(X = 1 | Y \leq 40cm) = \frac{P(Y \leq 40cm | X=1)P(X=1)}{P(Y \leq 40cm)} = \frac{0.3 \times 0.3}{0.2 \times 0.7 + 0.3 \times 0.3} = 0.46$$
 - ✓ Bayes decision 에 의해 연어라고 판단한다.

General Formulation

- Let $\{\omega_1, \dots, \omega_c\}$ be the finite set of c categories.
- Let $\{\alpha_1, \dots, \alpha_a\}$ be the finite set of a possible actions.
Ex. Action α_i = deciding that the true state is ω_i or others.
- The risk function $\lambda(\alpha_i|\omega_j)$ = risk incurred for taking action when the state of nature is ω_j .
- x = d –dimensional feature vector (random variable)
- $p(x|\omega_i)$ = likelihood probability density function for x for given ω_i
- $p(\omega_i)$ = prior probability that nature is in state ω_i .

Conditional Risk

- After the observation, the expected risk (conditional risk) is given by

$$R(\alpha_i|x) = \sum_{j=1}^c \lambda(\alpha_i|\omega_j)p(\omega_j|x)$$

- The decision action $\alpha(x)$ for given x is given

$$\alpha(x) = \arg \min_{\alpha_i} R(\alpha_i|x) = \sum_{j=1}^c \lambda(\alpha_i|\omega_j)p(\omega_j|x)$$

Two-Category Classification

- Action α_1 = deciding that the true state is ω_1
- Action α_2 = deciding that the true state is ω_2
- Let $\lambda_{ij} = \lambda(\alpha_i|\omega_j)$ be the risk incurred for deciding ω_i when true state is ω_j .
- *The conditional risks:*
$$R(\alpha_1|x) = \lambda_{11}p(\omega_1|x) + \lambda_{12}p(\omega_2|x)$$
$$R(\alpha_2|x) = \lambda_{21}p(\omega_1|x) + \lambda_{22}p(\omega_2|x)$$
- Decide ω_1 if $R(\alpha_1|x) < R(\alpha_2|x)$
 - or if $(\lambda_{21} - \lambda_{11})p(\omega_1|x) > (\lambda_{12} - \lambda_{22})p(\omega_2|x)$
 - or if $(\lambda_{21} - \lambda_{11})p(x|\omega_1)p(\omega_1) > (\lambda_{12} - \lambda_{22})p(x|\omega_2)p(\omega_2)$and ω_2 , otherwise

Two-Category Likelihood Ratio Test

- Under reasonable assumption that $\lambda_{12} > \lambda_{22}$ and $\lambda_{21} > \lambda_{11}$, (why?)
decide ω_1 if $\frac{p(x|\omega_1)}{p(x|\omega_2)} > \frac{(\lambda_{12}-\lambda_{22})p(\omega_2)}{(\lambda_{21}-\lambda_{11})p(\omega_1)} = T$
and ω_2 , otherwise.
- The ratio $\frac{p(x|\omega_1)}{p(x|\omega_2)}$ is called the *likelihood ratio*.
- We can decide ω_1 if the likelihood ratio exceeds a threshold T value that is independent of the observation x .

Minimum-Error-Rate Classification

- To give an **equal cost** to all errors, we define **zero-one risk function** as

$$\lambda(\alpha_i|\omega_j) = \begin{cases} 0, & i = j \\ 1, & i \neq j \end{cases}, \quad \text{for } i, j = 1, \dots, C$$

- The conditional risk **representing error rate** is

$$\begin{aligned} R(\alpha_i|x) &= \sum_{j=1}^C \lambda(\alpha_i|\omega_j) p(\omega_j|x) \\ &= \sum_{j \neq i}^C p(\omega_j|x) = 1 - p(\omega_i|x) \end{aligned}$$

- To **minimize** $R(\alpha_i|x)$, we **maximize** $p(\omega_i|x)$
Decide ω_i if $p(\omega_i|x) > p(\omega_j|x)$, for all $j \neq i$
(same as Bayes' decision rule)

Maximum Likelihood Estimation (MLE)

- The samples are i.i.d (pair die).

$$j^{th} \text{ class set } D_j = \{x_l | (x_l, \bar{\omega}_l) \in S_j\}, \quad S_j \subseteq S = \{(x_l, \bar{\omega}_l) | l = 1, \dots, N\}$$

- **Maximum likelihood estimation:** find $\hat{\theta}(D)$ to maximize $p(x|D)$

$$p(x|\omega_j) \approx p(x|D_j) \approx p(x|\hat{\theta}(D_j)), \quad \hat{\theta}(D_j) = \arg \max_{\theta} p(D_j|\theta)$$

- The i.i.d. assumption implies that

$$p(D_j|\theta_j) = \prod_{x \in D_j} p(x|\theta_j)$$

- Let D be a generic sample set of size $n = |D|$

- **Log-likelihood function:**

$$l(\theta; D) \equiv \ln p(D|\theta) = \sum_{k=1}^n \ln p(x_k|\theta)$$

$$\nabla_{\theta} l(\theta; D) = 0$$

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암진단은 피검사를 하여 암표지자 값을 보고 판단을 한다. 이 암표지자를 확률 변수 x 로 했을 때, 정상인과 암환자 모두 다음의 매개변수 θ 로 표현되는 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 의 형태를 가지고 있다. 샘플은 i.i.d. 특성을 만족한다. 정상인의 암표지자 값을 평균하면 0.01이 되고 암환자의 암표지자를 평균하면 0.1이 된다. 정상인과 암환자의 분포를 가장 잘 나타내는 $\hat{\theta}_1, \hat{\theta}_2$ 를 MLE 방법으로 추정하시오.

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암진단은 피검사를 하여 암표지자 값을 보고 판단을 한다. 이 암표지자를 확률 변수 X 로 했을 때, 정상인과 암환자 모두 다음의 매개변수 θ 로 표현되는 확률 분포는 $p(X = x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 의 형태를 가지고 각 샘플은 i.i.d. 라고 가정한다. 정상인의 암표지자 값을 평균하면 0.01이 되고 암환자의 암표지자를 평균하면 0.1이 된다. 정상인과 암환자의 분포를 가장 잘 나타내는 $\hat{\theta}_1, \hat{\theta}_2$ 를 MLE 방법으로 추정하시오.

- **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암진단은 피검사를 하여 암표지자 값을 보고 판단을 한다. 이 암표지자를 확률 변수 x 로 했을 때, 정상인과 암환자 모두 다음의 매개변수 θ 로 표현되는 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 의 형태를 가지고 각 샘플은 i.i.d. 라고 가정한다. 정상인의 암표지자 값을 평균하면 0.01이 되고 암환자의 암표지자를 평균하면 0.1이 된다. 정상인과 암환자의 분포를 가장 잘 나타내는 $\hat{\theta}_1, \hat{\theta}_2$ 를 MLE 방법으로 추정하시오.

- Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.
 - ✓ 정상인의 샘플 $n = 99000$ 개, 즉 $D_i = \{x_i | i = 1, \dots, n\}$ 를 가지고 MLE를 수행하자. Likelihood function 은 아래와 같이 정의한다.

$$p(D_i|\theta) = \prod_{i=1}^n \theta x_i e^{-\theta x_i}$$

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암진단은 피검사를 하여 암표지자 값을 보고 판단을 한다. 이 암표지자를 확률 변수 x 로 했을 때, 정상인과 암환자 모두 다음의 매개변수 θ 로 표현되는 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 의 형태를 가지고 각 샘플은 i.i.d. 라고 가정한다. 정상인의 암표지자 값을 평균하면 0.01이 되고 암환자의 암표지자를 평균하면 0.1이 된다. 정상인과 암환자의 분포를 가장 잘 나타내는 $\hat{\theta}_1, \hat{\theta}_2$ 를 MLE 방법으로 추정하시오.

■ **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

- ✓ 정상인의 샘플 $n = 99000$ 개, 즉 $D_i = \{x_i | i = 1, \dots, n\}$ 를 가지고 MLE를 수행하자. Likelihood function 은 아래와 같이 정의한다.

$$p(D_i|\theta) = \prod_{i=1}^n \theta x_i e^{-\theta x_i}$$

- ✓ 양변에 log를 위하여 log-likelihood 를 구하면 다음과 같다.

$$l(\theta) = n \log \theta + \sum_{i=1}^n \log x_i - \theta \sum_{i=1}^n x_i$$

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암진단은 피검사를 하여 암표지자 값을 보고 판단을 한다. 이 암표지자를 확률 변수 x 로 했을 때, 정상인과 암환자 모두 다음의 매개변수 θ 로 표현되는 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 의 형태를 가지고 각 샘플은 i.i.d. 라고 가정한다. 정상인의 암표지자 값을 평균하면 0.01이 되고 암환자의 암표지자를 평균하면 0.1이 된다. 정상인과 암환자의 분포를 가장 잘 나타내는 $\hat{\theta}_1, \hat{\theta}_2$ 를 MLE 방법으로 추정하시오.

■ **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

- ✓ 정상인의 샘플 $n = 99000$ 개, 즉 $D_i = \{x_i | i = 1, \dots, n\}$ 를 가지고 MLE를 수행하자. Likelihood function 은 아래와 같이 정의한다.

$$p(D_i|\theta) = \prod_{i=1}^n \theta x_i e^{-\theta x_i}$$

- ✓ 양변에 log를 위하여 log-likelihood 를 구하면 다음과 같다.

$$l(\theta) = n \log \theta + -\theta \sum_{i=1}^n x_i + \sum_{i=1}^n \log x_i$$

- ✓ 양변에 θ 에 대해 미분하여 그 값이 0이 되도록 θ 를 구하면

$$\frac{d}{d\theta} l(\theta) = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0, \quad \rightarrow \quad \hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^n x_i}$$

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암진단은 피검사를 하여 암표지자 값을 보고 판단을 한다. 이 암표지자를 확률 변수 x 로 했을 때, 정상인과 암환자 모두 다음의 매개변수 θ 로 표현되는 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 의 형태를 가지고 각 샘플은 i.i.d. 라고 가정한다. 정상인의 암표지자 값을 평균하면 0.01이 되고 암환자의 암표지자를 평균하면 0.1이 된다. 정상인과 암환자의 분포를 가장 잘 나타내는 $\hat{\theta}_1, \hat{\theta}_2$ 를 MLE 방법으로 추정하시오.

■ **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

- ✓ 정상인의 샘플 $n = 99000$ 개, 즉 $D_i = \{x_i | i = 1, \dots, n\}$ 를 가지고 MLE를 수행하자. Likelihood function 은 아래와 같이 정의한다.

$$p(D_i|\theta) = \prod_{i=1}^n \theta x_i e^{-\theta x_i}$$

- ✓ 양변에 log를 위하여 log-likelihood 를 구하면 다음과 같다.

$$l(\theta) = n \log \theta + -\theta \sum_{i=1}^n x_i + \sum_{i=1}^n \log x_i$$

- ✓ 양변에 θ 에 대해 미분하여 그 값이 0이 되도록 θ 를 구하면

$$\frac{d}{d\theta} l(\theta) = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0, \quad \rightarrow \quad \hat{\theta}_{MLE} = \frac{n}{\sum_{i=1}^n x_i}$$

- ✓ 여기서 정상인의 경우 암표지자 평균이 0.01 이므로 정상인 분포의 $\hat{\theta}_1$ 은 $\hat{\theta}_1 = 100$ 이 되고 암환자의 경우는 암표지자 평균이 0.1이므로 암환자 분포의 $\hat{\theta}_2$ 는 $\hat{\theta}_2 = 10$ 이 된다.

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암표지자 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 에서 환자의 검사결과로부터 θ 를 추정하였더니 정상인인 경우 $\hat{\theta}_1 = 100$, 암환자의 경우 $\hat{\theta}_2 = 10$ 으로 추정이 되었다. 암진단을 받으러 온 사람의 검사결과 $x = 0.06$ 으로 나왔다. 정상인을 암환자로 잘못 진단하였을 때 리스크를 1로 하고, 암환자를 정상인으로 잘못 진단 하였을 때 리스크를 10으로 설정 하였다. 정확히 진단하였을 때 리스크는 0으로 한다. 이 리스크를 감안하여 암환자 여부를 진단하시오.

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암표지자 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 에서 환자의 검사결과로부터 θ 를 추정하였더니 정상인인 경우 $\hat{\theta}_1 = 100$, 암환자의 경우 $\hat{\theta}_2 = 10$ 으로 추정이 되었다. 암진단을 받으러 온 사람의 검사결과 $x = 0.06$ 으로 나왔다. 정상인을 암환자로 잘못 진단하였을 때 리스크를 1로 하고, 암환자를 정상인으로 잘못 진단 하였을 때 리스크를 10으로 설정 하였다. 정확히 진단하였을 때 리스크는 0으로 한다. 이 리스크를 감안하여 암환자 여부를 진단하시오.
- **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암표지자 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 에서 환자의 검사결과로부터 θ 를 추정하였더니 정상인인 경우 $\hat{\theta}_1 = 100$, 암환자의 경우 $\hat{\theta}_2 = 10$ 으로 추정이 되었다. 암진단을 받으러 온 사람의 검사결과 $x = 0.06$ 으로 나왔다. 정상인을 암환자로 잘못 진단하였을 때 리스크를 1로 하고, 암환자를 정상인으로 잘못 진단 하였을 때 리스크를 10으로 설정 하였다. 정확히 진단하였을 때 리스크는 0으로 한다. 이 리스크를 감안하여 암환자 여부를 진단하시오.

■ **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

- ✓ 정상인 확률은 $p(\hat{\theta}_1) = 0.99$ 이고 암환자의 확률은 $p(\hat{\theta}_2) = 0.01$ 이다.

$$\text{질문: } R(\alpha_1|x = 0.06) = \lambda_{11}p(\hat{\theta}_1|x = 0.06) + \lambda_{12}p(\hat{\theta}_2|x = 0.06) = ?$$

$$R(\alpha_2|x = 0.06) = \lambda_{21}p(\hat{\theta}_1|x = 0.06) + \lambda_{22}p(\hat{\theta}_2|x = 0.06) = ?$$

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암표지자 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 에서 환자의 검사결과로부터 θ 를 추정하였더니 정상인인 경우 $\hat{\theta}_1 = 100$, 암환자의 경우 $\hat{\theta}_2 = 10$ 으로 추정이 되었다. 암진단을 받으러 온 사람의 검사결과 $x = 0.06$ 으로 나왔다. 정상인을 암환자로 잘못 진단하였을 때 리스크를 1로 하고, 암환자를 정상인으로 잘못 진단 하였을 때 리스크를 10으로 설정 하였다. 정확히 진단하였을 때 리스크는 0으로 한다. 이 리스크를 감안하여 암환자 여부를 진단하시오.

■ Sol. (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

- ✓ 정상인 확률은 $p(\hat{\theta}_1) = 0.99$ 이고 암환자의 확률은 $p(\hat{\theta}_2) = 0.01$ 이다.

$$\text{질문: } R(\alpha_1|x = 0.06) = \lambda_{11}p(\hat{\theta}_1|x = 0.06) + \lambda_{12}p(\hat{\theta}_2|x = 0.06) = ?$$

$$R(\alpha_2|x = 0.06) = \lambda_{21}p(\hat{\theta}_1|x = 0.06) + \lambda_{22}p(\hat{\theta}_2|x = 0.06) = ?$$

- ✓ $p(\hat{\theta}_1|x = 0.06) \propto p(x = 0.06|\hat{\theta}_1)p(\hat{\theta}_1) = 100 * 0.06e^{-100*0.06} * 0.99 = 0.0147$

- ✓ $p(\hat{\theta}_2|x = 0.06) \propto p(x = 0.06|\hat{\theta}_2)p(\hat{\theta}_2) = 10 * 0.06e^{-10*0.06} * 0.01 = 0.00329$

Exercise

- 서울대학교병원에 암진단을 받으러 온 사람은 10만명이다. 그중 1000명이 암환자로 판명이 난다. 암표지자 확률 분포는 $p(x|\theta) = \theta x e^{-\theta x}$, for $x > 0$ and $\theta > 0$ 에서 환자의 검사결과로부터 θ 를 추정하였더니 정상인인 경우 $\hat{\theta}_1 = 100$, 암환자의 경우 $\hat{\theta}_2 = 10$ 으로 추정이 되었다. 암진단을 받으러 온 사람의 검사결과 $x = 0.06$ 으로 나왔다. 정상인을 암환자로 잘못 진단하였을 때 리스크를 1로 하고, 암환자를 정상인으로 잘못 진단 하였을 때 리스크를 10으로 설정 하였다. 정확히 진단하였을 때 리스크는 0으로 한다. 이 리스크를 감안하여 암환자 여부를 진단하시오.

■ **Sol.** (힌트) 문장의 수치에 해당하는 내용과 질문을 수식으로 표현해 보세요.

- ✓ 정상인 확률은 $p(\hat{\theta}_1) = 0.99$ 이고 암환자의 확률은 $p(\hat{\theta}_2) = 0.01$ 이다.

$$\text{질문: } R(\alpha_1|x = 0.06) = \lambda_{11}p(\hat{\theta}_1|x = 0.06) + \lambda_{12}p(\hat{\theta}_2|x = 0.06) = ?$$

$$R(\alpha_2|x = 0.06) = \lambda_{21}p(\hat{\theta}_1|x = 0.06) + \lambda_{22}p(\hat{\theta}_2|x = 0.06) = ?$$

- ✓ 정상인 확률은 $p(\hat{\theta}_1) = 0.99$ 이고 암환자의 확률은 $p(\hat{\theta}_2) = 0.01$ 이다.
- ✓ $p(\hat{\theta}_1|x = 0.06) \propto p(x = 0.06|\hat{\theta}_1)p(\hat{\theta}_1) = 100 * 0.06e^{-100*0.06} * 0.99 = 0.0147$
- ✓ $p(\hat{\theta}_2|x = 0.06) \propto p(x = 0.06|\hat{\theta}_2)p(\hat{\theta}_2) = 10 * 0.06e^{-10*0.06} * 0.01 = 0.00329$
- ✓ $R(\alpha_1|x = 0.06) = 0 * 0.0147 + 10 * 0.00329 = 0.0329$
 $R(\alpha_2|x = 0.06) = 1 * 0.0147 + 0 * 0.00329 = 0.0147$

Error Probabilities and Integrals

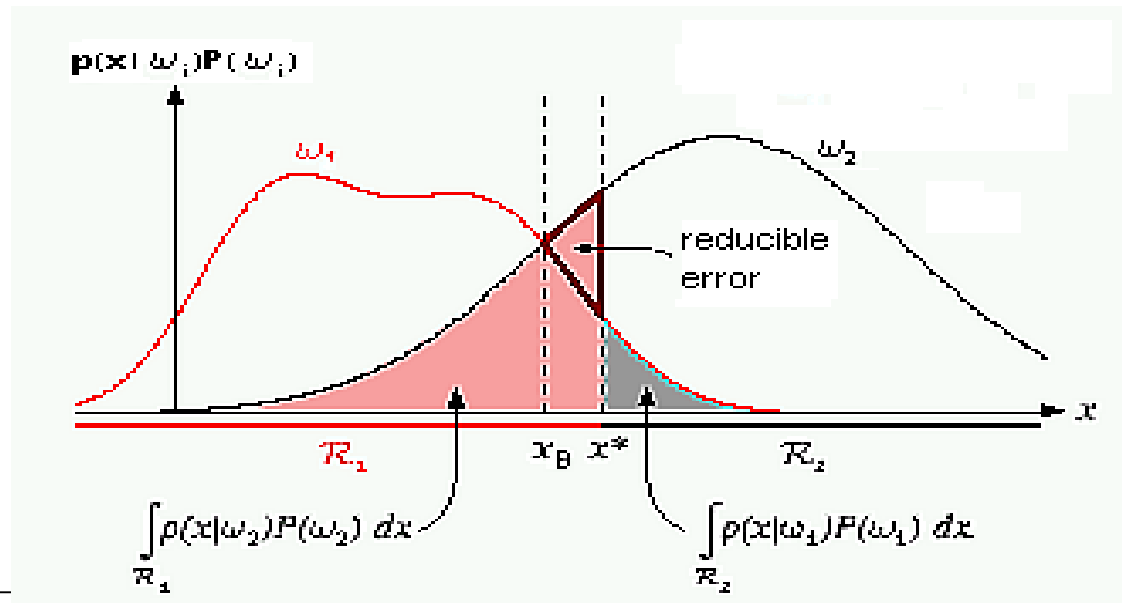
- Consider the 2-class problem and suppose that the feature space is divided into 2 regions R_1 and R_2 . There are 2 ways in which a classification error can occur.
 - An observation x falls in R_2 , and the true state is ω_1 .
 - An observation x falls in R_1 , and the true state is ω_2 .
- The error probability

$$\begin{aligned} P(\text{error}) &= P(x \in R_2 | \omega_1) p(\omega_1) + P(x \in R_1 | \omega_2) p(\omega_2) \\ &= \int_{R_2} p(x | \omega_1) p(\omega_1) dx + \int_{R_1} p(x | \omega_2) p(\omega_2) dx \end{aligned}$$

Error Probabilities and Integrals

- Because x^* is chosen arbitrarily, the probability of error is not as small as it might be.
- x_B = Bayes optimal decision boundary, and gives the lowest probability of error.
- Bayes classifier maximizes the correct probability.

$$P(\text{correct}) = \sum_{i=1}^C P(\mathbf{x} \in \mathcal{R}_i | \omega_i) p(\omega_i) = \sum_{i=1}^C \int_{\mathcal{R}_i} p(\mathbf{x} | \omega_i) p(\omega_i) d\mathbf{x}$$



Summary

- Bayes Formula
 - Priori probability
 - Likelihood
 - Posterior Probability
 - Bayes Decision
- Risk Formulation
 - Conditional Risk
 - Likelihood Ratio Test
 - Zero-one Loss Function (Bayes Decision)
- Maximum Likelihood Estimation (MLE)
- Error Probability