

作业3：贝叶斯决策与（非）参数估计

本次作业deadline为2019.4.9。编程问题请提供源代码，作业有任何问题请及时联系助教。

本次作业负责助教：张威

- 1. 贝叶斯估计 (20)
- 2. Parzen窗 (20)
- 3. 最大似然估计 (20)
- 4. 非参数估计与贝叶斯决策 (40)

1. 贝叶斯估计 (20)

请阅读《模式识别》49页，完成以下问题。

在贝叶斯估计中，给定有限样本集合 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 的情况下，贝叶斯估计量可以通过最小化期望风险获得，即

$$\theta^* = \arg \min_{\hat{\theta}} R(\hat{\theta}|\mathbf{X}) = \int_{\theta} \lambda(\hat{\theta}, \theta) p(\theta|\mathbf{X}) d\theta$$

其中， $\lambda(\hat{\theta}, \theta)$ 是定义的损失函数，当损失函数为平方误差损失函数，即 $\lambda(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$ 时，请证明，在给定样本集 \mathbf{X} 下， θ 的贝叶斯估计是下式。

$$\theta^* = E[\theta|\mathbf{X}] = \int_{\theta} \theta p(\theta|\mathbf{X}) d\theta$$

2. Parzen窗 (20)

请阅读《模式识别》56~58页，学习Parzen窗方法，完成以下问题。

假设 $P(x) \sim U(0, a)$ ，其中 $U(0, a)$ 指从0到a的均匀分布。假设Parzen窗的定义如下

$$\varphi(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

请回答以下问题。

(1) 证明使用以上Parzen窗对 $P(x)$ 估计的期望服从下式。

$$\bar{P}(x) = E[\hat{P}(x)] = \begin{cases} 0 & x < 0 \\ \frac{1}{a}(1 - e^{-\frac{x}{h}}) & 0 \leq x \leq a \\ \frac{1}{a}(e^{\frac{a}{h}} - 1)e^{-\frac{x}{h}} & a < x \end{cases}$$

- (2) 在一幅图中绘制出当 $a = 1, h = 1, \frac{1}{4}, \frac{1}{16}$ 时, $\bar{P}(x)$ 关于 x 的图像。
- (3) 当 $0 < x < a$ 时, h 需要满足什么条件时, 保证99%的 x 对应的 $\bar{P}(x)$ 的偏差才能保持在1%以内?
- (4) 当 $a = 1$ 时, 请给出 h 并绘制出在 $0 \leq x \leq 0.05$ 的 $\bar{P}(x)$ 。

3. 最大似然估计 (20)

- 运用最大似然估计的知识, 完成以下问题。
- (1) 令 $x_1, x_2, ..., x_n$ 是从正态分布 $N(\mu, \sigma^2)$ 下的抽样, 假设 μ 未知, σ^2 已知, 求 μ 的最大似然估计。如果加上条件 $\mu > \mu_0$, 请给出 μ 的最大似然估计。
 - (2) 请从正态分布 $N(0, 1)$ 中分别抽取10, 100, 1000个样本, 假设均值与方差均未知, 利用最大似然估计得到正态分布假设下的模型参数, 分别重复三次实验, 将同一抽样量下的三次重复实验估计的概率密度分布曲线绘制在一张图片内, 并与正态分布 $N(0, 1)$ 的概率密度分布曲线比较。
 - (3) 请从均匀分布 $U(0, 1)$ 中抽取100个样本, 假设均值与方差均未知, 利用最大似然估计正态分布假设下的模型参数, 绘制出估计得到的概率密度分布曲线图, 并与均匀分布 $U(0, 1)$ 的概率密度分布曲线图比较。
 - (4) 通过上述实验, 讨论模型的选择、样本量对参数估计的影响。

4. 非参数估计与贝叶斯决策 (40)

- 请生成600个样本数据, 其中350个的数据点采样于 $N(-2.5, 2)$, 记为负样本; 剩余250的数据采样于 $N(2.5, 1)$, 记为正样本。随机取出70%的数据作为训练集, 30%的数据作为测试集, 完成以下问题。
- (1) 利用高斯核 $Parzen$ 窗, 使用训练集中的数据对正样本和负样本分别进行非参数估计。
 - (2) 利用(1)的非参数估计的概率密度, 请使用最小错误率的贝叶斯决策对测试集样本进行预测, 给出测试集的错误率。
 - (3) 利用(1)中非参数估计的概率密度, 请使用最小风险的贝叶斯决策对测试集样本进行预测, 惩罚矩阵如下。

真实值 \ 预测值	正样本	负样本
	正样本	负样本
正样本	0	10
负样本	1	0

- (4) 请随机划分训练集和测试集样本三次, 重复上述步骤, 比较最小风险下的预测结果与最小错误率下的预测结果有什么不同。