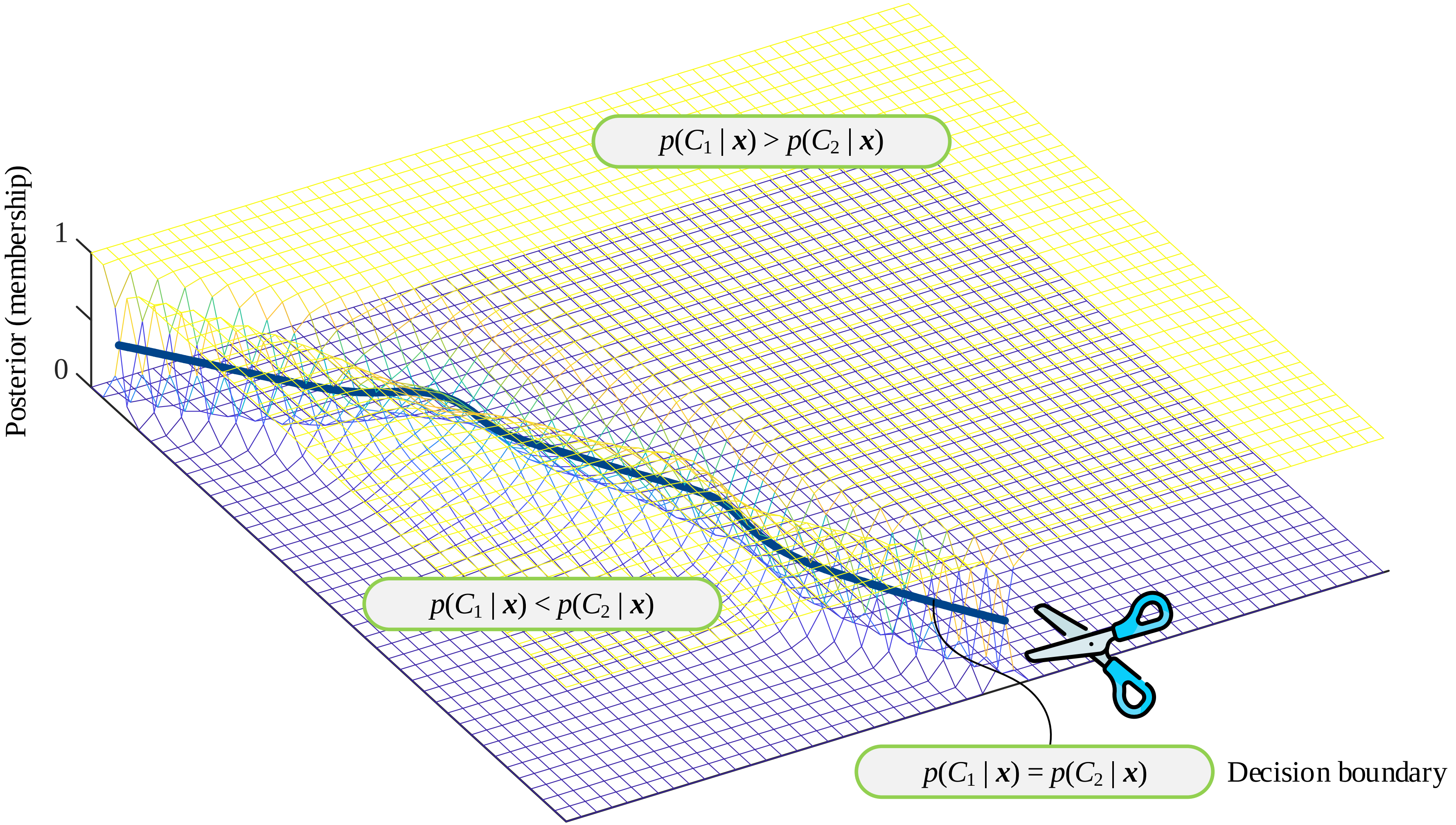
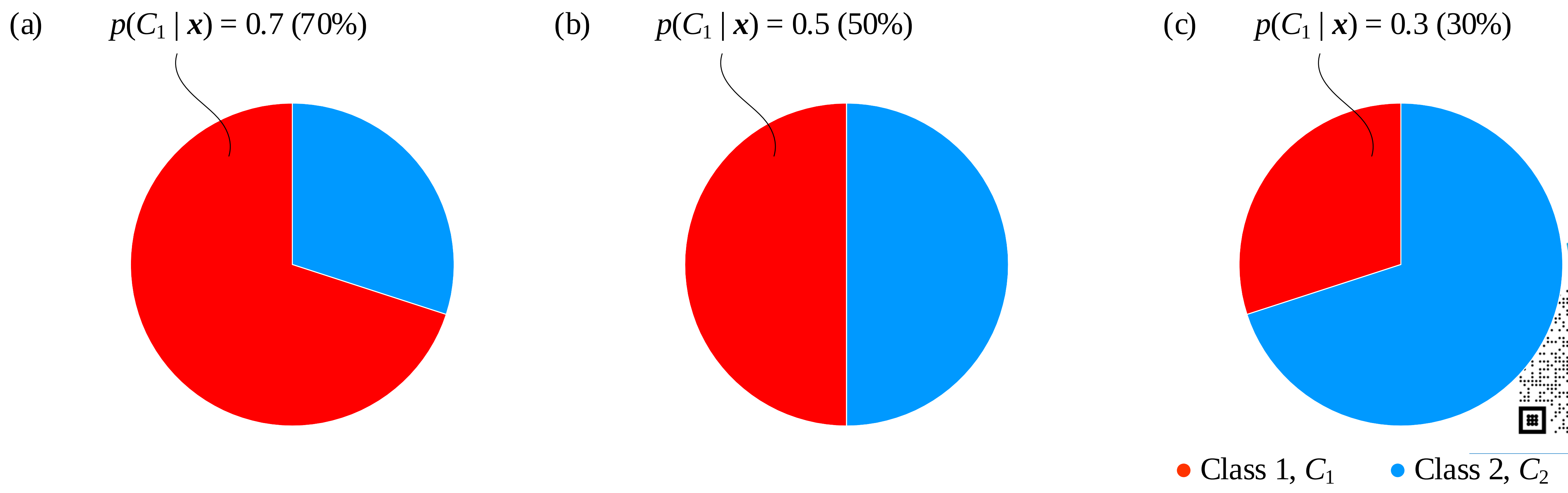
二分类，比较后验概率大小



bilibili

知乎

成员值 (membership score)

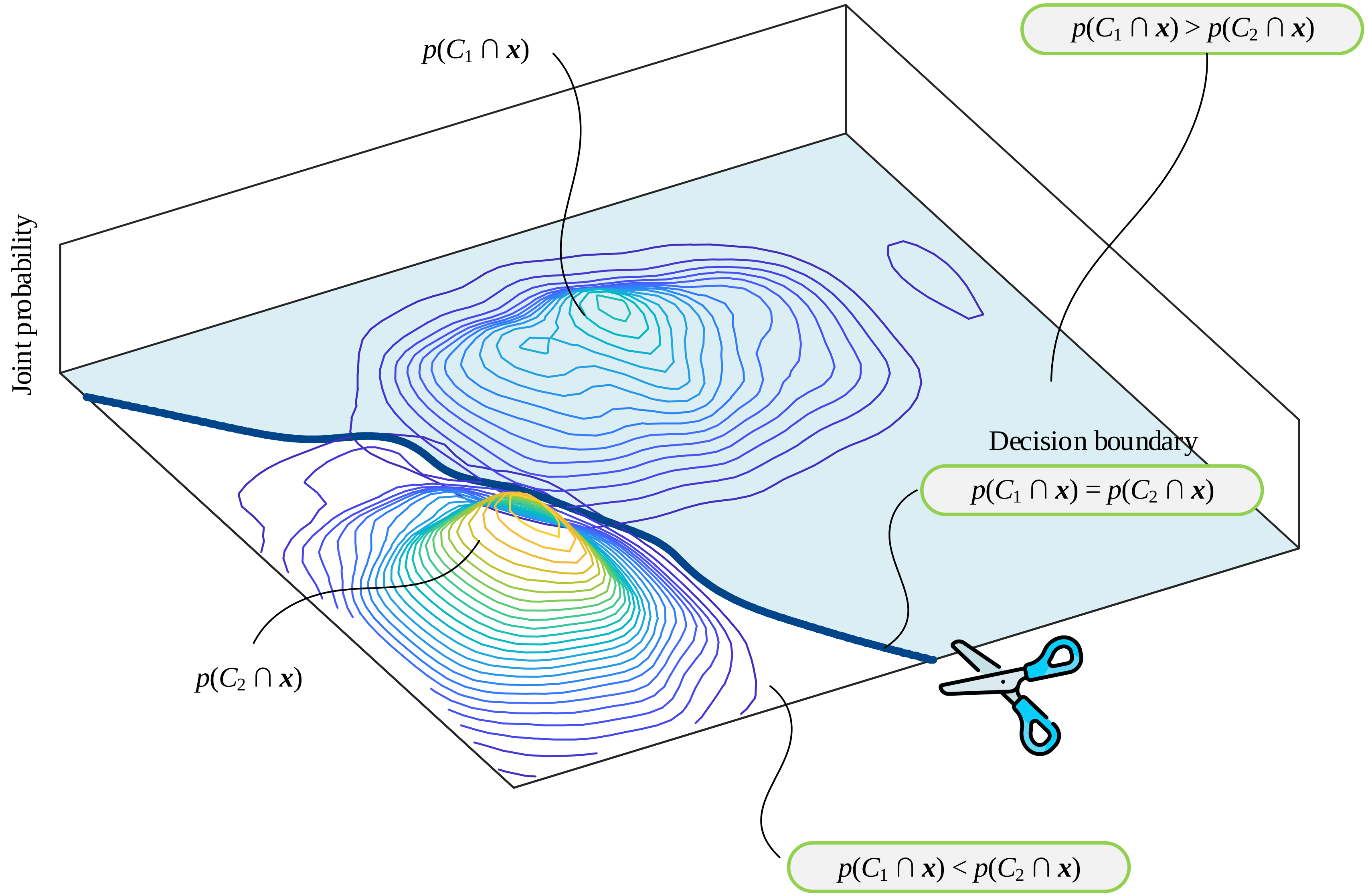


后验概率和联合概率

$$\begin{cases} \underbrace{p(C_1|\boldsymbol{x})}_{\text{Posterior}} = \dfrac{\overbrace{p(\boldsymbol{x} \cap C_1)}^{\text{Joint}}}{p(\boldsymbol{x})} \\ \underbrace{p(C_2|\boldsymbol{x})}_{\text{Posterior}} = \dfrac{\overbrace{p(\boldsymbol{x} \cap C_2)}^{\text{Joint}}}{p(\boldsymbol{x})} \end{cases}$$

$$\begin{cases} \underbrace{p(C_1|\boldsymbol{x})}_{\text{Posterior}} \propto \underbrace{p(\boldsymbol{x} \cap C_1)}_{\text{Joint}} \\ \underbrace{p(C_2|\boldsymbol{x})}_{\text{Posterior}} \propto \underbrace{p(\boldsymbol{x} \cap C_2)}_{\text{Joint}} \end{cases}$$

二分类，比较联合概率大小



朴素贝叶斯分类原理：最大化后验概率

$$\hat{y} = \underset{C_k}{\arg\max}\, p\left(C_k \middle| \mathbf{x}\right)$$

bilibili

知乎

“最大化后验概率”，等价于“最大化联合概率”

$$p\left(C_k \mid \boldsymbol{x}\right) \propto p\left(C_k \cap \boldsymbol{x}\right)$$

$$\hat{y} = \underset{C_k}{\arg\max}\, p\left(C_k \cap \boldsymbol{x}\right)$$

bilibili

知乎

回顾贝叶斯定理

$$\underbrace{p\left(C_k \mid \boldsymbol{x}\right)}_{\text{Posterior}} = \frac{\overbrace{p\left(\boldsymbol{x} \cap C_k\right)}^{\text{Joint}}}{p\left(\boldsymbol{x}\right)} = \frac{\overbrace{p\left(\boldsymbol{x} \mid C_k\right)}^{\text{Likelihood}} \overbrace{\mathrm{P}\left(C_k\right)}^{\text{Prior}}}{\underset{\text{Evidence}}{p\left(\boldsymbol{x}\right)}}$$

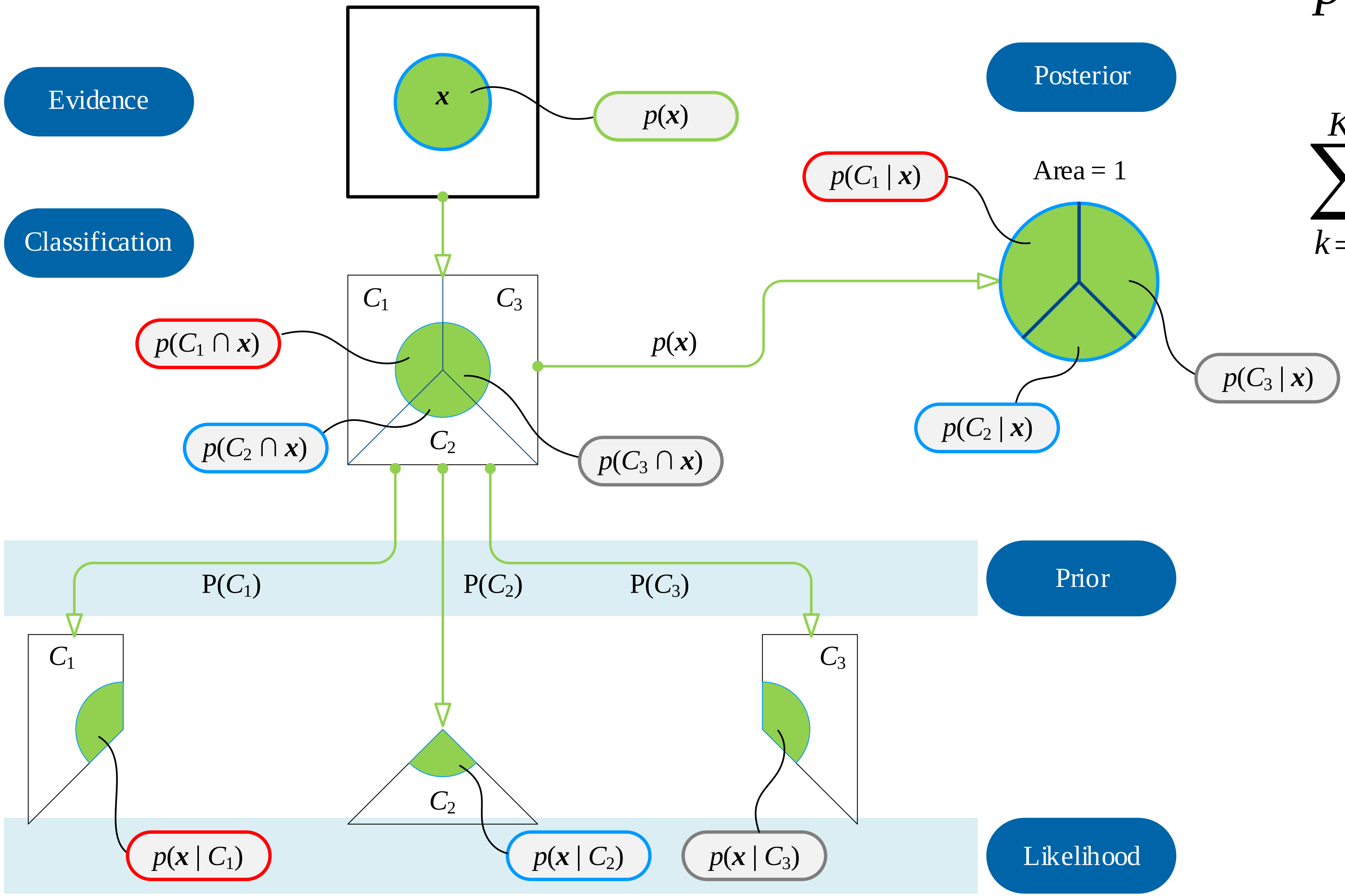
bilibili

生姜

知乎

生姜

三分类，贝叶斯定理



$$p(\boldsymbol{x}) = \sum_{k=1}^{K} \underbrace{p(C_k \cap \boldsymbol{x})}_{\text{Joint}}$$

$$\sum_{k=1}^{K} \underbrace{p(C_k | \boldsymbol{x})}_{\text{Posterior}} = 1$$

朴素之处：假设特征之间条件独立

$$\begin{cases} p(\mathbf{x}|C_1) = p(x_1|C_1)\,p(x_2|C_1) \\ p(\mathbf{x}|C_2) = p(x_1|C_2)\,p(x_2|C_2) \end{cases}$$

$$p(\mathbf{x}|C_k) = p(x_1|C_k)\,p(x_2|C_k)\ldots p(x_D|C_k) = \prod_{j=1}^{D} p(x_j|C_k)$$

特征独立

$$p(\boldsymbol{x}) = p(x_1, x_2) = p(x_1 \cap x_2) = p(x_1)\,p(x_2)$$

$$p(\boldsymbol{x}) = p(x_1)\,p(x_2)\ldots p(x_D) = \prod_{j=1}^{D} p(x_j)$$

“特征条件独立”条件下，联合概率计算式

$$p\left(C_k \cap \boldsymbol{x}\right) = p\left(\boldsymbol{x} \middle| C_k\right) \mathrm{P}\left(C_k\right) = \mathrm{P}\left(C_k\right) \prod_{j=1}^{D} p\left(x_j \middle| C_k\right)$$

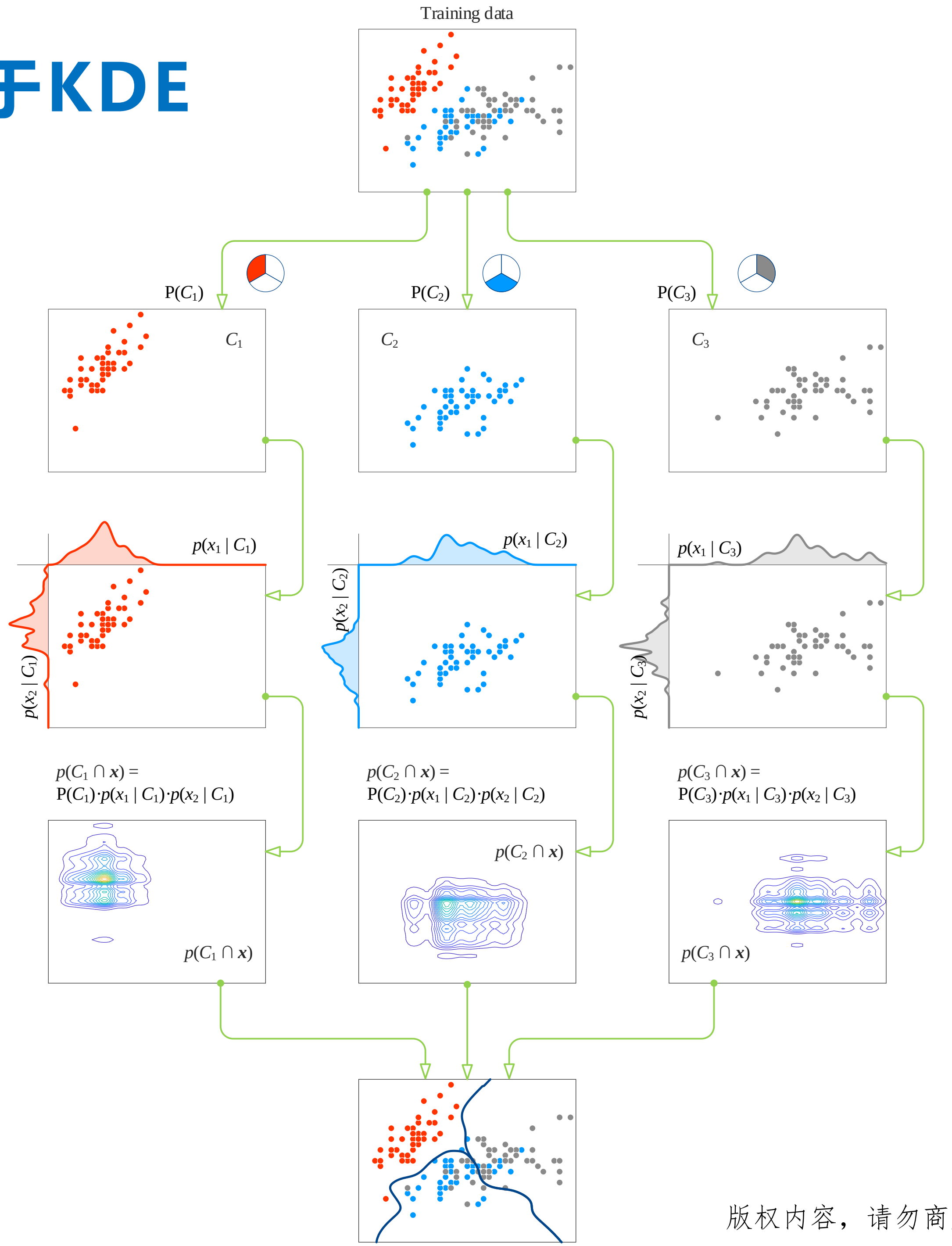
bilibili

知乎

朴素贝叶斯优化问题

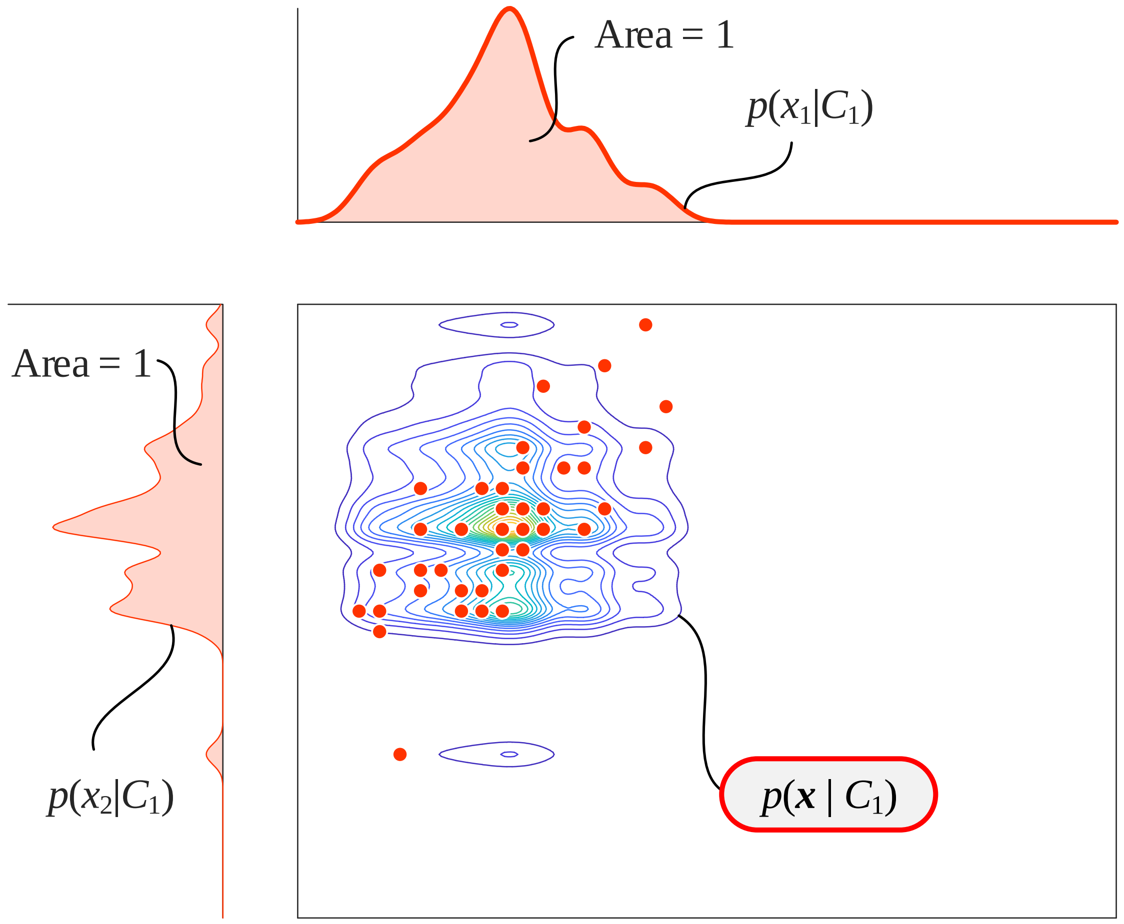
$$\hat{y} = \arg\max_{C_k} \mathrm{P}(C_k) \prod_{j=1}^{D} p(x_j | C_k)$$

朴素贝叶斯分类过程，基于KDE



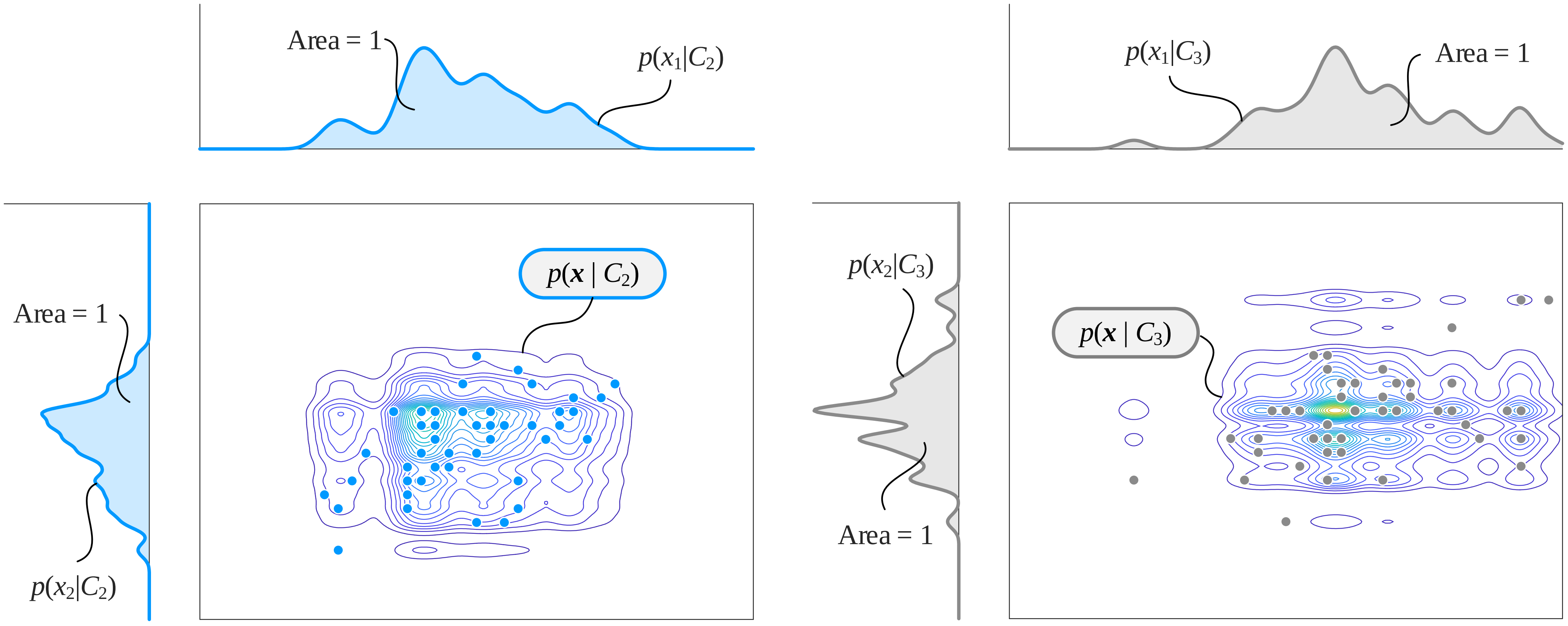
似然概率

$$p(\boldsymbol{x}|C_1) = p(x_1, x_2|C_1) = p(x_1|C_1) \cdot p(x_2|C_1)$$



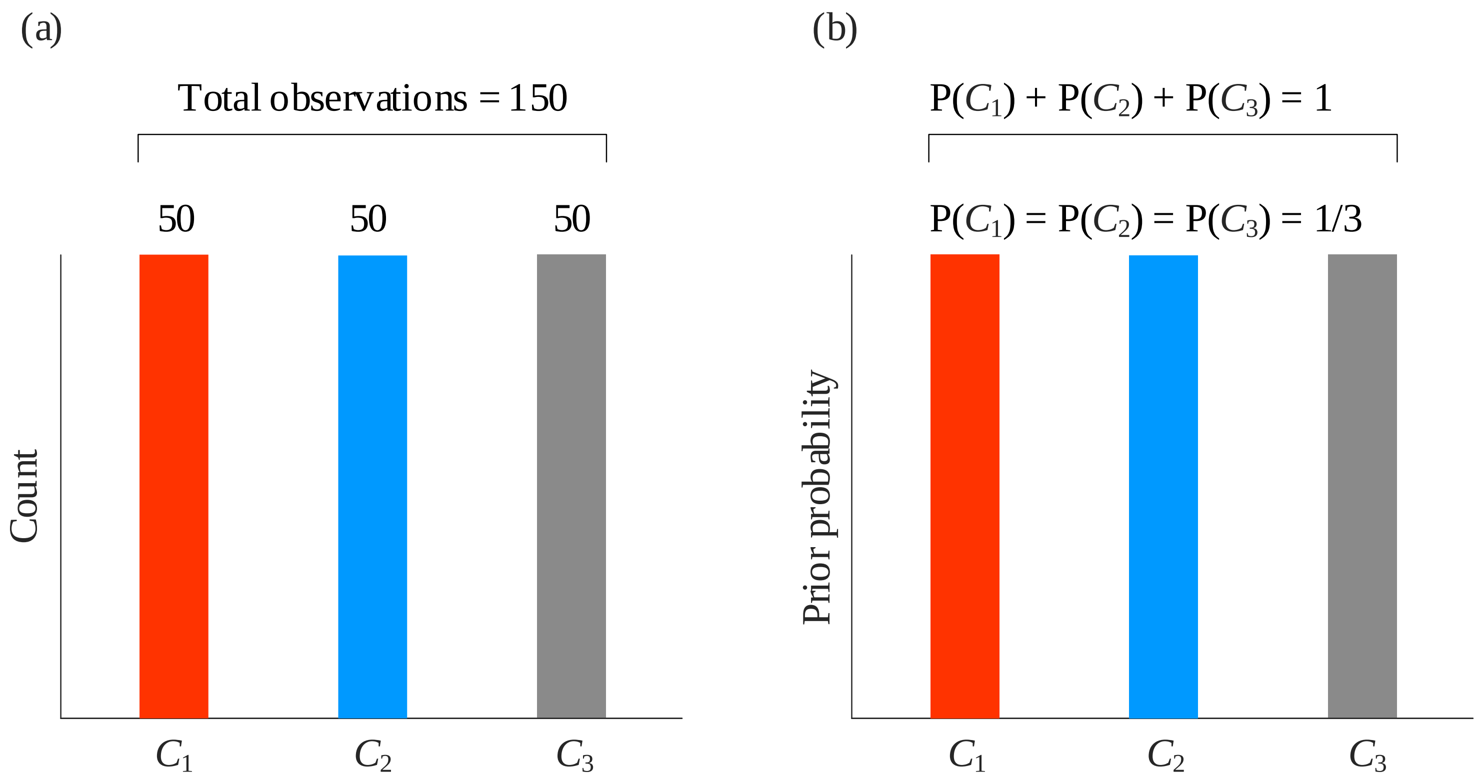
似然概率

$$\begin{cases} p(\mathbf{x}|C_2) = p(x_1, x_2|C_2) = p(x_1|C_2)\cdot p(x_2|C_2) \\ p(\mathbf{x}|C_3) = p(x_1, x_2|C_2) = p(x_1|C_3)\cdot p(x_2|C_3) \end{cases}$$



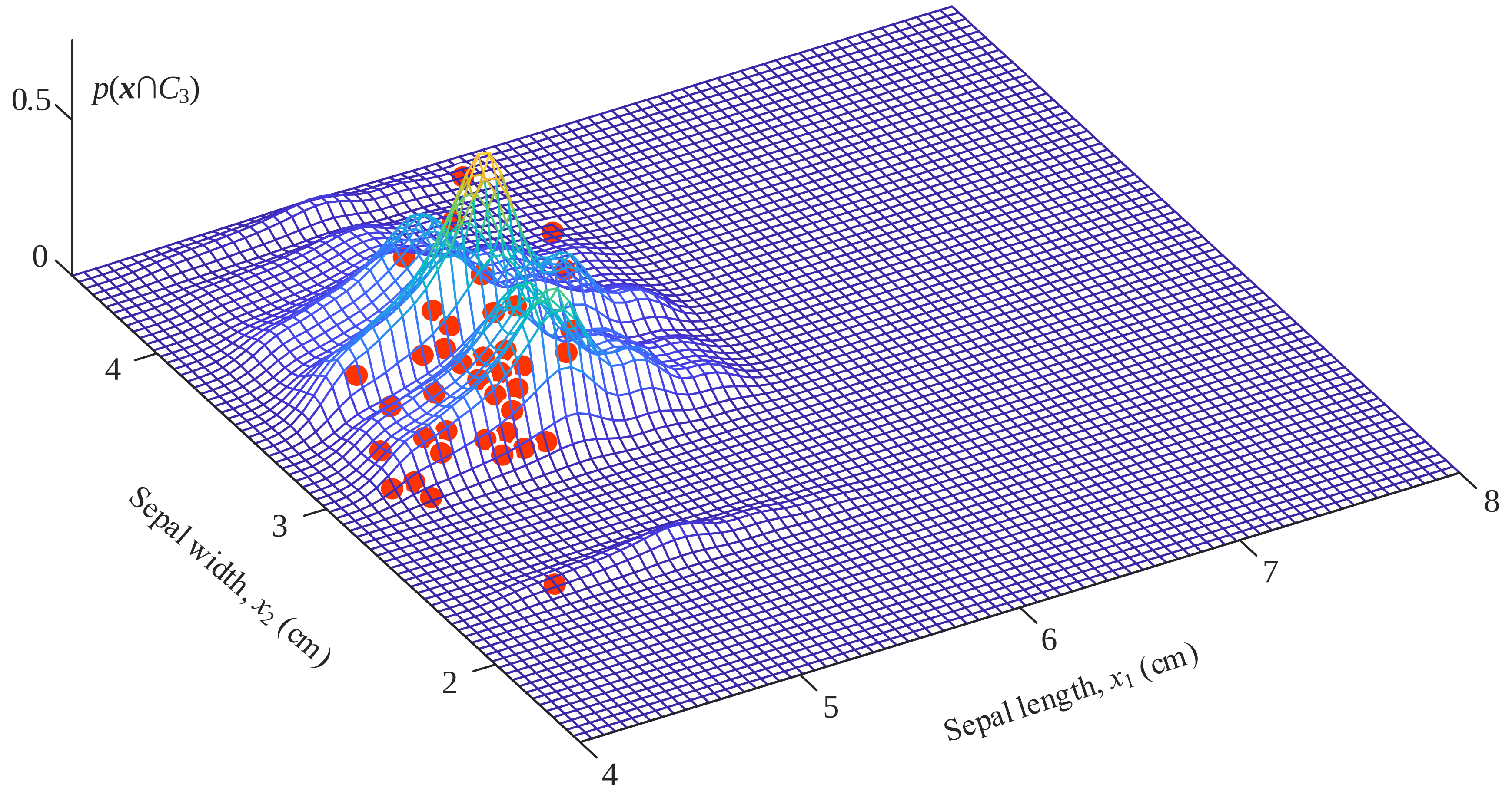
先验概率

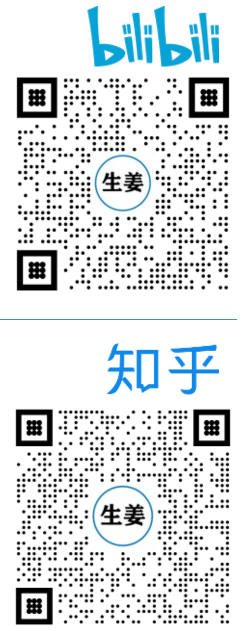
$$\mathrm{P}\left(C_1\right)=\frac{\operatorname{count}\left(C_1\right)}{\operatorname{count}(\Omega)}, \quad \mathrm{P}\left(C_2\right)=\frac{\operatorname{count}\left(C_2\right)}{\operatorname{count}(\Omega)}, \quad \mathrm{P}\left(C_3\right)=\frac{\operatorname{count}\left(C_3\right)}{\operatorname{count}(\Omega)},$$



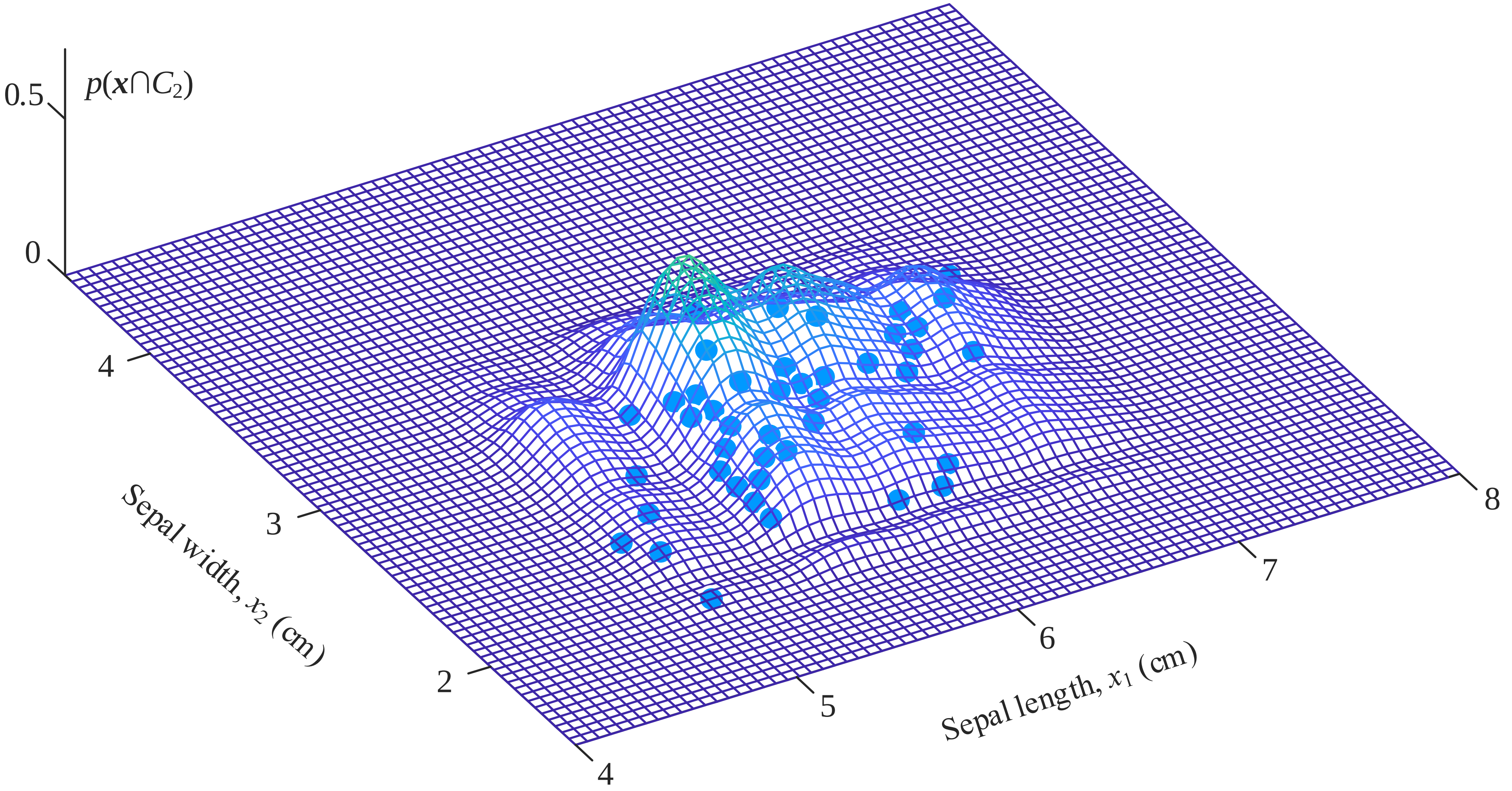
联合概率

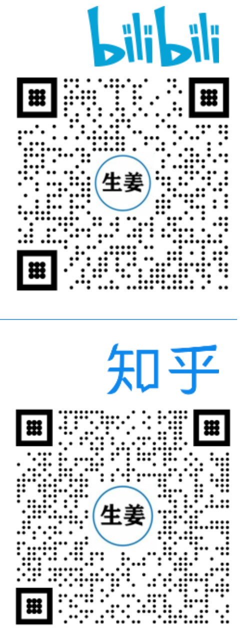
$$p(C_1 \cap \boldsymbol{x}) = p(\boldsymbol{x} | C_1) \mathrm{P}(C_1)$$



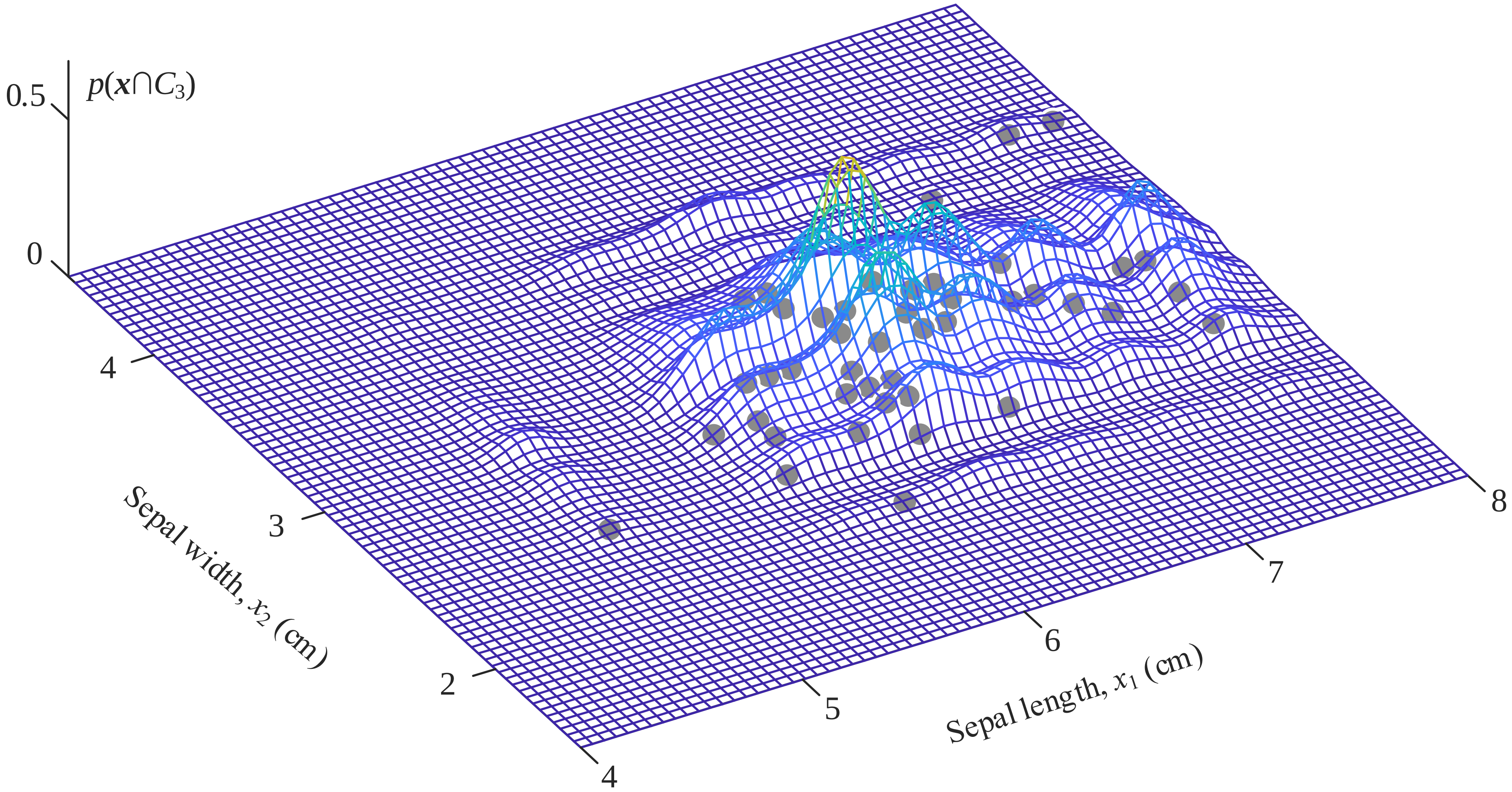


联合概率

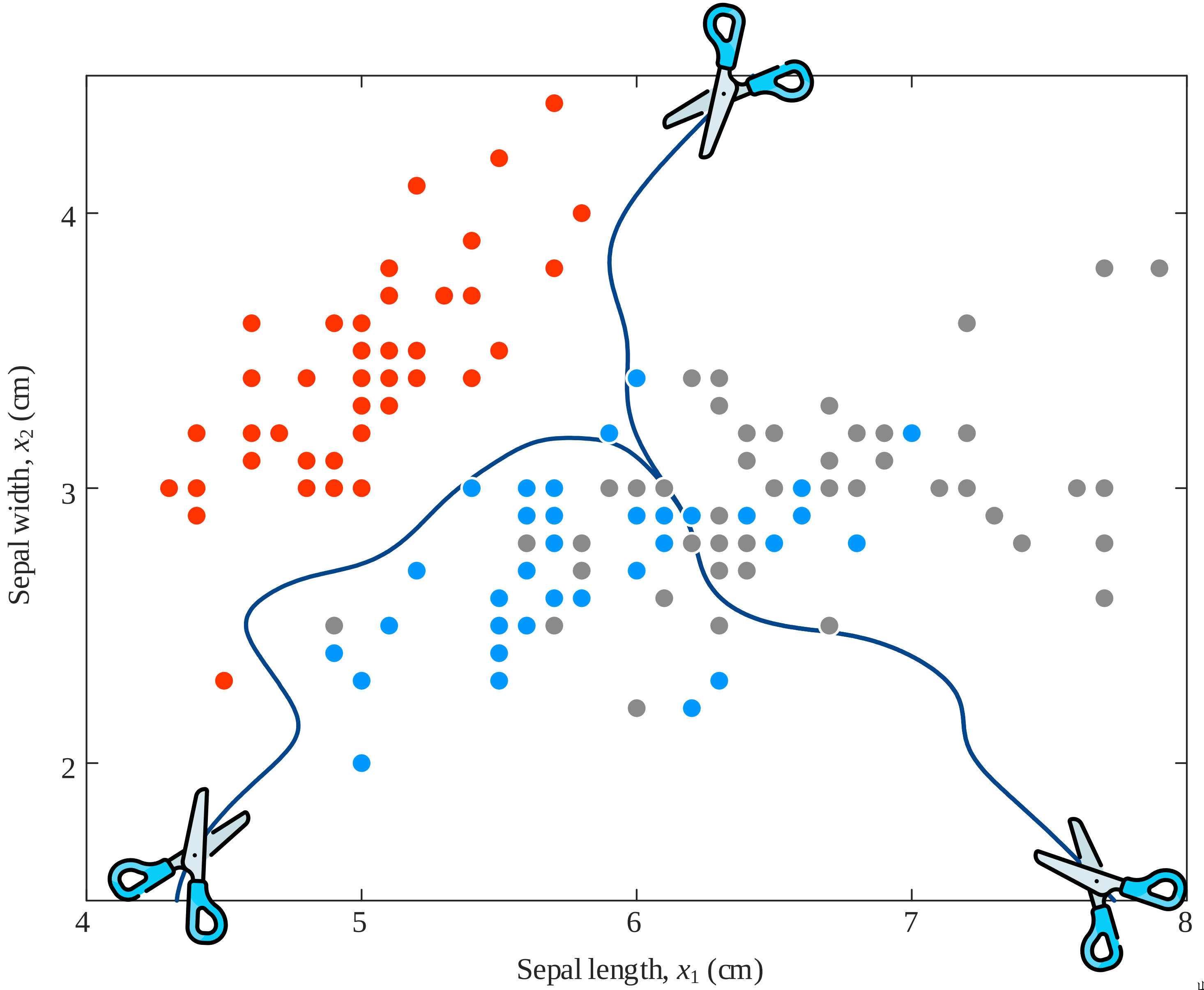

0.5
0
$p(\boldsymbol{x} \cap C_2)$
4
3
2
Sepal width, x_2 (cm)
4
5
6
7
8
Sepal length, x_1 (cm)


bilibili
生姜
知乎
生姜

联合概率



朴素贝叶斯决策边界，基于核密度估计KDE

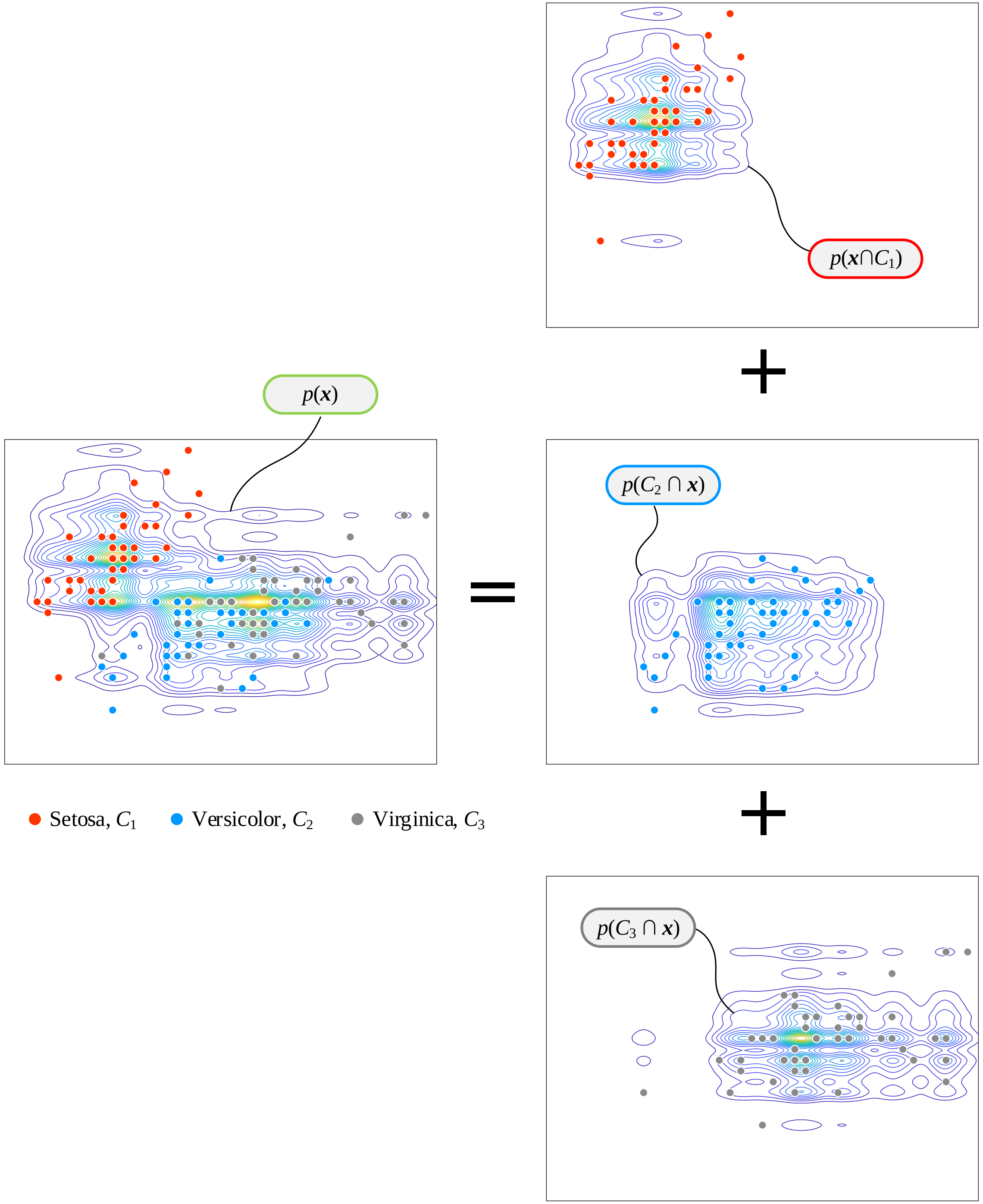


证据因子

$$p(\boldsymbol{x}) = \sum_{k=1}^{K}\left[\mathrm{P}(C_k)\prod_{j=1}^{D}p(x_j|C_k)\right]$$

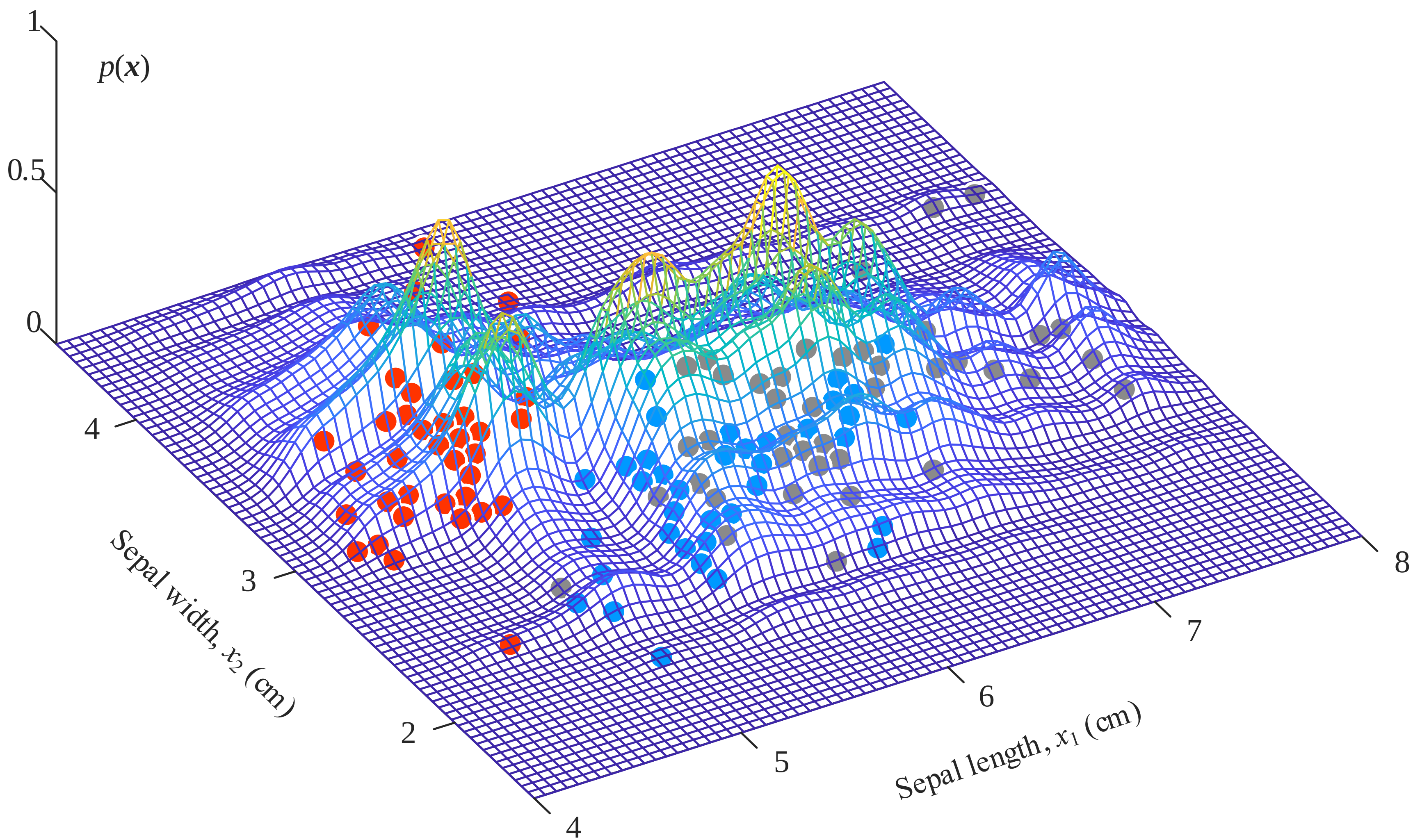
$$\begin{aligned} p(\boldsymbol{x}) &= p(C_1 \cap \boldsymbol{x}) + p(C_2 \cap \boldsymbol{x}) + p(C_3 \cap \boldsymbol{x}) \\ &= p(\boldsymbol{x}|C_1)\mathrm{P}(C_1) + p(\boldsymbol{x}|C_2)\mathrm{P}(C_2) + p(\boldsymbol{x}|C_3)\mathrm{P}(C_3) \end{aligned}$$

证据因子

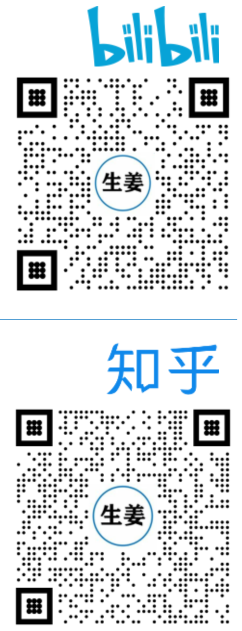


$p(x)$
$p(x \cap C_1)$
$p(C_2 \cap x)$
$p(C_3 \cap x)$
=
+
+
Setosa, C_1
Versicolor, C_2
Virginica, C_3

证据因子



1
0.5
0
$p(\boldsymbol{x})$
4
3
2
Sepal width, x_2 (cm)
4
5
6
7
8
Sepal length, x_1 (cm)

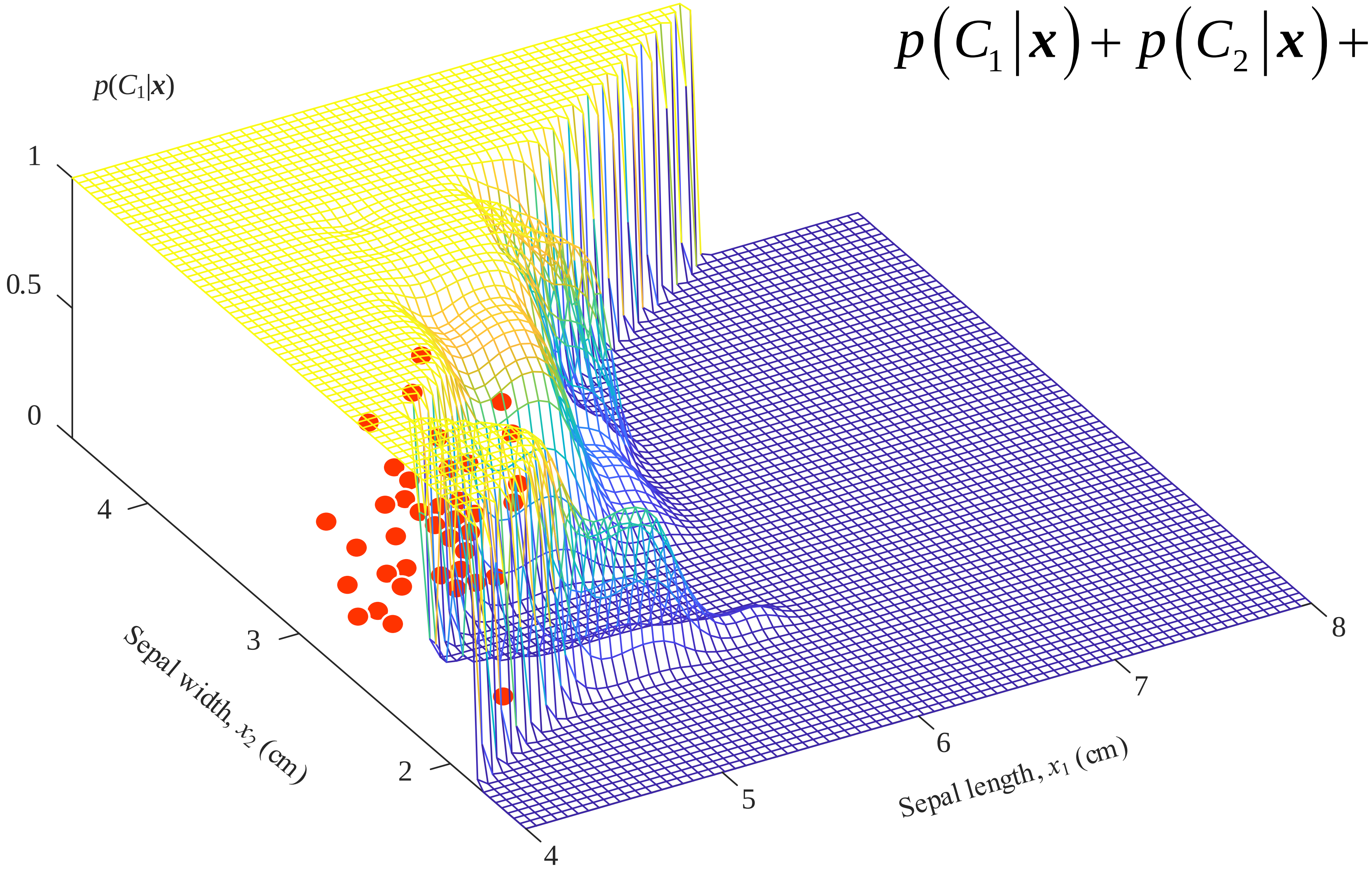


bilibili
生姜
知乎
生姜

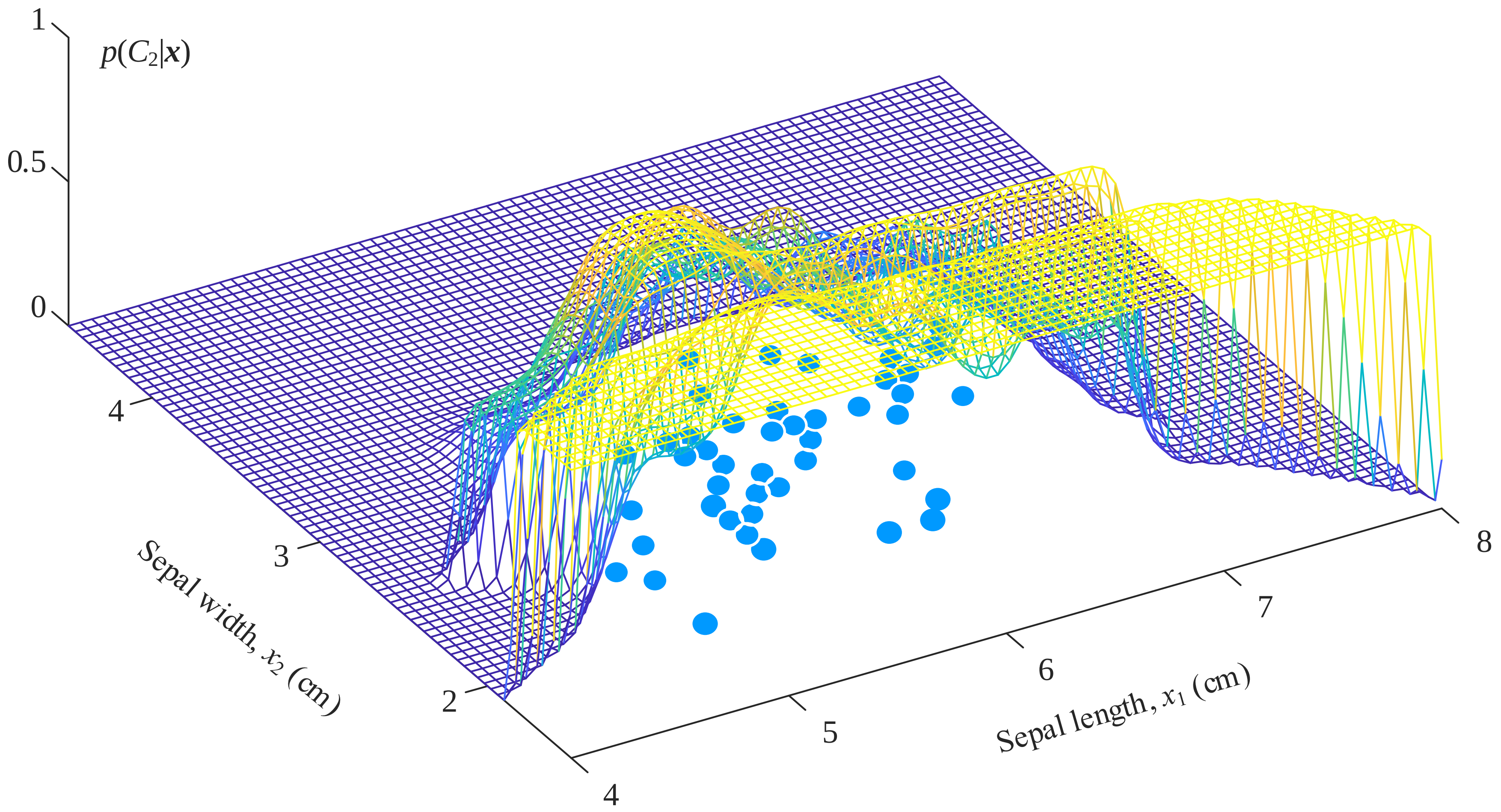
后验概率

$$p\left(C_1 \middle| \boldsymbol{x}\right) = \frac{p\left(C_1 \cap \boldsymbol{x}\right)}{p\left(\boldsymbol{x}\right)}$$

$$p\left(C_1 \middle| \boldsymbol{x}\right) + p\left(C_2 \middle| \boldsymbol{x}\right) + p\left(C_3 \middle| \boldsymbol{x}\right) = 1$$

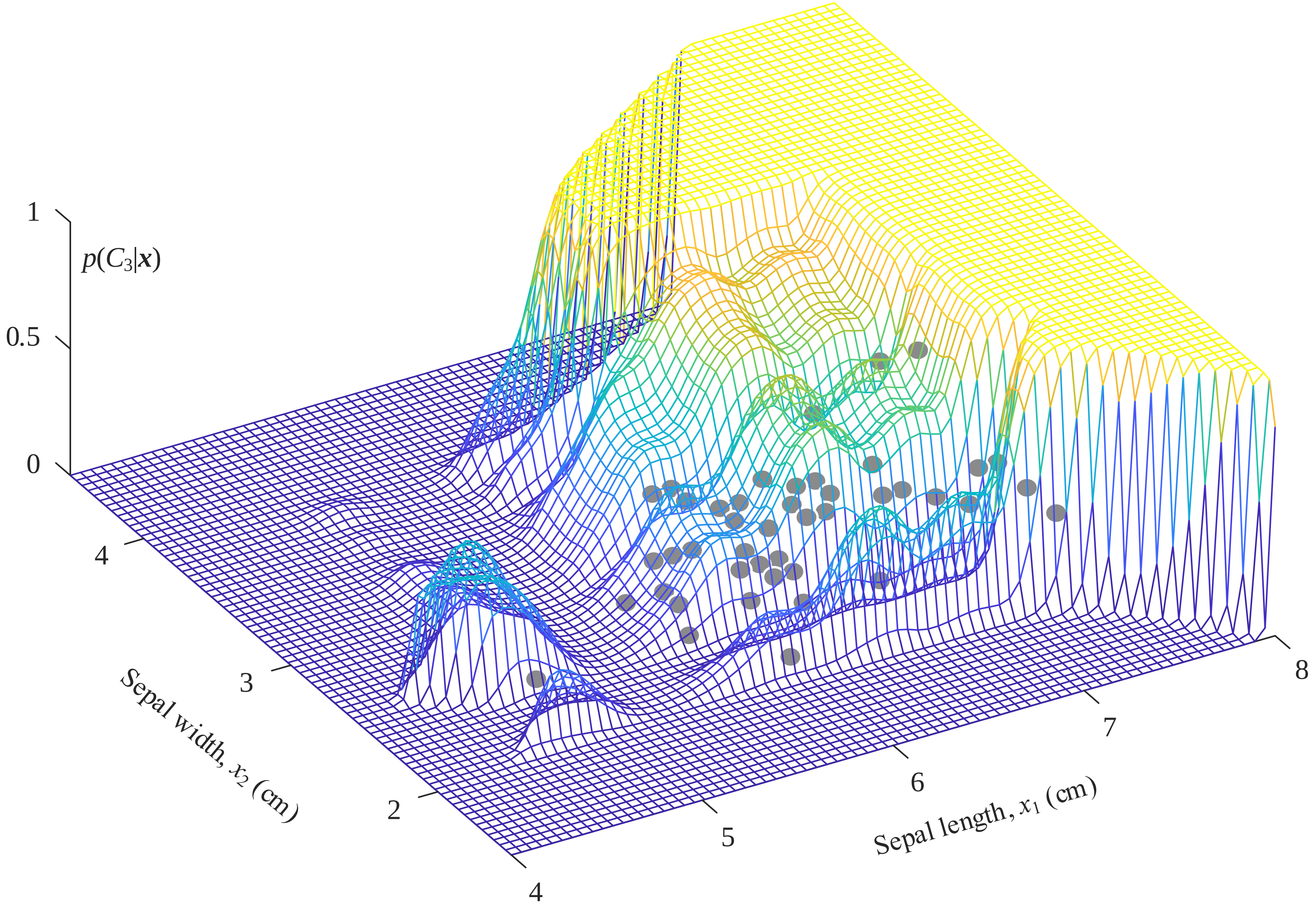


后验概率



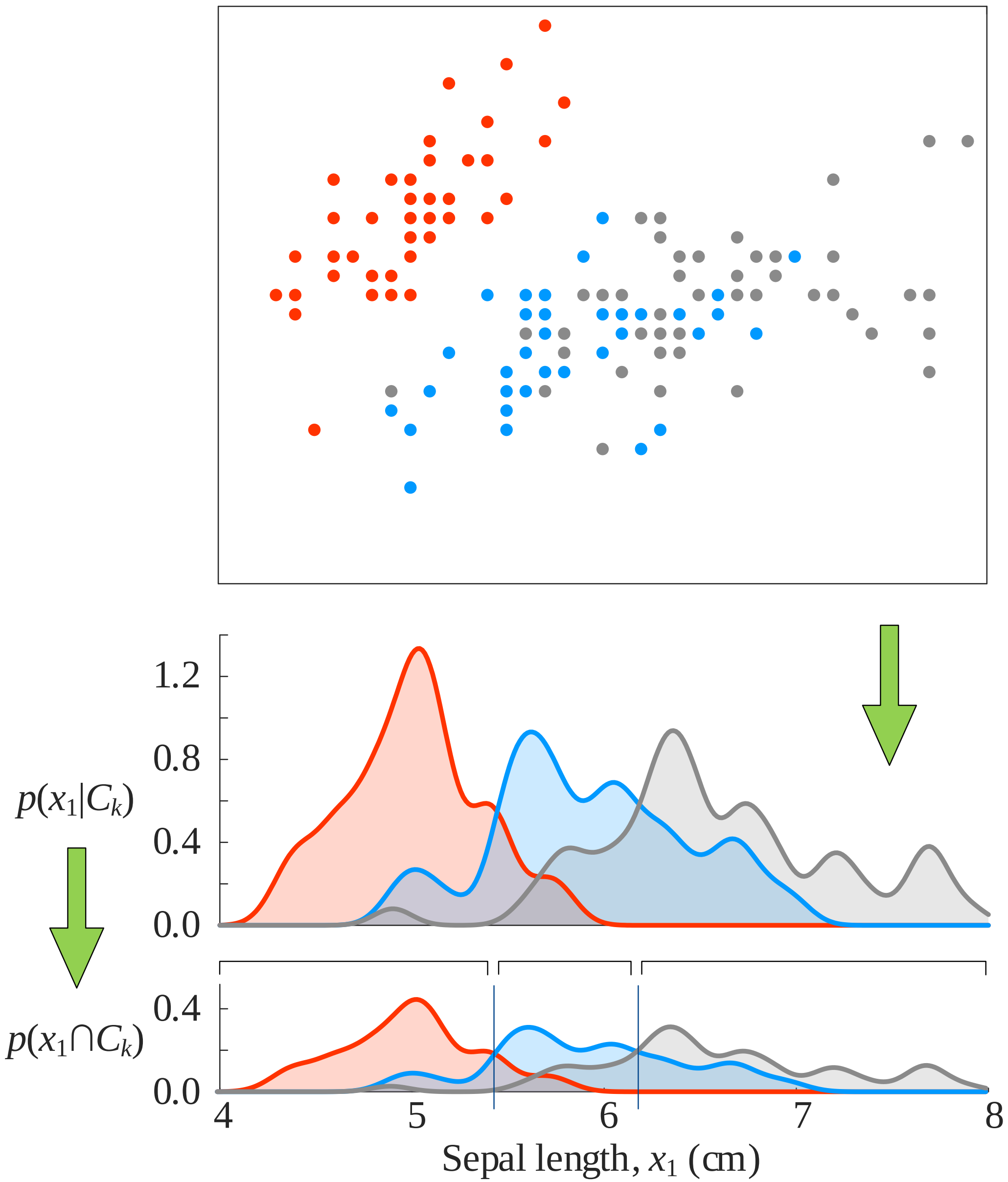
$p(C_2|\mathbf{x})$
1
0.5
0
4
3
2
Sepal width, x_2 (cm)
4
5
6
7
8
Sepal length, x_1 (cm)

后验概率

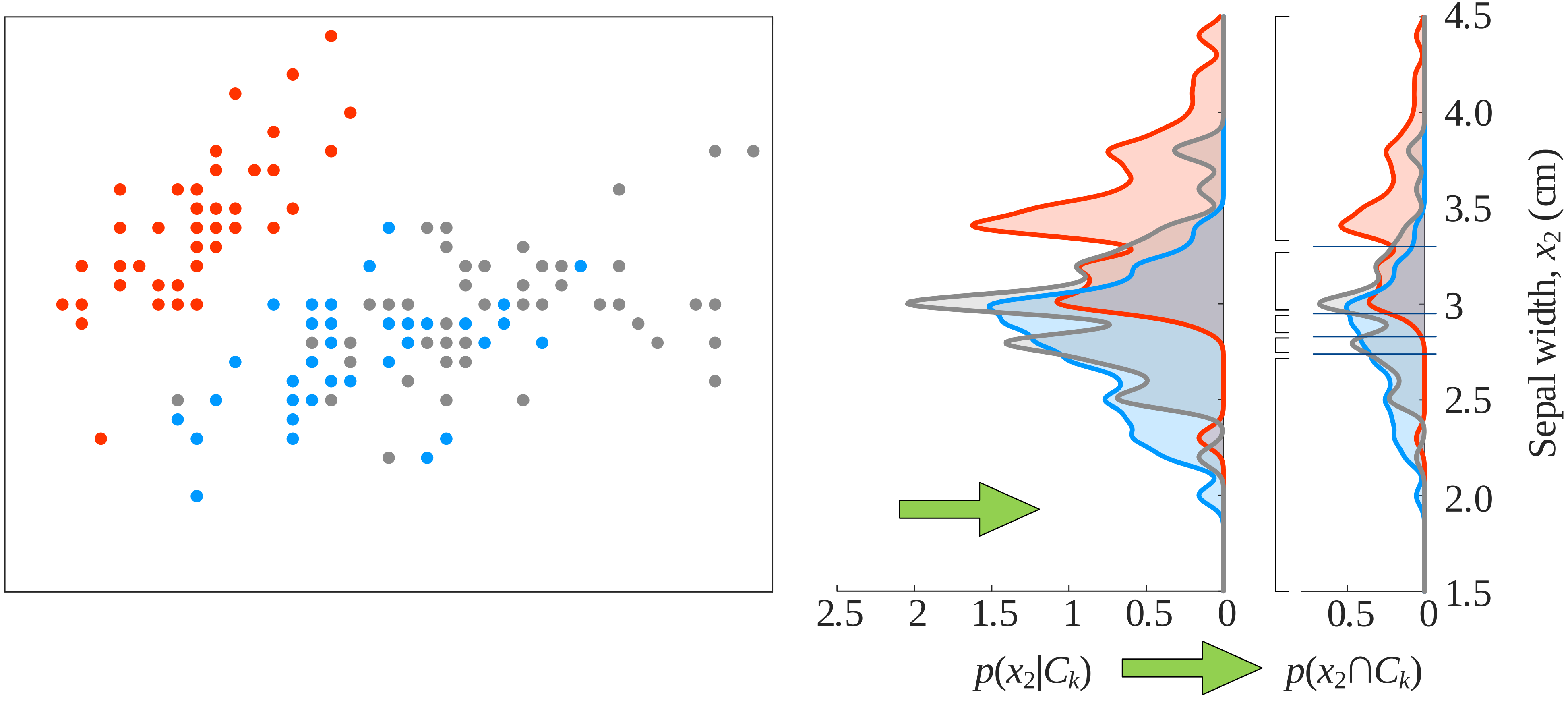


1
0.5
0
$p(C_3|\boldsymbol{x})$
4
3
2
Sepal width, x_2 (cm)
4
5
6
7
8
Sepal length, x_1 (cm)

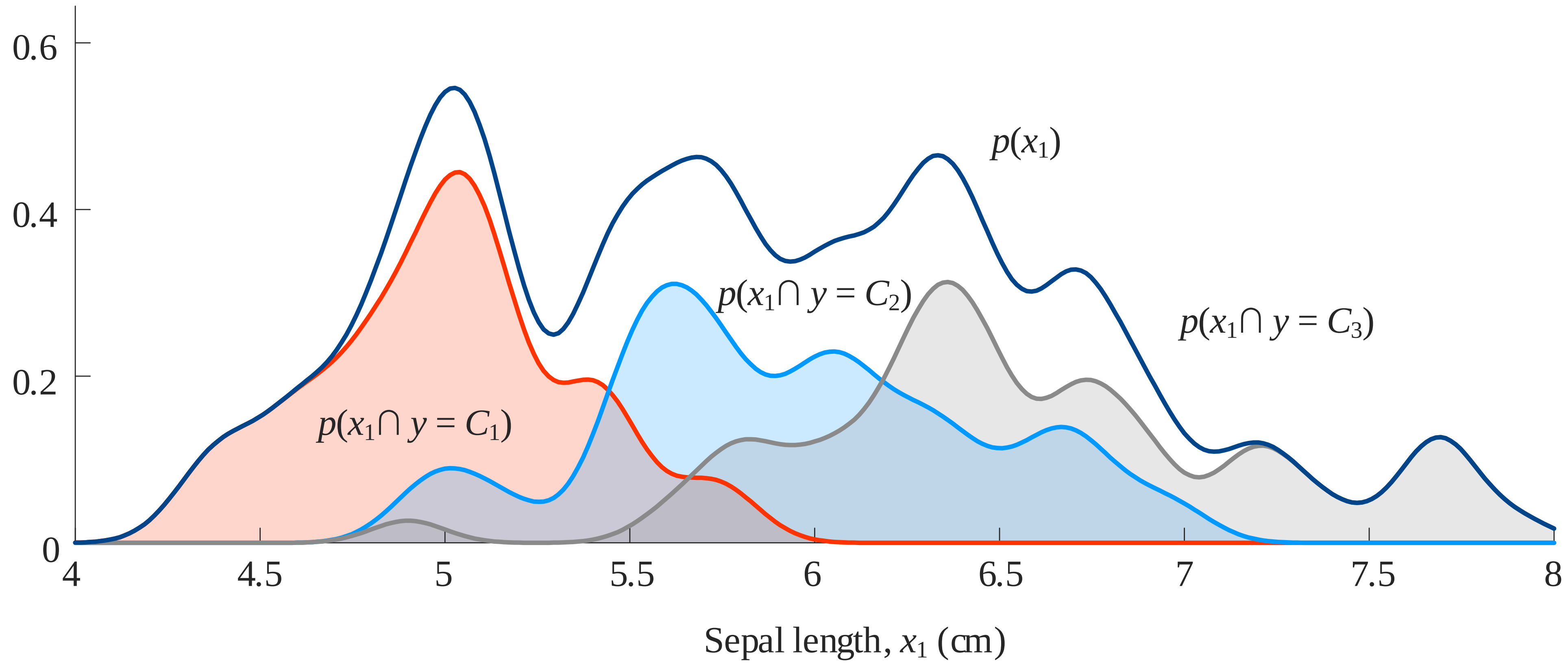
似然概率密度到联合概率，花萼长度特征



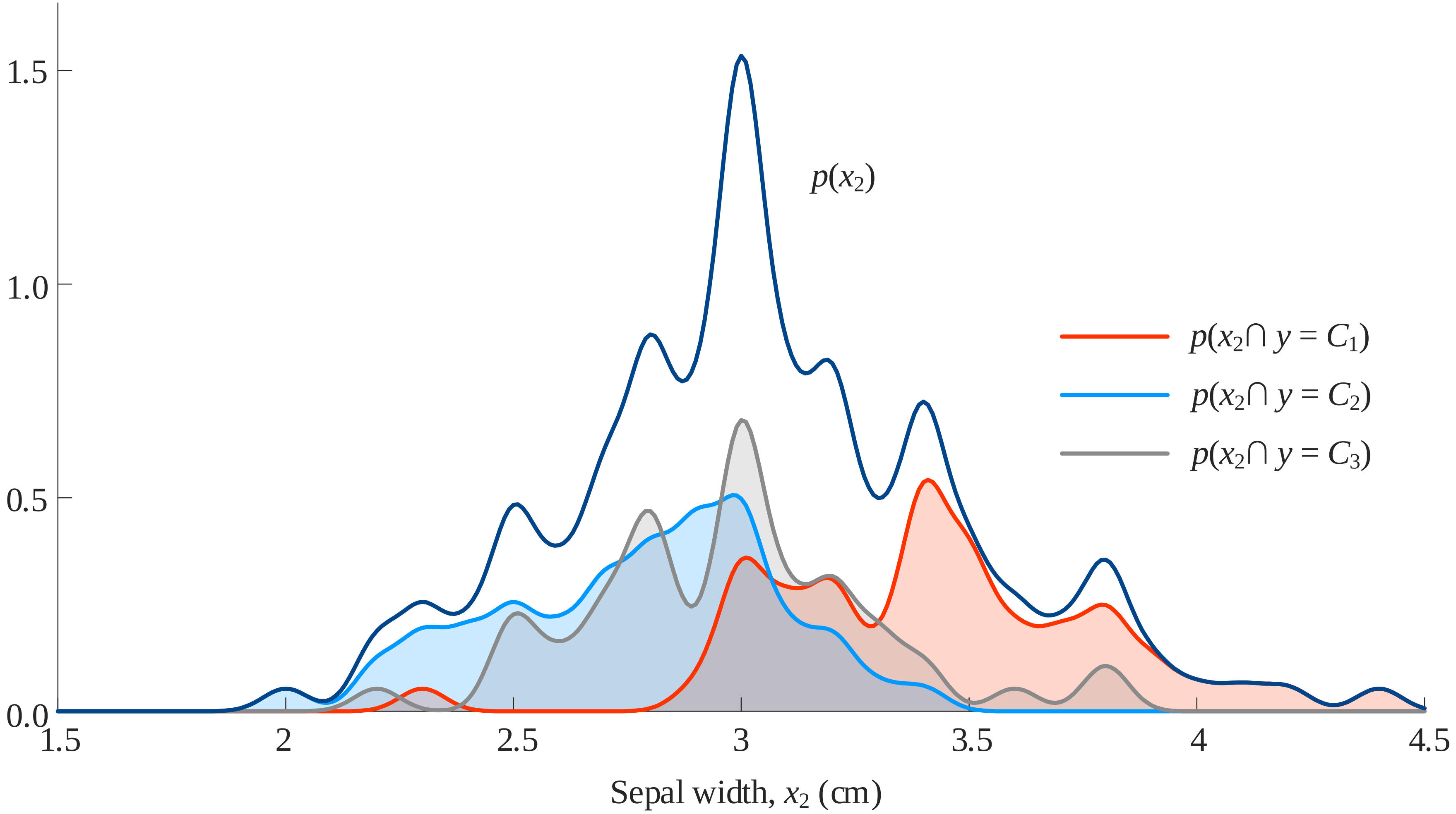
似然概率密度到联合概率，花萼宽度特征



证据因子/边际概率，花萼长度特征



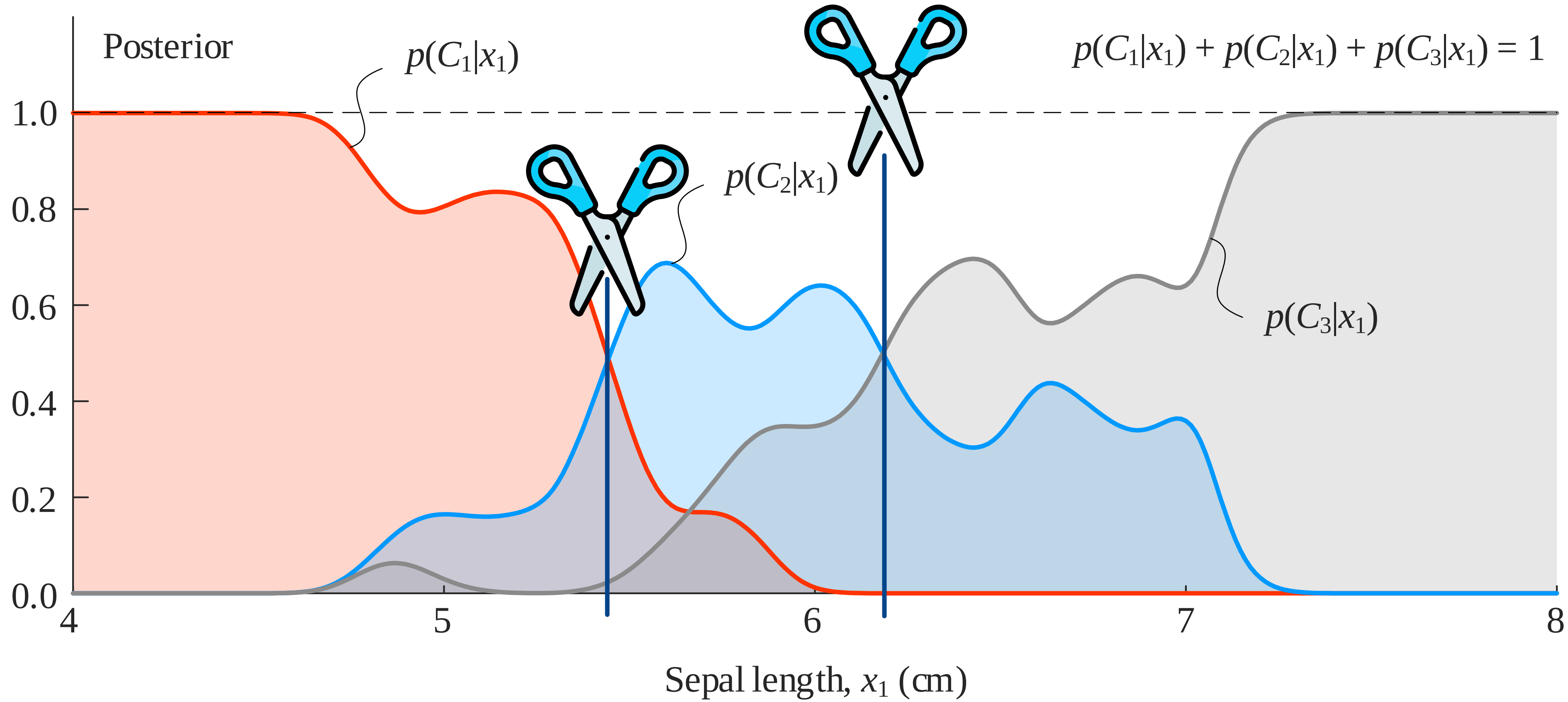
证据因子/边际概率，花萼宽度特征



bilibili

知乎

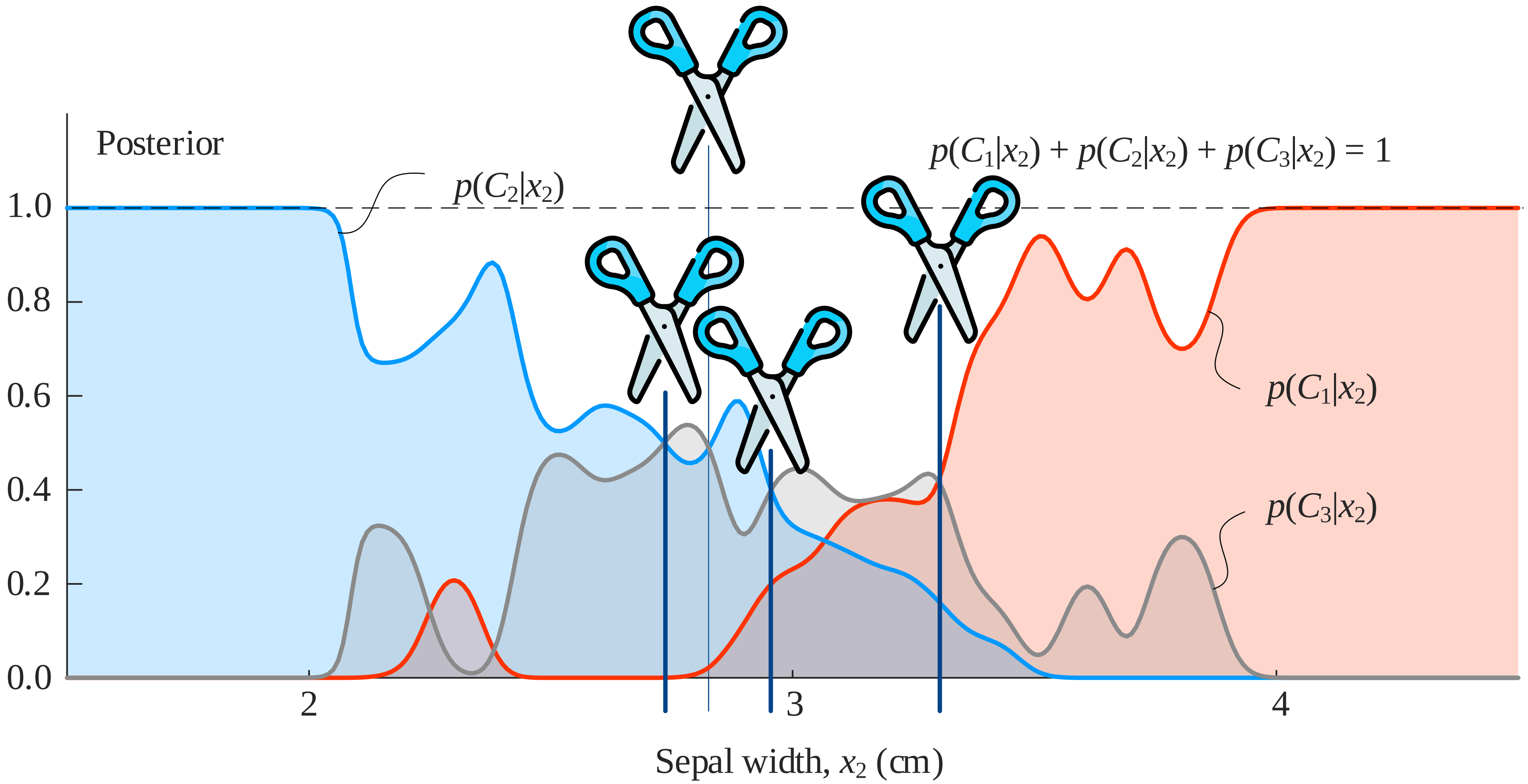
后验概率，花萼长度特征



bilibili

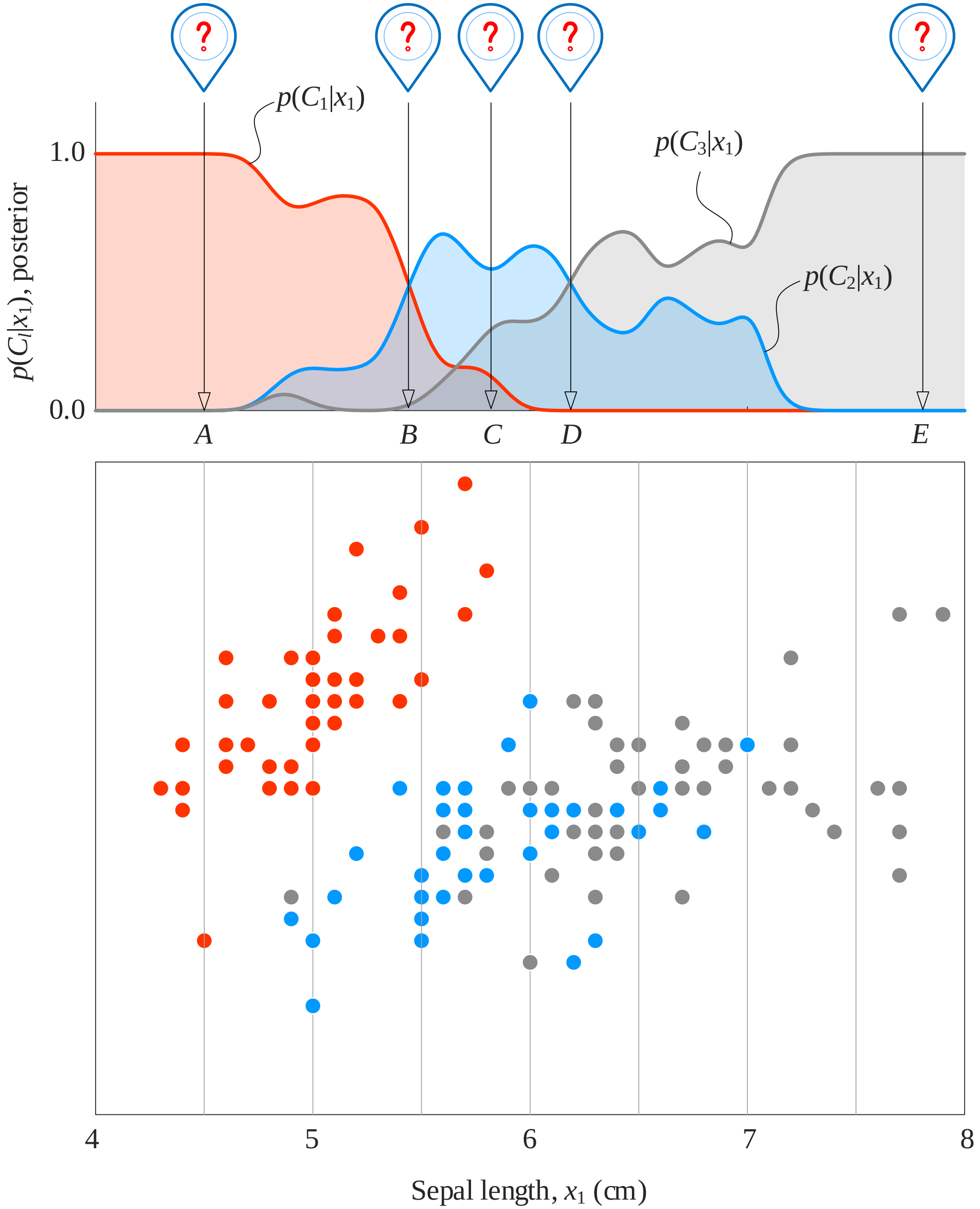
知乎

后验概率，花萼宽度特征

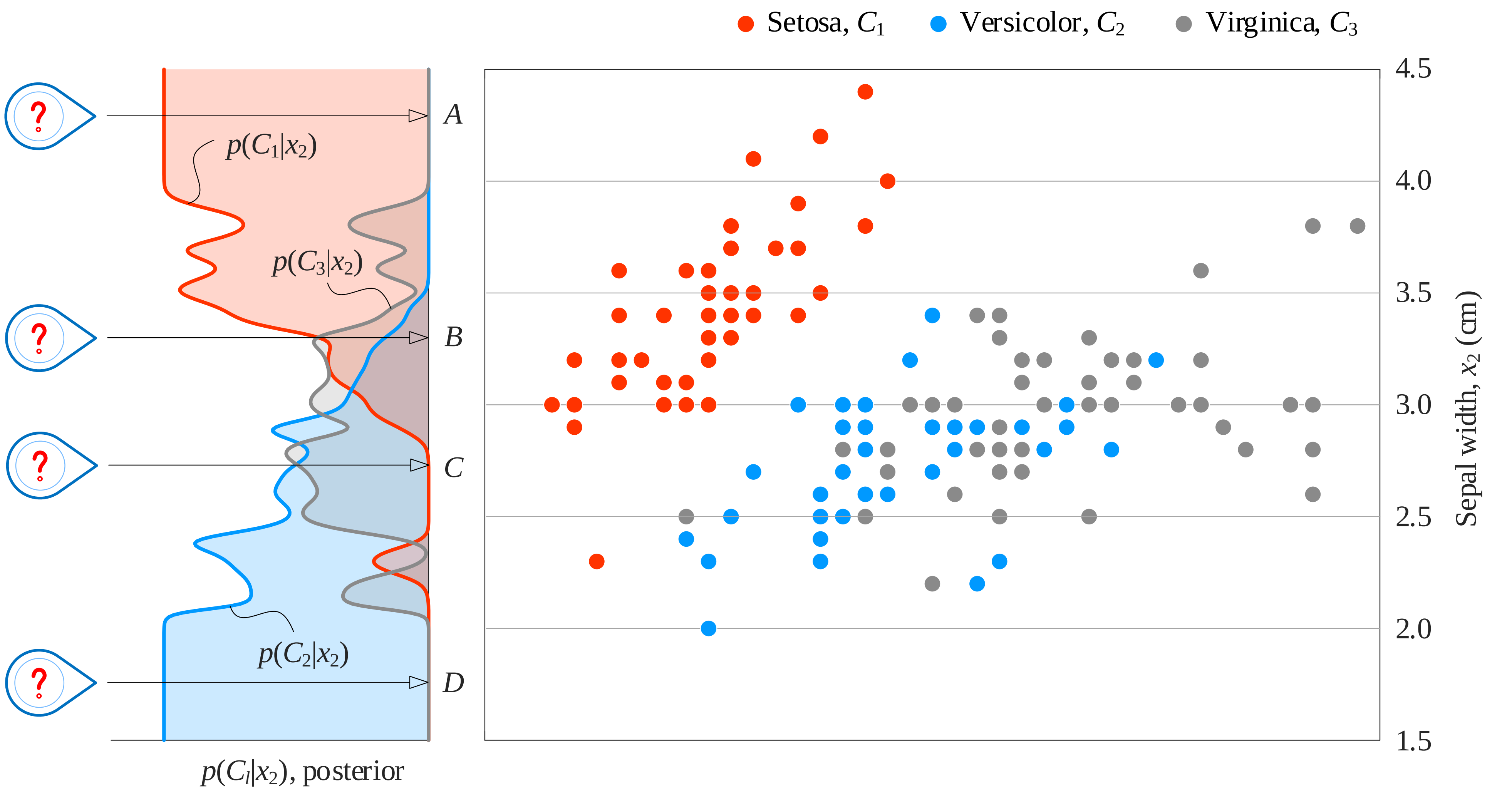


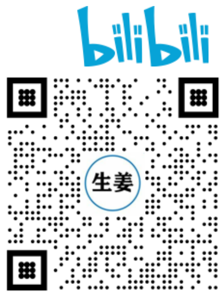
Posterior
$p(C_2|x_2)$
$p(C_1|x_2) + p(C_2|x_2) + p(C_3|x_2) = 1$
1.0
0.8
0.6
0.4
0.2
0.0
$p(C_1|x_2)$
$p(C_3|x_2)$
2
3
4
Sepal width, x_2 (cm)

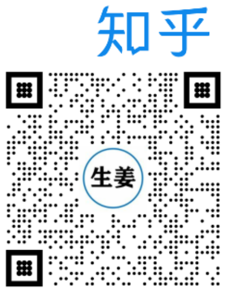
利用花萼长度特征后验概率，进行分类预测



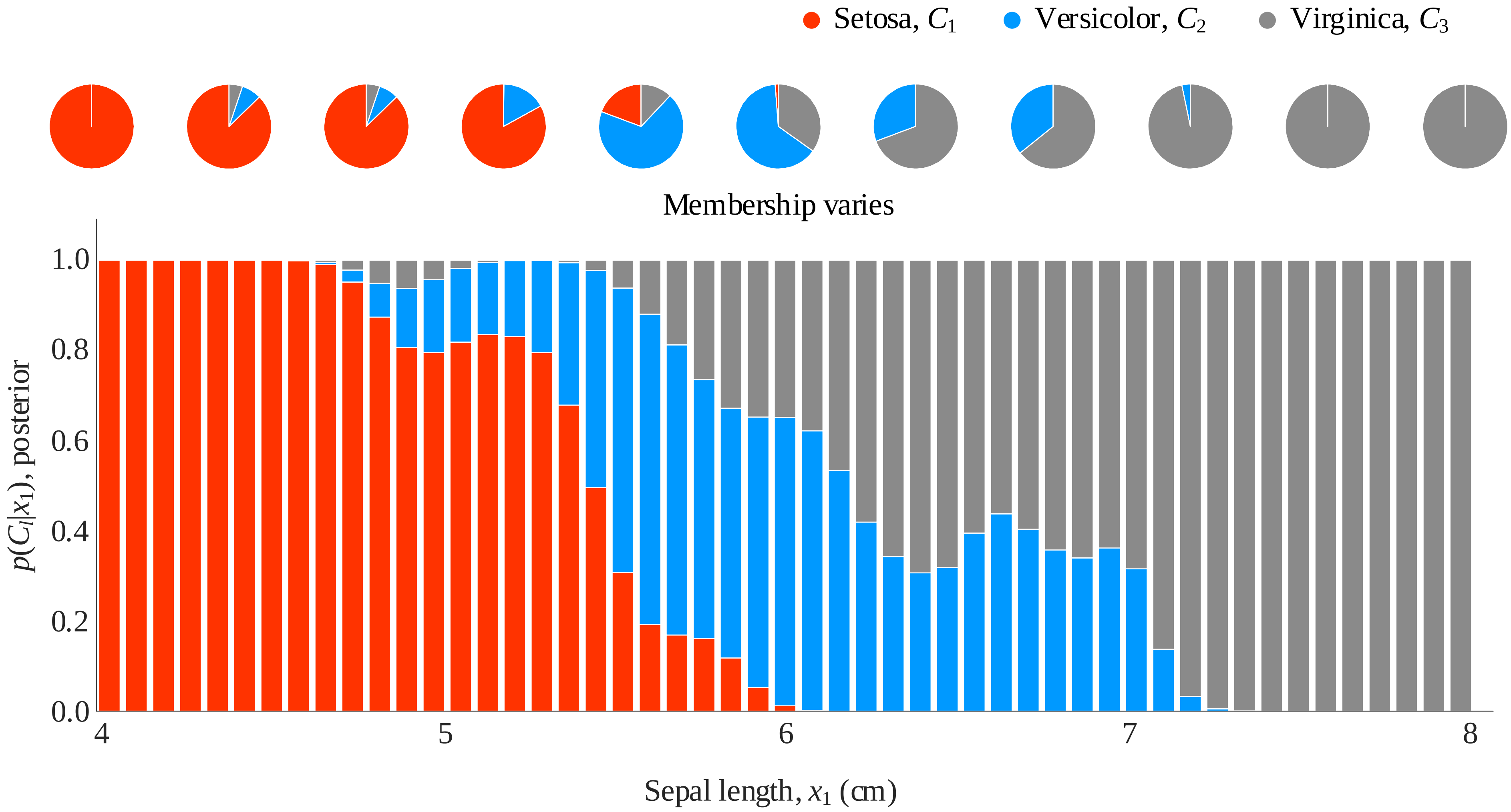
利用花萼宽度特征后验概率，进行分类预测



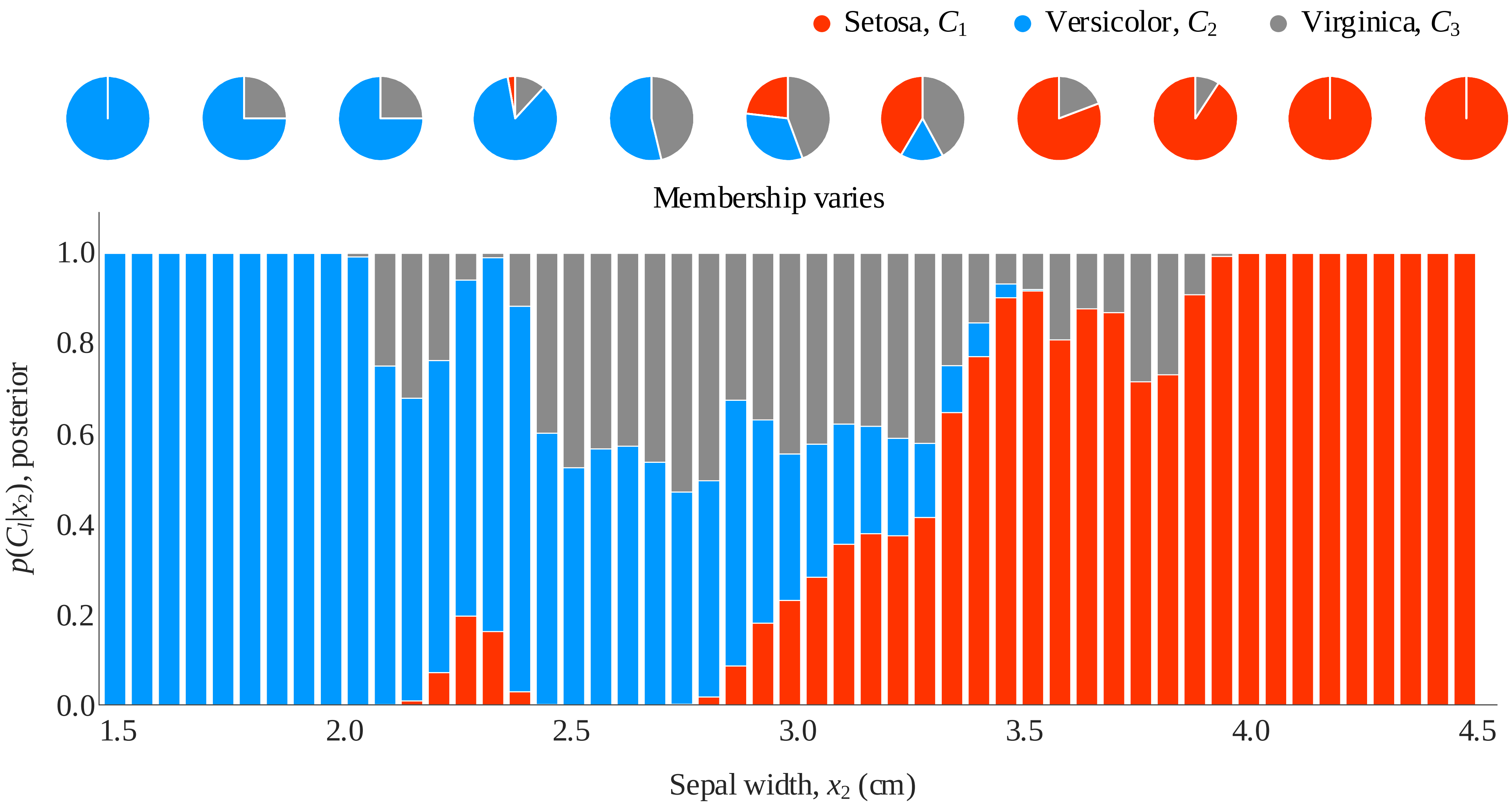




堆积直方图和饼图，利用花萼长度特征成员值确定分类



堆积直方图和饼图，利用花萼宽度特征成员值确定分类



bilibili

知乎

比较四种概率密度函数曲线随特征变化趋势