基本思路如下，具体细节不一定精确

1. 20 个样本的性别和身高测量值(单位 cm)如下：
男： 176,174,179,186,180,173,179,165，
女： 174,160，158，162，167，171，163，160，165，167，163，168
a) 假设样本身高均服从高斯分布，用最大似然估计的方法估计男女生身高分布，并对身高如下的三个样本进行分类： 170， 165，175。
b) 假设样本身高均服从高斯分布，且男生身高的先验分布为 N(175,0)，女生身高的先验分布为 N(165,10)，试用贝叶斯估计的方法估计男女生身高分布，并对身高如下的三个样本进行分类： 170，165，175。
要求： 1.写出运算过程； 2.使用计算器

解：
a）最大似然估计

$\mu_{男}=\frac{1}{n}\sum x_k=\frac{1}{8}*(\mathbf{176+174+179+186+180+173+179+165})$= 176.50 (cm)

$\hat{\varepsilon}_{男}=\frac{1}{n}\sum(x_k-\mu_{男})^2=\frac{1}{8}*[(\mathbf{176-176.5})^2+(\mathbf{174-176.5})^2+(\mathbf{179-176.5})^2+(\mathbf{186-176.5})^2+(\mathbf{180-176.5})^2+(\mathbf{173-176.5})^2+(\mathbf{179-176.5})^2+(\mathbf{165-176.5})^2]$ = 33.25

类似地，$\mu_{女}=164.83$，$\hat{\varepsilon}_{女}$=20.81

P(男)=8/20=0.4，P(女)=12/20=0.6

对于 170 的样本，

$P(x|男)*P(男)=\frac{1}{\sqrt{2\pi\hat{\varepsilon}_{男}}}\exp\left(-\frac{1}{2}\frac{(x-\mu_{男})^2}{\hat{\varepsilon}_{男}}\right)=\frac{1}{\sqrt{2*3.14*33.25}}\exp\left(-\frac{1}{2}\frac{(x-176.50)^2}{33.25}\right)*0.4=0.0147$

$P(x|女)*P(女)=\frac{1}{\sqrt{2\pi\hat{\varepsilon}_{女}}}\exp\left(-\frac{1}{2}\frac{(x-\mu_{女})^2}{\hat{\varepsilon}_{女}}\right)=\frac{1}{\sqrt{2*3.14*20.81}}\exp\left(-\frac{1}{2}\frac{(x-164.83)^2}{20.81}\right)*0.6=0.0276$

P(x|男)* P(男) < P(x|女)* P(女),所以该样本是女性

类似的，对于 165 样本 P(165|男)< P(170|男),而 P(165|女)>P(170|女)该样本是女性

对于 175 样本　　P(x|男)* P(男)=0.0268，　P(x|女)* P(女)=0.0044
P(x|男)* P(男) > P(x|女)* P(女)该样本是男性

b)贝叶斯估计

$$\hat{\mu} = \frac{\frac{\mu_0}{N} + \frac{\sigma_0^{2p}}{\hat{\sigma}^{2p}} \frac{\sum_{k=1}^{N} x_k}{N}}{\frac{1}{N} + \frac{\sigma_0^{2p}}{\hat{\sigma}^{2p}}}$$

$$\hat{\sigma}^{2p} = \sum_{k=1}^{N} \frac{||x - \hat{\mu}||^2}{N}$$

这道题涉及迭代（参见上面公式，更新$\hat{\mu}$需要$\hat{\sigma}^{2p}$，更新$\hat{\sigma}^{2p}$需要$\hat{\mu}$），编程算还行，手工计算太琐碎，不管了，只要编程时候会套公式就行

2. 已知 12 个女生样本身高分布如上，试用 k 近邻密度估计方法估计女生在身高区间 160，161，…，170 各点的密度。

解：

女生身高分布为 174,160，158，162，167，171，163，160，165，167，163，168

假设取 3 近邻（通常题目会指定），套用公式$\widehat{p(x)} = \frac{K}{N \times V_k}$，其中 N=12，K=3，Vk 即区间大小，一维情况下为 x 与其第 K 个近邻距离的 2 倍

$\widehat{p(160)} = \frac{3}{12\times4}$=0.0625 （160 的三个近邻是 160 160 和 162，最远近邻 162 的距离是 2,2 倍即 Vk=4）

类似地，$\widehat{p(161)}$=$\frac{3}{12\times2}$=0.125 （161 的三个近邻是 160 160 和 162，最远近邻 162 的距离是 1,2 倍即 Vk=2）

$\widehat{p(162)}$=$\frac{3}{12\times2}$=0.125 （162 的三个近邻是 162 163 和 163，最远近邻 163 的距离是 1,2 倍即 Vk=2）

其它类似……