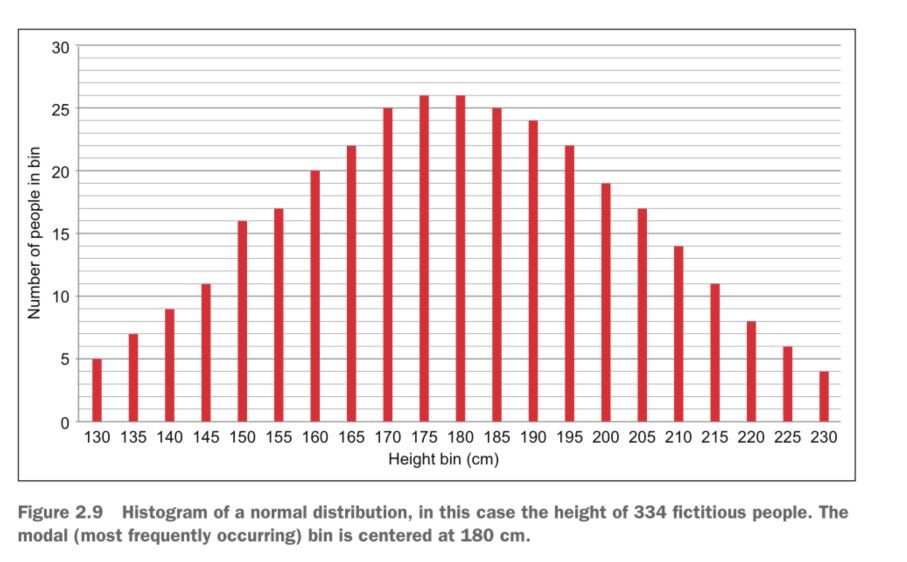
**一、人均身高为例**

大概可以取一个X~N(180,25)左右的正态分布 155~205占据68.27%

130~230占据95%的比例

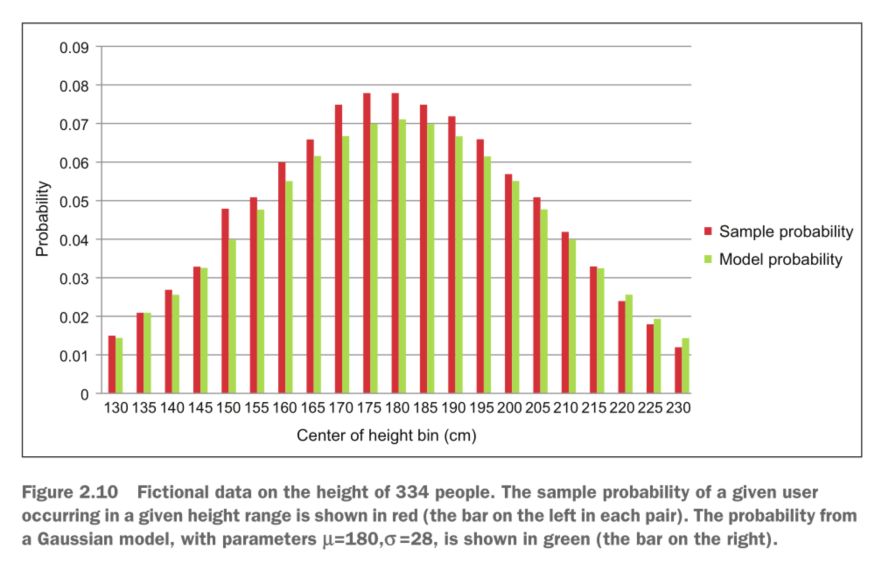
**二、高斯模型**

实际上我们并不知道对应的均值和方差,而高斯模型就是为了取出最正确的均值和方差。

**三、判定标准**

每一个标量在总体量中对应的概率与高斯预估模型中根据预估均值和方差构成的概率分布函数算出来的概率对比，拟合度越高，则模型越精确。

**四、EM算法**

通过观察采样的概率值和模型概率值的接近程度，来判断一个模型是否拟合良好。然后我们通过调整模型以让新模型更适配采样的概率值。反复迭代这个过程很多次，直到两个概率值非常接近时，我们停止更新并完成模型训练。