

# MA263 概率论与数理统计大作业

薛昊天 518021910506

## 1 问题 1

### 1.1 题目描述

$X \sim N(\mu_1, \sigma_1^2)$ ,  $Y \sim N(\mu_2, \sigma_2^2)$ ,  $\eta \sim B(1, p)$  则有  $Z = X + \eta Y$  服从混合高斯分布。利用以上规则生成 10000 个数据，画出频率分布直方图。改变参数，观察频率分布直方图的“峰”的变化。

## 2 实验及结论

利用 Numpy 库的 random 函数可以生成服从任意参数的正态分布随机数，以及服从 0-1 分布的随机数。根据题目描述生成 10000 个 Z 的数值，再利用 Matplotlib.pyplot 库中的 hist 功能可以画出频率分布直方图。

### 2.1 0-1 分布中 p 对图像的影响

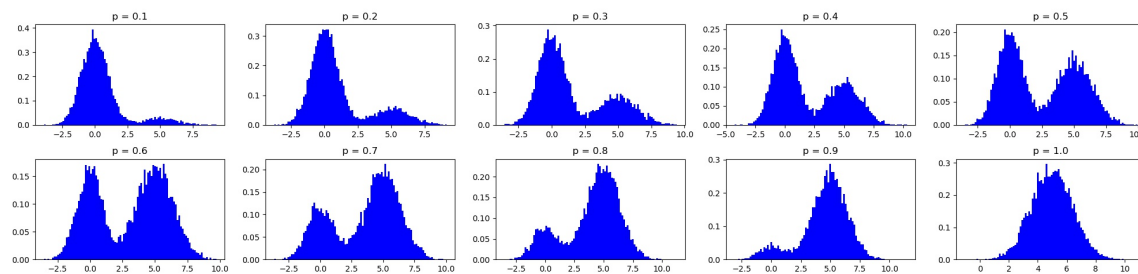


图 1

由于  $p$  对于和式中的  $\eta Y$  有着直接的影响，这里不是一般性，固定其它所有变量。可以从图 1 中看出， $p$  很小的时候，图像几乎只体现为  $X$  的分布特点；随着  $p$  值逐渐增大，图像的峰变为两个且左边的峰逐渐减弱而右边的峰逐渐增大，随后当  $p=1$  时，两个峰趋于中和且峰值对应的横坐标点位于  $\mu_1 + \mu_2$  处。

### 2.2 $\mu$ 对图像的影响

$\mu$  是正态分布中的均值，图像上反应的正态分布峰值横坐标的值，在这里不失一般性，固定其它变量，只改变  $\mu_2$ 。

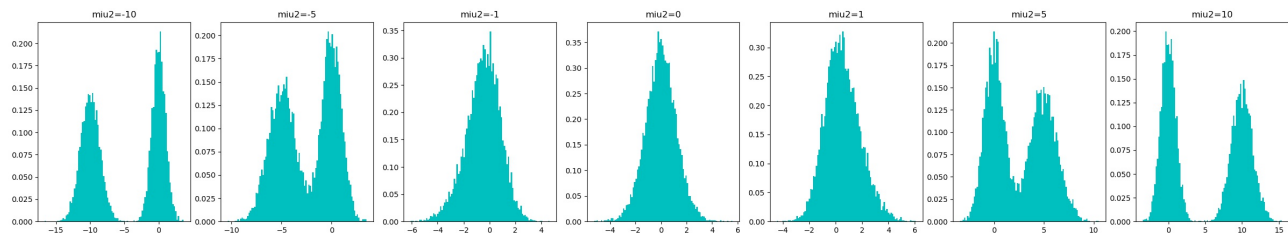


图 2

改变时 Y 变量的均值从-5 到 10，我们可以从图 2 中看出，Y 所对应的峰值的坐标从-5 到 10 移动，中途“穿过”了 X 所对应的峰，且可以看出途中两个峰的高度基本保持不变，X 对应的峰位置基本保持不变。

## 2.3 $\sigma$ 对图像的影响

和 2.2 的研究方法一致

### 2.3.1 只改变 $\sigma_1$

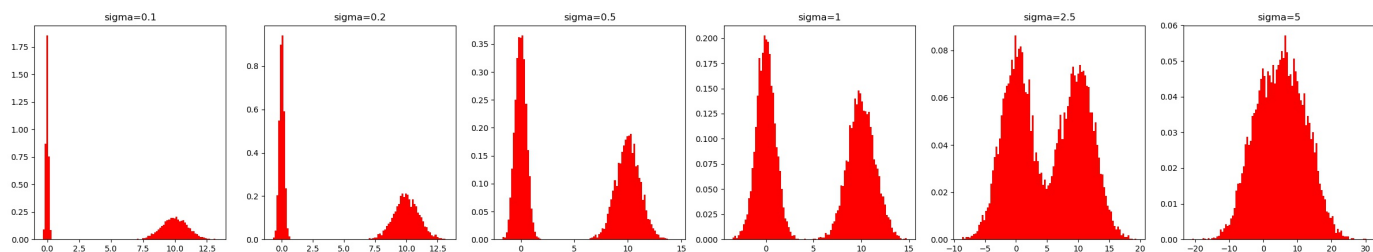


图 3.1

从图 3.1 中可以看出：当 X 的方差很小的时候，左边峰也就是 X 变量所对应的峰就会很“尖”，此时右边 Y 所对应的峰很平。当 X 的方差逐渐变大后，左边的峰逐渐趋平，而右边的峰逐渐变尖，峰值逐渐变大；最后两个峰合成一个峰。

### 2.3.2 只改变 $\sigma_2$

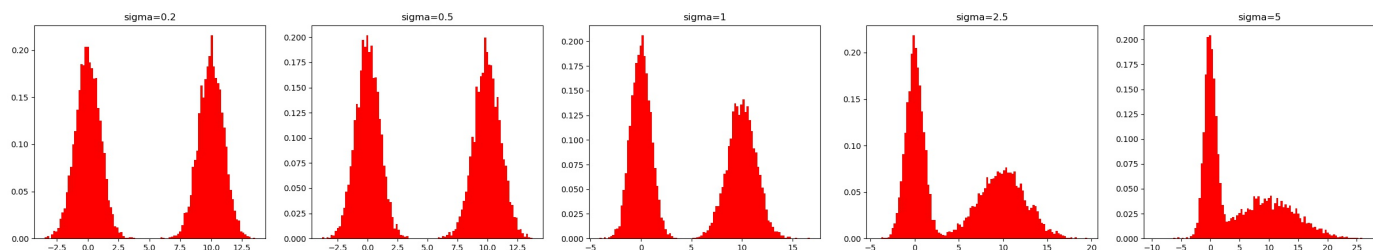


图 3.2

如图 3.2, 和 3.1 类似。当 Y 的方差逐渐变大的时候，右边的峰变矮变平，左边的峰值保持不变，变尖。最后当方差足够大，一定会合二为一（图中为完全体现，但事实如此）。

## 3 问题 2

### 3.1 题目描述

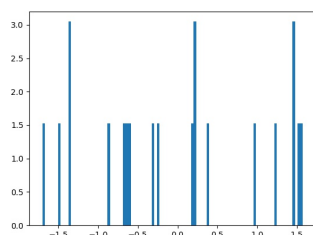
自己设定参数, 用计算机生成 1000 组, 每组  $n$  个混合高斯分布随机数: 第  $i$  组为  $(Z_{i,1}, Z_{i,2}, \dots, Z_{i,n})$

$$U_i = \frac{\sum_{k=1}^n Z_{i,k} - nEZ}{\sqrt{nDZ}}$$

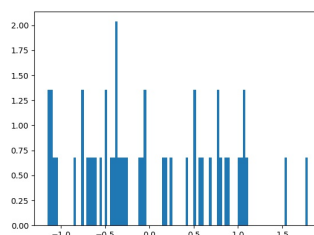
画出  $U_1, U_2, \dots, U_{1000}$  的频率分布直方图, 讨论不同  $n$  对峰的影响。

### 3.2 实验以及结论

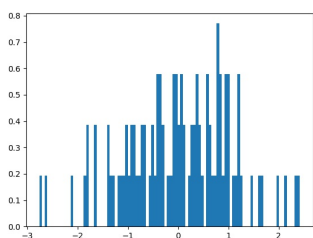
按照题目要求, 设定  $\mu_1 = 0, \mu_2 = 5, \sigma_1 = \sigma_2 = 1, p = 0.5$  分别取  $n = 20, 50, 100, 1000$  画出频率分布直方图。可以看出当  $n$  很小的时候图像十分离散, 有很多不连续峰。当  $n$  越大的时候,  $U$  的分布越来越趋向于正态分布。且由中心极限定律可以知道当  $n$  足够大的时候,  $U$  趋向于标准正态分布。所以我们可以得出, 当  $n$  很大时, 随机量相互抵消, 最后近似服从标准正态分布, 这也正是中心极限定律的思想。



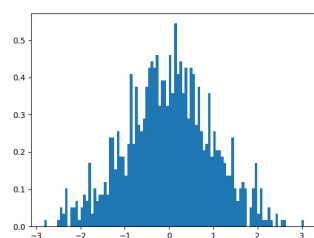
(a)  $n = 20$



(b)  $n = 50$



(c)  $n = 100$



(d)  $n = 1000$

## 4 致谢

感谢熊老师上课详细的理论知识的介绍, 感谢老师助教以及同学们课后的答疑解惑。