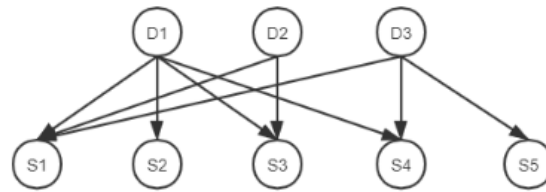9.1

（1） 如下图

（2） $P(D_1, D_2, D_3, S_1, S_2, S_3, S_4, S_5) =$
$P(D_1)P(D_2)P(D_3)P(S_1|D_1, D_2, D_3)P(S_2|D_1)P(S_3|D_1, D_2)P(S_4|D_1, D_3)P(S_5|D_3)$

（3） 23、255。

（4） D_1

（5） D_1, D_2。S_3已知时，V 字结构$D_1 \to S_3 \leftarrow D_2$已激活，此时观测到$S_2 = 1$，增大了$D_1 = 1$的概率，同时使$D_2 = 1$的概率减小

9.2
第(3)小问强调长程相互作用、上下文等要点即可

9.3
(1)

设计的真实值为：初始概率$[0.65, 0.35]$

转移矩阵$[[0.95, 0.05], [0.1, 0.9]]$

发射概率：正常骰子全为 1/6；作弊骰子 6 的概率为 0.5，其他为 0.1

```
start prob: [0.64975571 0.35024429]
trans prob:
[[0.94706525 0.05293475]
 [0.08998176 0.91001824]]
emission prob:
[[0.16412205 0.15644102 0.17805723 0.1736882  0.17438372 0.15330777]
 [0.10439933 0.10223966 0.09475692 0.09505405 0.10343664 0.50011341]]
```

(2)

(12). $e_0(6) = 0.15$, $e_1(6) = 0.50$ $z_0 = 0.65$ $z_1 = 0.35$

$$\begin{bmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{bmatrix} = \begin{bmatrix} 0.95 & 0.05 \\ 0.09 & 0.91 \end{bmatrix}$$

$\alpha_1(0) = 0.65 \times 0.15 = 9.75 \times 10^{-2}$

$\alpha_1(1) = 0.35 \times 0.5 = 17.5 \times 10^{-2}$

$\alpha_2(0) = (9.75 \times 10^{-2} \times 0.95 + 17.5 \times 10^{-2} \times 0.09) \times 0.15 = 1.63 \times 10^{-2}$

$\alpha_2(1) = (9.75 \times 10^{-2} \times 0.05 + 17.5 \times 10^{-2} \times 0.91) \times 0.50 = 8.21 \times 10^{-2}$

$\alpha_3(0) = (1.63 \times 10^{-2} \times 0.95 + 8.21 \times 10^{-2} \times 0.09) \times 0.15 = 0.34 \times 10^{-2}$

$\alpha_3(1) = (1.63 \times 10^{-2} \times 0.05 + 8.21 \times 10^{-2} \times 0.91) \times 0.50 = 3.78 \times 10^{-2}$

由前向算法，结果为 $4.12*10^{-2}$

$\beta_4(0) = \beta_4(1) = 1$

$\beta_3(0) = 1 \cdot 0.95 \cdot 0.15 + 1 \cdot 0.05 \cdot 0.5 = 0.167$

$\beta_3(1) = 1 \cdot 0.09 \cdot 0.15 + 1 \cdot 0.91 \cdot 0.5 = 0.469$

$\beta_2(0) = 0.167 \cdot 0.95 \cdot 0.15 + 0.469 \cdot 0.05 \cdot 0.5 = 3.56 \times 10^{-2}$

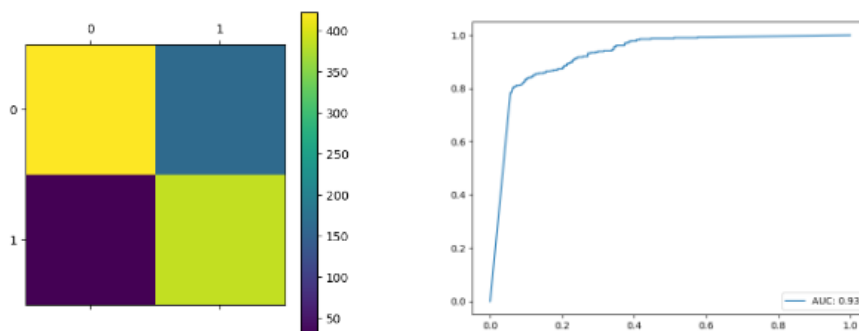$\beta_2(1) = 0.167 \cdot 0.09 \cdot 0.15 + 0.469 \cdot 0.91 \cdot 0.5 = 21.5 \times 10^{-2}$

$P(O) = 3.56 * 10^{-2} * 0.65 * 0.15 + 21.5 * 10^{-2} * 0.35 * 0.5 = 4.11 * 10^{-2}$

在误差范围内结果一致

(3) 第 12 次开始作弊

9.4

测试正确率一般在 0.82 附近，AUC 能达到 0.9 以上。混淆矩阵与 ROC 如下：

ROC 有人是一条折线，即除了（0，0）和（1，1）中间只有一个点，是因为调用函数的时候把分类结果输入进去了，而不是输入的预测的概率值。