

Problem 1:

(1)

$$\text{LHS} = E[ax + bY]$$

$$= \iint p(x, y) [ax + by] dx dy$$

$$= a \iint p(x, y) x dy dx + b \iint p(x, y) y dx dy$$

$$\text{RHS} = aE[X] + bE[Y]$$

$$= a \int p(x) x dx + b \int p(y) y dy$$

Remark: by law of total probability, note X, Y can take continuous

$$\int p(x, y) dy = p(x)$$

$$\int p(x, y) dx = p(y)$$

$$\text{thus RHS} = a \int \int p(x, y) dy x dx + b \iint p(x, y) dx y dy$$

$$= a \int \int p(x, y) x dy dx + b \iint p(x, y) y dx dy$$

$$= \text{LHS}$$

$$\text{thus } E[ax + bY] = aE[X] + bE[Y]$$

(2)

Remark

$$(*) \text{ var}(X) = E[(X - \mu_x)^2] = E[X^2 - 2E[X]X + \mu_x^2]$$

$$\stackrel{\text{by } \textcircled{1}}{=} E[X^2] - 2E[X]E[X] + E[\mu_x^2]$$

$$= E[X^2] + (E[X])^2 - 2E[X]E[X]$$

$$= E[X^2] + E[X]^2 - 2E[X]E[X] = E[X^2] - (E[X])^2$$

Recall: let X, Y be two real-valued random variables

$$(**) \text{ then } E[XY] = \iint p(x, y) xy dx dy \quad \text{since } X, Y \text{ independent}$$

$$= \iint p(x) p(y) xy dx dy$$

$$= \int p(x) x dx \cdot \int p(y) y dy$$

$$= E[X] E[Y]$$

by (*)

method (I)

$$\begin{aligned} \text{Var}(aX + bY) &= E[(aX + bY)^2] - (E[aX + bY])^2 \\ &\stackrel{\text{by (I)}}{=} E[a^2X^2 + 2abXY + b^2Y^2] - (aE[X] + bE[Y])^2 \\ &\stackrel{\text{by (I)}}{=} a^2E[X^2] + 2abE[XY] + b^2E[Y^2] - a^2(E[X])^2 - b^2(E[Y])^2 - 2abE[X]E[Y] \\ &\text{since } X, Y \text{ independent by (**) } 2abE[XY] - 2abE[X]E[Y] \\ &= 2ab(E[XY] - E[X]E[Y]) = 0 \end{aligned}$$

$$\begin{aligned} \text{thus } \text{Var}(aX + bY) &= a^2[E[X^2] - (E[X])^2] + b^2[E[Y^2] - (E[Y])^2] \\ &= a^2 \text{Var}(X) + b^2 \text{Var}(Y) \end{aligned}$$

method (II)

$$\begin{aligned} \text{Var}(aX + bY) &= \iint p(x, y) (ax + by - a\mu_x - b\mu_y)^2 dx dy \quad \text{since } X, Y \text{ independent } p(x, y) = p(x)p(y) \\ &= \iint p(x)p(y) [(ax - a\mu_x) + (by - b\mu_y)]^2 dx dy \\ &= a^2 \int p(x) (x - \mu_x)^2 \int p(y) dy dx \\ &\quad + b^2 \int p(y) (y - \mu_y)^2 \int p(x) dx dy \\ &\quad + \iint p(x)p(y) 2 \cdot [aby - abx\mu_y - a\mu_x y + ab\mu_x\mu_y] dx dy \\ &= a^2 \text{Var}(X) + b^2 \text{Var}(Y) \\ &\quad + 2ab \int x p(x) dx \int y p(y) dy - 2ab \int x p(x) dx \mu_y \int p(y) dy \\ &\quad - 2ab \mu_x \int p(x) dx \int y p(y) dy + 2ab \mu_x \int p(x) dx \mu_y \int p(y) dy \\ &= a^2 \text{Var}(X) + b^2 \text{Var}(Y) \\ &\quad + 2ab \mu_x \mu_y - 2ab \mu_x \mu_y - 2ab \mu_x \mu_y + 2ab \mu_x \mu_y \\ &= a^2 \text{Var}(X) + b^2 \text{Var}(Y) \end{aligned}$$

note

$$\begin{aligned} \int y p(y) dy &= \mu_y \\ \int x p(x) dx &= \mu_x \\ \int p(y) dy &= \int p(x) dx = 1 \end{aligned}$$

Problem 2

①

$$\begin{aligned} E[X] &= \int p(x) x dx = \int_0^a p(x) x dx + \int_a^b p(x) x dx + \int_b^{\infty} p(x) x dx \\ &= \int_0^a 0 \cdot x dx + \int_a^b \frac{1}{b-a} x dx + \int_b^{\infty} 0 \cdot x dx \\ &= \frac{1}{b-a} \cdot \frac{1}{2} x^2 \Big|_a^b \\ &= \frac{1}{b-a} \cdot \frac{1}{2} (b^2 - a^2) \\ &= \frac{b+a}{2} \end{aligned}$$

②

$$\begin{aligned} \text{Var}(X) &= \int p(x) (x - \mu_x)^2 dx = \int_0^a 0 \cdot (x - \mu_x)^2 dx + \int_a^b \frac{1}{b-a} (x - \mu_x)^2 dx + \int_b^{\infty} 0 \cdot (x - \mu_x)^2 dx \\ &= \frac{1}{b-a} \cdot \frac{1}{3} (x - \mu_x)^3 \Big|_a^b \\ &= \frac{1}{b-a} \cdot \frac{1}{3} \cdot \left[\left(\frac{b-a}{2} \right)^3 - \left(\frac{a-b}{2} \right)^3 \right] \\ &= \frac{1}{24} \cdot \frac{1}{b-a} [(b-a)(b-a-a+b)] \quad \text{note } b-a > 0 \\ &= \frac{1}{24} (b-a) (2b-2a) \\ &= \frac{1}{12} (b-a)^2 \end{aligned}$$

Problem 3

$$\text{let } a_n = (1-\theta)^n$$

$$\begin{aligned} \textcircled{1} \quad \sum_{x=1}^{\infty} P(X) &= \sum_{x=1}^{\infty} (1-\theta)^{x-1} \theta = \theta \sum_{x=1}^{\infty} (1-\theta)^{x-1} \quad \left\{ \begin{array}{l} \text{then } \frac{a_{n+1}}{a_n} = \frac{(1-\theta)^{(x+1)}}{(1-\theta)^{(x-1)}} = (1-\theta) \end{array} \right. \\ &\text{by summation formula of geometric progression} \quad \text{where } 0 < 1-\theta < 1 \\ &= \theta \frac{1 - (1-\theta)^n}{1 - (1-\theta)} \quad \text{where } n \rightarrow \infty, (1-\theta)^n \rightarrow 0 \\ &= 1 \end{aligned}$$

$$a_1 = (1-\theta)^0 \theta = \theta$$

$$\begin{aligned} \textcircled{2} \quad E[X] &= \sum_{i=1}^{\infty} i p(i) \\ &= p(1) + 2p(2) + 3p(3) + \dots \\ &= (p(1) + p(2) + p(3) + \dots) + (p(2) + p(3) + \dots) + (p(3) + \dots) + \dots \\ &= P(X \geq 1) + P(X \geq 2) + \dots \end{aligned}$$

$$\text{Note } P(X \geq i) = (1-\theta)^{i-1} \quad i \in \mathbb{Z}^+$$

$$\text{thus } E[X] = \sum_{i=1}^n (1-\theta)^{i-1}$$

$$\begin{aligned} &= \frac{1 - (1-\theta)^n}{1 - (1-\theta)} \quad n \rightarrow \infty \\ &= \frac{1}{\theta} \end{aligned}$$

Since $0 < 1-\theta < 1$

by Problem 1 (2)(*)

$$\textcircled{3} \quad \text{Var}(X) = E[(X - \mu_X)^2] = E[X^2] - (E[X])^2$$

$$\text{let } f(x) = \sum_{k=2}^{\infty} x^k = \frac{1}{1-x}$$

$$\text{then } f'(x) = \sum_{k=2}^{\infty} k x^{k-1} = \frac{1}{(1-x)^2}$$

$$f''(x) = \sum_{k=2}^{\infty} k(k-1) x^{k-2} = \frac{2}{(1-x)^3}$$

$$\text{note that } E[X^2] = \theta \sum_{k=1}^{\infty} k^2 (1-\theta)^{k-1}$$

$$= \theta (1-\theta) \sum_{k=1}^{\infty} k(k-1) (1-\theta)^{k-2} + \theta \sum_{k=1}^{\infty} k (1-\theta)^{k-1}$$

$$= \theta (1-\theta) \frac{2}{\theta^3} + \frac{\theta}{\theta^2} = \frac{2-2\theta}{\theta^2} + \frac{1}{\theta} = \frac{2}{\theta^2} - \frac{1}{\theta}$$

$$\begin{aligned}\text{Var}(X) &= E[X^2] - (E[X])^2 \\ &= \frac{2}{\theta^2} - \frac{1}{\theta} - \frac{1}{\theta^2} = \frac{1}{\theta^2} - \frac{1}{\theta}\end{aligned}$$

Problem 4

1. when independent random feature values $\{X_1, \dots, X_n\}$ are added, their properly normalized sum tends toward a normal distribution, even if the original feature value X_i themselves are not normally distributed
2. see last page of this homework
3. As n gets larger, the figure will be more like normal distribution bell-like shape.
As n gets larger the variance decreases which is agree with the fact that $\text{var} = \frac{\sigma^2}{n}$ where σ is a finite variance of the feature of uniform distribution
4. see last page of this homework

Problem 5

$$p(x) = \sum_{k=1}^K \alpha_k p_k(x; \theta_k) \quad \text{where} \quad \sum_{k=1}^K \alpha_k = 1, \quad 0 \leq \alpha_k \leq 1$$

each $p_k(x; \theta_k)$ is pdf $\theta_k = \{\mu_k, \sigma_k^2\}$

(1)

sum to 1:
$$\int p(x) dx = \int \sum_{k=1}^K \alpha_k p_k(x; \theta_k) dx$$

$$= \sum_{k=1}^K \alpha_k \int p_k(x; \theta_k) dx = \sum_{k=1}^K \alpha_k \cdot 1 = 1$$

since each $p_k(x; \theta_k)$ is pdf
then $\int p_k(x; \theta_k) dx = 1$

nonnegative since each $p_k(x; \theta_k)$ is pdf then $p_k(x; \theta_k) \geq 0$ for $\forall x$

since $\alpha_k \in [0, 1]$ for $\forall k$

then $\forall k, \alpha_k p_k(x; \theta_k) \geq 0$ then $p(x) = \sum_{k=1}^K \alpha_k p_k(x; \theta_k) \geq 0$

(2) (a) $\mu = E[X] = \int x p(x) dx = \int x \sum_{k=1}^K \alpha_k p_k(x; \theta_k) dx$

$$= \sum_{k=1}^K \alpha_k \int x p_k(x; \theta_k) dx$$

$$= \sum_{k=1}^K \alpha_k \mu_k$$

(b) $E[X^2] = \int x^2 p(x) dx = \sum_{k=1}^K \alpha_k \int x^2 p_k(x; \theta_k) dx$

$$= \sum_{k=1}^K \alpha_k [\text{var}_k[X] + (E_k[X])^2]$$

$$= \sum_{k=1}^K \alpha_k (\sigma_k^2 + \mu_k^2)$$

$$(E[X])^2 = \left(\sum_{k=1}^K \alpha_k \mu_k \right)^2$$

$$\sigma^2 = \text{var}(X) = E[X^2] - (E[X])^2 = \sum_{k=1}^K \alpha_k (\sigma_k^2 + \mu_k^2) - \left(\sum_{k=1}^K \alpha_k \mu_k \right)^2$$

Problem 6

$$\Sigma = \begin{pmatrix} \sigma_1^2 & r\sigma_1\sigma_2 \\ r\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

$$(1) \quad 0 < c = p(\underline{x}) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} (\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})}$$

$$\text{let } k = (2\pi)^{d/2} |\Sigma|^{1/2} > 0$$

$$\text{then } c k = e^{-\frac{1}{2} (\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})}$$

$$-2 \log(ck) = (\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu}) = (x_1-\mu_1, x_2-\mu_2) \frac{\begin{pmatrix} \sigma_2^2 & -r\sigma_1\sigma_2 \\ -r\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}}{|\Sigma|} \begin{pmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{pmatrix}$$

$$0 < -2|\Sigma| \log(ck) = \begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} \sigma_2^2 & -r\sigma_1\sigma_2 \\ -r\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{where let } \begin{matrix} a = x_1 - \mu_1 \\ b = x_2 - \mu_2 \end{matrix}$$

$$= (a\sigma_2^2 - br\sigma_1\sigma_2, -ar\sigma_1\sigma_2 + b\sigma_1^2) \begin{pmatrix} a \\ b \end{pmatrix}$$

$$= a^2\sigma_2^2 - abr\sigma_1\sigma_2 - abr\sigma_1\sigma_2 + b^2\sigma_1^2$$

$$= \sigma_2^2(x_1-\mu_1)^2 - 2\sigma_1\sigma_2r(x_1-\mu_1)(x_2-\mu_2) + \sigma_1^2(x_2-\mu_2)^2 = (**)$$

by <https://en.wikipedia.org/wiki/Ellipse>

the (**) is the ellipse $\sigma_2^2(x_1-\mu_1)^2 + \sigma_1^2(x_2-\mu_2)^2 - 2\sigma_1\sigma_2r(x_1-\mu_1)(x_2-\mu_2) = -2|\Sigma| \log(ck) > 0$

rotate $\arctan\left(\frac{1}{-2\sigma_1\sigma_2r} (\sigma_1^2 - \sigma_2^2 - \sqrt{(\sigma_2^2 - \sigma_1^2)^2 + (2\sigma_1\sigma_2r)^2})\right)$ counter clock wise

(2) see last page

Problem 7

(1)

```

for i=1:floor(N/2)
    U1 vector = Uniform(0,1) * ones(1, floor(d/2))
    U2 vector = Uniform(0,1) * ones(1, floor(d/2))
    for i=1:length(U1 vector)
        Z[i, 2i+1] = box-muller(U1 vector(i), U2 vector(i)) % which give 2*floor(N/2) normal distribution numbers N(0,1)
    end
    A = chol(Z)

    Z = Z(1:N)
    X = A*Z + mu
output X

```

(2)

get box-muller involves multiplication $O(d^2)$
 cholesky involves getting eigenvalues matrix $O(d^3)$
 getting X need matrix multiply for N values $O(Nd)$
 $O(Nd^2 + d^3)$

(3) since \underline{Z} is ^{multi} normal distribution and $A\underline{Z} + \underline{\mu}$ is linear transformation then \underline{X} is also normal distribution.

$$E[X] \stackrel{\text{by Problem 1}}{=} A E[Z] + \underline{\mu} = A \cdot 0 + \underline{\mu} = \underline{\mu}$$

$$\text{Var}[X] = E[(X - \underline{\mu})(X - \underline{\mu})^T] = E[(A\underline{Z})^T(A\underline{Z})] = E[A\underline{Z}\underline{Z}^T A^T] = E[A] \cdot E[\underline{Z}\underline{Z}^T] \cdot E[A^T]$$

note $E[\underline{Z}\underline{Z}^T] = \begin{pmatrix} E[t_1^2] & E[t_1 t_2] & & \\ E[t_2 t_1] & E[t_2^2] & & \\ & & \ddots & \\ E[t_n t_1] & E[t_n t_2] & & E[t_n^2] \end{pmatrix}$

note $E[t_i^2] = \text{Var}(t_i) + E[t_i]^2$
 $= 1 + 0 = 1$
 independent
 $E[t_i t_j] = E[t_i] E[t_j] = 0 \cdot 0 = 0$
 $i \neq j$

thus $E[\underline{Z}\underline{Z}^T] = I$

thus $\text{Var}[X] = E[A] I E[A^T] = E[A A^T] = \Sigma$

Problem 8

$$\textcircled{1} \quad P(C=1|X) = \frac{1}{1 + e^{c \cdot d_0 - d_1^T X}} \quad (\Rightarrow) \quad \log \frac{P(C=1|X)}{P(C=0|X)} \text{ is linear in } X$$

$$P(C=0|X) = 1 - P(C=1|X) = \frac{e^{c \cdot d_0 - d_1^T X}}{1 + e^{c \cdot d_0 - d_1^T X}}$$

$$\text{then } \log \frac{P(C=1|X)}{P(C=0|X)} = \log \frac{1/(1 + e^{c \cdot d_0 - d_1^T X})}{e^{c \cdot d_0 - d_1^T X}/(1 + e^{c \cdot d_0 - d_1^T X})} = \log e^{d_0 + d_1^T X} = d_0 + d_1^T X$$

$$P(C=1|X) = \frac{1}{1 + e^{c \cdot d_0 - d_1^T X}} \quad (\Leftarrow) \quad \log \frac{P(C=1|X)}{P(C=0|X)} \text{ is linear in } X$$

$$\begin{aligned} (\Leftarrow) \quad P(C=1|X) &= \frac{P(X|C=1) P(C=1)}{P(X|C=1) P(C=1) + P(X|C=0) P(C=0)} \\ &= \frac{1}{1 + \frac{P(X|C=0) P(C=0)}{P(X|C=1) P(C=1)}} \\ &= \frac{1}{1 + \frac{P(C=0)}{P(C=1)} e^{d_0 + d_1^T X}} \end{aligned}$$

$$\text{wlog let } e^{d_0} = \frac{P(C=0)}{P(C=1)} e^{d_1}$$

$$= \frac{1}{1 + e^{d_0 + d_1^T X}}$$

② let $\Sigma^{-1} = (k_{ij}) = \begin{pmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ k_{21} & & \ddots & \\ & & & k_{nn} \end{pmatrix}$

$$(X - \mu_2)^t = (X^1 - \mu_2^1, X^2 - \mu_2^2, \dots, X^n - \mu_2^n)$$

$$(X - \mu_2) = \begin{pmatrix} X^1 - \mu_2^1 \\ \vdots \\ X^n - \mu_2^n \end{pmatrix}$$

$$(X - \mu_2)^t \Sigma^{-1} = \left(\sum_{i=1}^n k_{i1} (X^i - \mu_2^i), \dots, \sum_{i=1}^n k_{in} (X^i - \mu_2^i) \right)$$

$$\begin{aligned} \textcircled{1} = (X - \mu_2)^t \Sigma^{-1} (X - \mu_2) &= \sum_{j=1}^n \left\{ (X^j - \mu_2^j) \sum_{i=1}^n k_{ij} (X^i - \mu_2^i) \right\} \\ &= \sum_{j=1}^n (X^j - \mu_2^j) \left(\sum_{i=1}^n k_{ij} X^i - \sum_{i=1}^n k_{ij} \mu_2^i \right) \\ &= \sum_{j=1}^n X^j \sum_{i=1}^n k_{ij} X^i - \sum_{j=1}^n X^j \sum_{i=1}^n k_{ij} \mu_2^i - \sum_{j=1}^n \mu_2^j \sum_{i=1}^n k_{ij} X^i + \sum_{j=1}^n \mu_2^j \sum_{i=1}^n k_{ij} \mu_2^i \\ &= \sum_{j=1}^n X^j \sum_{i=1}^n k_{ij} X^i - \sum_{j=1}^n X^j \sum_{i=1}^n k_{ij} \mu_2^i - \sum_{j=1}^n \mu_2^j \sum_{i=1}^n k_{ij} X^i + \sum_{j=1}^n \mu_2^j \sum_{i=1}^n k_{ij} \mu_2^i \end{aligned}$$

$$= X^t \Sigma^{-1} X - \mu_2^t \Sigma^{-1} X - \mu_2^t \Sigma^{-1} X + \mu_2^t \Sigma^{-1} \mu_2$$

$$\textcircled{2} = (X - \mu_1)^t \Sigma^{-1} (X - \mu_2) = X^t \Sigma^{-1} X - \mu_1^t \Sigma^{-1} X - \mu_1^t \Sigma^{-1} X + \mu_1^t \Sigma^{-1} \mu_2$$

$$\begin{aligned} \textcircled{1} - \textcircled{2} &= -2\mu_2^t \Sigma^{-1} X - 2\mu_1^t \Sigma^{-1} X + \mu_1^t \Sigma^{-1} \mu_1 - \mu_2^t \Sigma^{-1} \mu_2 \\ &= (-2\mu_2^t \Sigma^{-1} + 2\mu_1^t \Sigma^{-1}) X + \mu_1^t \Sigma^{-1} \mu_1 - \mu_2^t \Sigma^{-1} \mu_2 \end{aligned}$$

thus

$$P(X|C=1) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-1/2 (X-\mu_1)^t \Sigma^{-1} (X-\mu_1)}$$

$$P(X|C=0) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-1/2 (X-\mu_2)^t \Sigma^{-1} (X-\mu_2)}$$

$$\frac{P(C=1|X)}{P(C=0|X)} = \frac{P(X|C=1) P(C=1)}{P(X|C=0) P(C=0)}$$

$$\text{thus } \log \frac{P(C=1|X)}{P(C=0|X)} = \log \frac{P(C=1)}{P(C=0)} - \log \left(\frac{e^{-1/2 (X-\mu_1)^t \Sigma^{-1} (X-\mu_1)}}{e^{-1/2 (X-\mu_2)^t \Sigma^{-1} (X-\mu_2)}} \right)$$

$$\log \frac{P(C=1|X)}{P(C=0|X)} = -\frac{1}{2} [(X-\mu_1)^t \Sigma^{-1} (X-\mu_1) - (X-\mu_2)^t \Sigma^{-1} (X-\mu_2)] + \log \left(\frac{P(C=1)}{P(C=0)} \right)$$

by page before

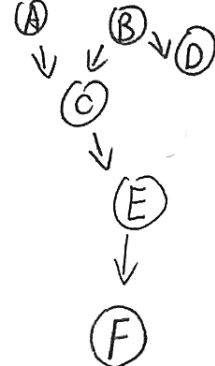
$$= \underbrace{(\mu_2^t \Sigma^{-1} - \mu_1^t \Sigma^{-1}) X}_{\text{let } = d^t} - \frac{1}{2} \underbrace{\mu_1^t \Sigma^{-1} \mu_1 + \frac{1}{2} \mu_2^t \Sigma^{-1} \mu_2}_{\text{let } = d_0} + \log \left(\frac{P(C=1)}{P(C=0)} \right)$$

$$= d^t X + d_0$$

by ① $P(C=0|X)$ is in the form of a logistic function

Problem 9

(1)



$$P(a,b,c,d,e,f) = P(f|e) P(e|c) P(d|b) P(c|a,b) P(b|a)$$

(2) for a, b (need $k-1$)

for c need $(k-1)(k^2)$ since 2 parent

for d 1 parent $(k-1)k$

for e 1 parent $(k-1)k$

for f 1 parents $(k-1)k$

$$\text{parameter} = (k-1)k^2 + 3(k-1)k + 2(k-1)$$

(3) $k^6 - 1 =$

$$(4) P(f_k | a^*, d^*) = \sum_{a,b,c,d,e} P(a,b,c,d,e,f | a^*, d^*)$$

$$= \sum_{a,b,c,d,e} P(f|e) P(e|c) P(d^*|b) P(c|a^*, b) P(b|a^*)$$

$$= P(a^*) \sum_{b,c,d,e} P(f|e) P(e|c) P(d^*|b) P(c|a^*, b) P(b)$$

$$= P(a^*) \sum_{b,c,e} P(f|e) P(e|c) P(c|a^*, b) g_1(c,b), \text{ where } g_1(c,b) = P(b|a^*)$$

$$= P(a^*) \sum_{c,e} P(f|e) P(e|c) \sum_b P(c|a^*, b) g_1(c,b)$$

$$= P(a^*) \sum_{c,e} P(f|e) P(e|c) g_2(c,e)$$

$$= P(a^*) \sum_e P(f|e) \sum_c P(e|c) g_2(c,e)$$

$$= P(a^*) \sum_e P(f|e) g_3(c,e)$$

$$= P(a^*) g_4(c,e)$$

$$\text{where } g_2(c) = \sum_b P(c|a^*, b) g_1(c,b)$$

$$\text{where } g_3(c) = \sum_e P(e|c) g_2(c,e)$$

$$\text{where } g_4(c) = \sum_e P(f|e) g_3(c,e)$$

5. graphical model

$$K^2 + K^2 + K^2 + K^2 = 4K^2$$

$O(nK^2)$ where n is unobserved nodes

saturated model

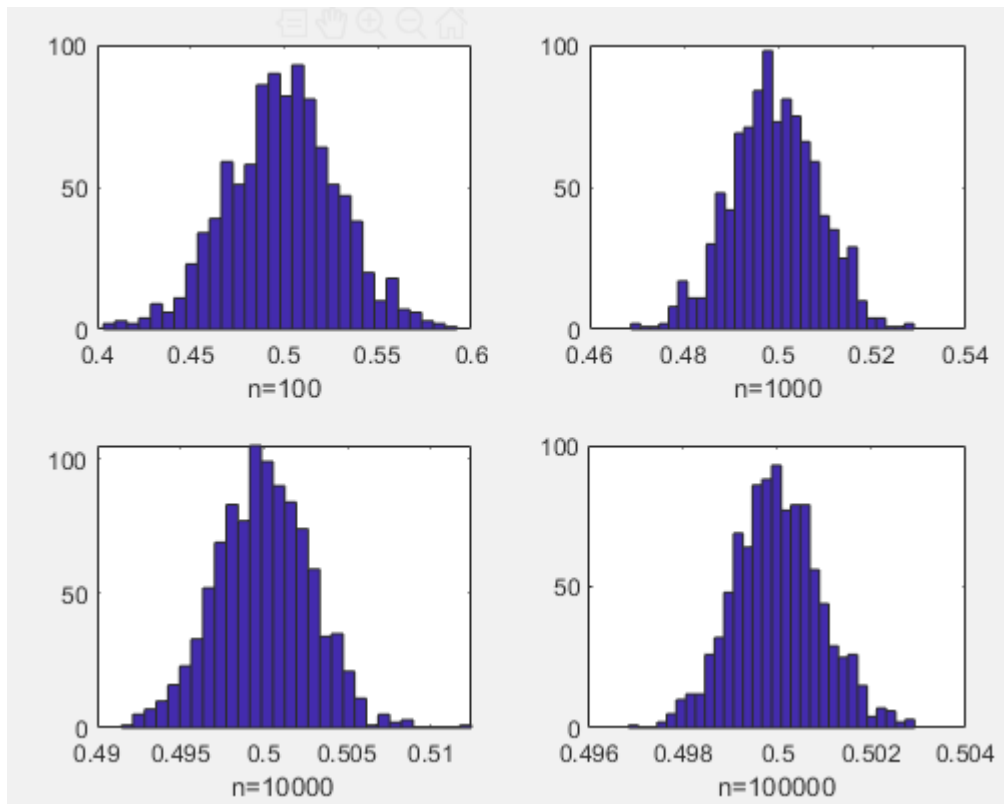
$$O(K^n)$$

6.

$$O(K^3) \text{ since } \sum_{a,b} P(c|a,b) P(a) P(b) P(d^*|b)$$

will have $(K+1)K^2 = K^3$ complexity

Problem 4



```
>> cs274ahwlprob4
      n = 100      n = 1000      n= 10000      n= 100000

mean_lst =

    0.500315260996849    0.499742398791919    0.499985623765814    0.500038984282360

var_lst =

    0.029361103601407    0.009081463005706    0.002908322176448    0.000879601183466

error_mean =

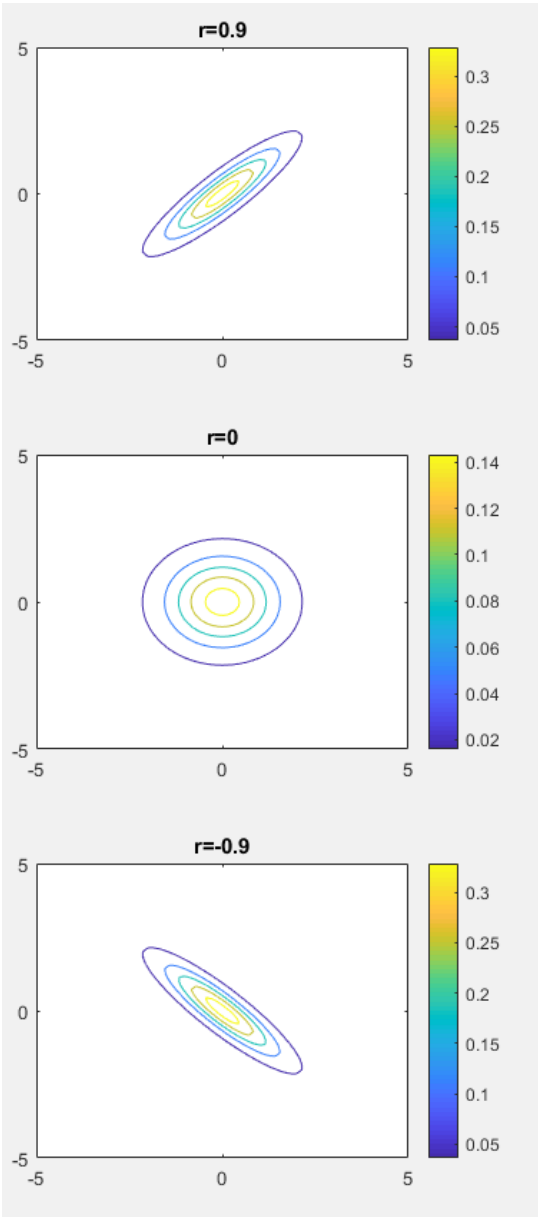
    1.0e-03 *

    0.315260996848643    0.257601208080827    0.014376234185931    0.038984282359600

error_var =

    0.029352770268074    0.009081379672373    0.002908321343114    0.000879601175132
```

Problem 6



The below are the P values respect to the isocontour.

	yellow	light green	deep green	light blue	deep blue
Figure 1	0.328613832616992	0.255588536479883	0.182563240342773	0.109537944205664	0.036512648068555
Figure 2	0.143239448782706	0.111408460164327	0.079577471545948	0.047746482927569	0.015915494309190
Figure 3	0.328613832616992	0.255588536479883	0.182563240342773	0.109537944205664	0.036512648068555