1.4.31

(a) It's known that $Y|X=x \sim B(n, x)$, $X \sim \mathcal{U}[0,1]$

So $\mathbb{E}Y = \mathbb{E}[\mathbb{E}(Y|X)] = \mathbb{E}[nx] = n\mathbb{E}X = \frac{n}{2}$

$$\begin{aligned}
\operatorname{Var}(Y) &= \mathbb{E}[\operatorname{var}(Y|X)] + \operatorname{var}(\mathbb{E}(Y|X)) \\
&= \mathbb{E}[nX(1-X)] + \operatorname{var}(nX) \\
&= n\mathbb{E}X - n\mathbb{E}X^2 + n^2\operatorname{var}(X) \\
&= n\mathbb{E}X - n(\operatorname{var}(X) + (\mathbb{E}X)^2) + n^2\operatorname{var}(X) \\
&= \frac{n}{2} - n\left(\frac{1}{12} + \frac{1}{4}\right) + n^2 \cdot \frac{1}{12} \\
&= \frac{n}{6} + \frac{n^2}{12}
\end{aligned}$$

(b) Here, Y is discrete and X is continuous.

We consider the probability $\mathbb{P}[Y \in A, X \in B]$

By the law of total probability.

$$\begin{aligned}
\mathbb{P}[Y \in A, X \in B] &= \int_B \mathbb{P}[Y \in A | X = x] f_X(x)\, dx \\
&= \int_B \left(\sum_{y \in A} \mathbb{P}[Y = y | X = x] \right) f_X(x)\, dx \\
&= \sum_{y \in A} \int_B \mathbb{P}[Y = y | X = x] f_X(x)\, dx \qquad (*)
\end{aligned}$$

Therefore, we can use the function $\mathbb{P}[Y=y|X=x] f_X(x)$ to describe the joint distribution of (X, Y). The probability on any given set of (X, Y) can be calculated by equation $(*)$.

So, in this question, the joint distribution of (X, Y) is $\mathbb{P}[Y=y|X=x] f_X(x) = \binom{n}{y} x^y (1-x)^{n-y}$, $0 < x < 1$, $y = 0, 1, \cdots n$

(c) $\mathbb{P}[Y=y] = \mathbb{P}[Y=y, X\in\mathbb{R}]$

$= \mathbb{P}[Y=y, 0<X<1]$

By (b), we have

$$\mathbb{P}[Y=y, X\in[0,1]] = \int_0^1 \binom{n}{y} x^y (1-x)^{n-y}\,dx$$

$$= \binom{n}{y} \text{Beta}(y+1, n-y+1)$$

$$= \frac{n!}{(n-y)!\,y!} \cdot \frac{y!\,(n-y)!}{(n+1)!}$$

$$= \frac{1}{n+1}$$

1.4.35

(a) $X|P \sim \text{binomial}(P)$, $i=1\cdots n$

$P \sim \text{beta}(\alpha, \beta)$

We have, $\text{var}(X) = \mathbb{E}(\text{var}(X|P)) + \text{var}(\mathbb{E}(X|P))$ (conditional variance identity)

$$= \mathbb{E}(nP(1-P)) + \text{var}(nP)$$

$$= n\mathbb{E}P - n\mathbb{E}P^2 + n^2\,\text{var}(P)$$

$$= n\mathbb{E}P(1-\mathbb{E}P) + n(\mathbb{E}P)^2 - n\mathbb{E}P^2 + n^2\,\text{var}(P)$$

$$= n\mathbb{E}P(1-\mathbb{E}P) - n\,\text{var}(P) + n^2\,\text{var}(P)$$

$$= n\mathbb{E}P(1-\mathbb{E}P) + n(n-1)\,\text{var}(P)$$

(b) $Y|\Lambda \sim \text{Poisson}(\Lambda)$, $\Lambda \sim \Gamma(\alpha, \beta)$

So $\text{var}(Y) = \mathbb{E}(\text{var}(Y|\Lambda)) + \text{var}(\mathbb{E}(Y|\Lambda))$

$$= \mathbb{E}(\Lambda) + \text{var}(\Lambda)$$

$$= \mathbb{E}\Lambda + \frac{\alpha}{\beta^2}$$

$$= \frac{\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)^2 \cdot \frac{1}{\alpha} = \mu + \frac{1}{\alpha}\mu^2$$

2. 4.39

(a) By defition 4.6.2

$$\mathbb{P}(X_1 = x_1 \cdots X_n = x_n) = \frac{m!}{x_1! \cdots x_n!} p_1^{x_1} \cdots p_n^{x_n}, \quad \sum_i p_i = 1, \ \sum_i x_i = m, \ x_i \geq 0$$

$\forall j$, when $X_j = x_j$, by $\sum_i x_i = m$, we have $\sum_{i \neq j} x_i = m - x_j$.

Note that $\mathbb{P}[X_1 = x_1 \cdots X_n = x_n]$

$$= \frac{m!}{(m-x_j)!\, x_j!} p_j^{x_j} \frac{(m-x_j)!}{x_1! \cdots x_{j-1}!\, x_{j+1}! \cdots x_n!} p_1^{x_1} \cdots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \cdots p_n^{x_n}$$

So
$$\mathbb{P}[X_j = x_j] = \sum_{\substack{x_1 \cdots x_{j-1}, x_{j+1}, \cdots x_n \\ : \sum_{i \neq j} x_i = m - x_j}} \mathbb{P}[X_1 = x_1 \cdots X_n = x_n]$$

$$= \frac{m!}{(m-x_j)!\, x_j!} p_j^{x_j} \sum_{\substack{x_1 \cdots x_{j-1}, x_{j+1}, \cdots x_n \\ : \sum_{i \neq j} x_i = m - x_j}} \frac{(m-x_j)!}{x_1! \cdots x_{j-1}!\, x_{j+1}! \cdots x_n!} p_1^{x_1} \cdots p_{j-1}^{x_{j-1}} p_{j+1}^{x_{j+1}} \cdots p_n^{x_n}$$

$$\begin{pmatrix}\text{multinomial} \\ \text{Theorem}\end{pmatrix} = \frac{m!}{(m-x_j)!\, x_j!} \cdot \left(\sum_{i \neq j} p_i\right)^{m-x_j}$$

$$\left(\sum_i p_i = 1\right) = \frac{m!}{(m-x_j)!\, x_j!} p_j^{x_j} (1 - p_j)^{m-x_j}$$

$\sim$ binomial (m, p_j)

Similarly, $\forall i \neq j$, the joint distribution of (X_i, X_j) is

$\mathbb{P}[X_i = x_i, X_j = x_j]$

$$= \frac{m!}{(m-x_i-x_j)!\, x_i!\, x_j!} p_i^{x_i} p_j^{x_j} \left(\sum_{k \neq i,j} p_k\right)^{m-x_i-x_j}$$

$$= \frac{m!}{(m-x_i-x_j)!\, x_i!\, x_j!} p_i^{x_i} p_j^{x_j} (1 - p_i - p_j)^{m-x_i-x_j}$$

So $\mathbb{P}[X_i|X_j = x_j] = \dfrac{\mathbb{P}[X_i = x_i,\ X_j = x_j]}{\mathbb{P}[X_j = x_j]}$

$$= \frac{\dfrac{m!}{(m-x_i-x_j)!\, x_i!\, x_j!} p_i^{x_i} p_j^{x_j} (1-p_i-p_j)^{m-x_i-x_j}}{\dfrac{m!}{(m-x_j)!\, x_j!} p_j^{x_j} (1-p_j)^{m-x_j}}$$

$$= \frac{(m-x_j)!}{(m-x_j-x_i)!\, x_i!} \left(\frac{p_i}{1-p_j}\right)^{x_i} \left(\frac{1-p_j-p_i}{1-p_j}\right)^{m-x_j-x_i}$$

$\sim$ binomial $\left(m-x_j,\ \dfrac{p_i}{1-p_j}\right)$

(b) From $X_i|X_j=x_j \sim$ binomial $\left(m-x_j,\ \dfrac{p_i}{1-p_j}\right)$

we have $\mathbb{E}(X_i|X_j) = \dfrac{(m-X_j)p_i}{1-p_j}$ $\quad mp_j(1-p_j)+m^2p_j^2$

From $X_j \sim$ binomial (m, p_j), $\mathbb{E}X_j = mp_j$, $\mathbb{E}X_j^2 = mp_j(1-p_j)+m^2p_j^2$

So $\mathbb{E}(X_iX_j) = \mathbb{E}(\mathbb{E}(X_iX_j|X_j)) = \mathbb{E}(X_j \cdot \mathbb{E}(X_i|X_j))$

$$= \mathbb{E}\left[X_j \cdot \frac{(m-X_j)p_i}{1-p_j}\right]$$

$$= \frac{p_i}{1-p_j} m \cdot \mathbb{E}X_j - \frac{p_i}{1-p_j}\mathbb{E}X_j^2$$

$$= \frac{p_i}{1-p_j} m \cdot mp_j - \frac{p_i}{1-p_j}(mp_j(1-p_j)+m^2p_j^2)$$

Therefore $\operatorname{Cov}(X_i, X_j) = \mathbb{E}X_iX_j - \mathbb{E}X_i\mathbb{E}X_j$

$$= \frac{p_i}{1-p_j} m \cdot mp_j - \frac{p_i}{1-p_j}(mp_j(1-p_j)+m^2p_j^2) - mp_i mp_j$$

$$= \frac{p_ip_j}{1-p_j} m^2 - mp_ip_j - \frac{p_ip_j^2}{1-p_j} m^2 - m^2p_ip_j$$

$$= -mp_ip_j + \frac{m^2}{1-p_j}\left(p_ip_j - p_ip_j^2 - p_ip_j(1-p_j)\right)$$

$$= -mp_ip_j$$

2.4.42

By Theorem 4.2.0, we have

$$\mathbb{E}(g(X)\cdot h(Y)) = \mathbb{E}(g(X))\,\mathbb{E}(h(Y))$$

Thus. $\operatorname{cov}(XY, Y)$

$$= \mathbb{E}(XY\cdot Y) - \mathbb{E}(XY)\mathbb{E}(Y)$$

$$= \mathbb{E}(X\cdot Y^2) - \mathbb{E}(X)\mathbb{E}(Y)\mathbb{E}(Y)$$

$$= \mathbb{E}X\left(\mathbb{E}Y^2 - (\mathbb{E}Y)^2\right) = \mathbb{E}X\left(\operatorname{var}(Y)\right)$$

$$= \mu_X \sigma_Y^2$$

$$\operatorname{var}(XY) = \mathbb{E}(X^2Y^2) - (\mathbb{E}XY)^2$$

$$= (\mathbb{E}X^2)(\mathbb{E}Y^2) - (\mathbb{E}X)^2(\mathbb{E}Y)^2$$

$$= (\sigma_X^2 + \mu_X^2)(\sigma_Y^2 + \mu_Y^2) - \mu_X^2\mu_Y^2$$

$$= \sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2$$

$$\operatorname{var}(Y) = \sigma_Y^2$$

Therefore. $\rho_{XY,Y} = \dfrac{\operatorname{cov}(XY, Y)}{\sqrt{\operatorname{var}(XY)\cdot\operatorname{var}(Y)}}$

$$= \frac{\mu_X\sigma_Y^2}{(\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2)^{\frac{1}{2}}\,\sigma_Y}$$

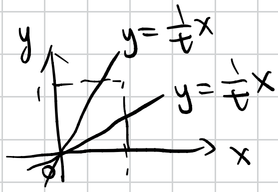
$$= \frac{\mu_X\sigma_Y}{(\sigma_X^2\sigma_Y^2 + \sigma_X^2\mu_Y^2 + \sigma_Y^2\mu_X^2)^{\frac{1}{2}}}$$

3. 4.51

(a) Since X, Y, Z are independent uniform $(0,1)$ r.v.

we have $f_{X,Y}(x,y) = 1, \quad 0 \le x \le 1, \ 0 \le y \le 1$.

Consider the region $X/Y \le t \Rightarrow Y \ge \frac{1}{t}X$



① $0 < t < 1$

$$\mathbb{P}[X/Y \le t] = \mathbb{P}[Y \ge \tfrac{1}{t}X]$$

$$= \text{area}(\text{Triangle}\{(0,0),(0,1),(0,t)\})$$

$$= \frac{t}{2}$$

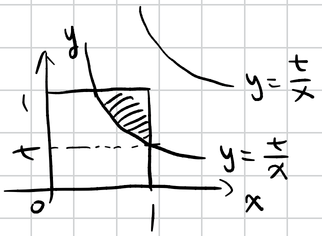
② $t \ge 1$. $\mathbb{P}[X/Y \le t] = \mathbb{P}[Y \ge \frac{1}{t}X]$

$$= 1 - \text{area}(\text{Triangle}\{(0,0),(1,0),(0,\tfrac{1}{t})\})$$

$$= 1 - \frac{1}{2t}$$

③ $t \le 0$, obviously. $\mathbb{P}[X/Y \le t] = 0$

Consider the region $XY \le t \Rightarrow Y \le \frac{1}{X}t$



① $0 \le t \le 1$

$$\mathbb{P}[XY \le t] = 1 - \text{area}(\text{"shaded"})$$

$$= 1 - \int_t^1 \int_{\frac{t}{x}}^1 1 \, dy \, dx$$

$$= 1 - \int_t^1 1 - \frac{t}{x} \, dx$$

$$= 1 - (1-t) - t \ln x \Big|_t^1$$

$$= t - t\ln t$$

② $t < 0$, obviously. $\mathbb{P}[XY \le t] = 0$

③ $t > 1$, obviously. $\mathbb{P}[XY \le t] = 1$

(b) $\mathbb{P}[XY/Z \le t] = \mathbb{P}[XY \le Zt]$ (law of total probability)

$$= \int_0^1 \mathbb{P}[XY \le Zt \mid Z = z]\, f_Z(z)\, dz$$

Since X, Y, Z are independent, we have $XY \mid Z \overset{d}{=} XY$

then $\mathbb{P}[XY \le Zt \mid Z = z] = \mathbb{P}[XY \le zt]$, $0 < z < 1$

① $t < 0$. obviously, $\mathbb{P}[XY/Z \le t] = 0$

② $0 \le t < 1$, $\mathbb{P}[XY/Z \le t] = \int_0^1 \mathbb{P}[XY \le zt]\, dz$

$$= \int_0^1 zt - zt \ln zt \, dz \qquad (\forall\, 0 < z < 1,\ 0 < zt < 1)$$

$$= \frac{t}{2} z^2 \Big|_0^1 - t\left(-\frac{z^2}{4} + \frac{1}{2} z^2 \ln zt\right)\Big|_0^1$$

$$= \frac{3}{4} t - \frac{1}{2} t \ln t$$

③ $t \geq 1$ $\mathbb{P}[XY/Z \le t] = \int_0^1 \mathbb{P}[XY \le zt]\, dz$

$$= \int_0^{\frac{1}{t}} \mathbb{P}[XY \le zt]\, dz + \int_{\frac{1}{t}}^1 \mathbb{P}[XY \le zt]\, dz$$

$$= \frac{t}{2} z^2 \Big|_0^{\frac{1}{t}} - t\left(-\frac{z^2}{4} + \frac{1}{2} z^2 \ln zt\right)\Big|_0^{\frac{1}{t}}$$

$$+ \int_{\frac{1}{t}}^1 1 \cdot dz$$

$$= \frac{1}{2t} + \frac{1}{4t} + 1 - \frac{1}{t}$$

$$= 1 - \frac{1}{4t}$$

By ①②③.

$$\mathbb{P}[XY/Z \le t] = \begin{cases} 0, & t < 0 \\ \frac{3}{4} t - \frac{1}{2} t \ln t, & 0 \le t < 1 \\ 1 - \frac{1}{4t}, & 1 \le t \end{cases}$$

3. 4.52

Let the first hit point be (X_1, Y_1), the second hit point be (X_2, Y_2).

The distance between (X_1, Y_1), (X_2, Y_2) is

$$\left[(X_1-X_2)^2+(Y_1-Y_2)^2\right]^{\frac{1}{2}} = R, \quad R>0$$

Note that, $X_1, X_2, Y_1, Y_2 \overset{iid}{\sim} N(0,1)$, then

$X_1 - X_2 \sim N(0,2)$, $Y_1 - Y_2 \sim N(0,2)$

and $X_1 - X_2 \perp\!\!\!\perp Y_1 - Y_2$. This is guaranteed by the property of normal distribution.

Further, $\frac{1}{\sqrt{2}}(X_1-X_2) \sim N(0,1)$, $\frac{1}{\sqrt{2}}(Y_1-Y_2) \sim N(0,1)$.

By the definition of chi-squared distribution

$$\left[\frac{1}{\sqrt{2}}(X_1-X_2)\right]^2 + \left[\frac{1}{\sqrt{2}}(Y_1-Y_2)\right]^2 \sim \chi^2(2)$$

i.e. $\frac{1}{2}R^2 \sim \chi^2(2)$

By the density function of $\chi^2(2)$, it's easy to get

$$f_R(r) = \left|\frac{dq}{dr}\right| f_Q(q) = r \cdot \frac{1}{2^1 \Gamma(1)}\left(\frac{1}{2}r^2\right)^0 e^{-\frac{1}{4}r^2}$$

$$= \frac{1}{2} r e^{-\frac{1}{4}r^2}, \quad r \geq 0$$

This is a Rayleigh distribution with $\sigma^2 = 2$.

4. 4.53

The event Ax^2+Bx+C has real roots is equivalent to

$$\Delta = B^2-4AC \geq 0 \iff B^2 \geq 4AC \iff 2\log B \geq \log 4 + \log A + \log C$$

Note that $A, B, C \overset{iid}{\sim} \mathcal{U}[0,1]$, so we have

$-\log A, -\log B, -\log C \overset{iid}{\sim} \text{Exponential}(1)$

Further, by the relationship between gamma and exponential distribution, we have

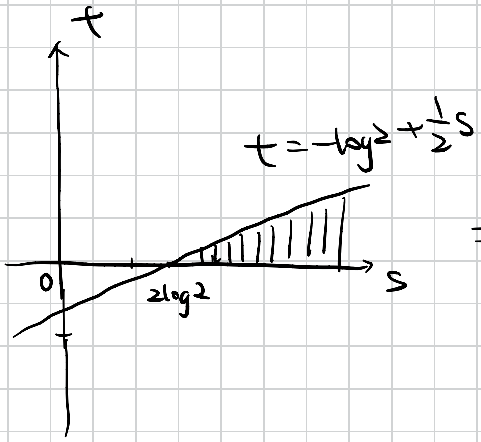
$$-\log A - \log C \sim \text{gamma}(2,1)$$

Let $S = -\log A - \log C$, $T = -\log B$. since A, B, C are independent, S, T are independent as well.

Then
$$f_{S,T}(s,t) = \frac{1}{\Gamma(2)} s e^{-s} \cdot \frac{1}{\Gamma(1)} e^{-t}$$
$$= s e^{-(s+t)}, \quad s \geq 0,\ t \geq 0$$

So, $\mathbb{P}[\text{real roots}] = \mathbb{P}[2\log B \geq \log 4 + \log A + \log C]$

$$= \mathbb{P}[-2\log B \leq -\log 4 - \log A - \log C]$$
$$= \mathbb{P}[2T \leq -\log 4 + S]$$
$$= \mathbb{P}[T \leq -\log 2 + \tfrac{1}{2}S]$$



$$= \int_{2\log 2}^{+\infty} ds \int_0^{-\log 2 + \frac{1}{2}s} f_{S,T}(s,t)\, dt$$
$$= \int_{2\log 2}^{+\infty} s e^{-s} ds \int_0^{-\log 2 + \frac{1}{2}s} e^{-t}\, dt$$
$$= \int_{2\log 2}^{+\infty} s e^{-s} \cdot (1 - e^{\log 2 - \frac{1}{2}s})\, ds$$
$$= \int_{2\log 2}^{+\infty} s e^{-s} (1 - 2e^{-\frac{1}{2}s})\, ds$$
$$= e^{-s}(-1-s)\Big|_{2\log 2}^{+\infty} - 2e^{-\frac{3}{2}s}\left(-\frac{4}{9} - \frac{2}{3}s\right)\Big|_{2\log 2}^{+\infty}$$
$$= e^{-\log 4}(1 + 2\log 2) - 2e^{-\frac{3}{2}\cdot 2\log 2}\left(\frac{4}{9} + \frac{2}{3}\cdot 2\log 2\right)$$
$$= \frac{1}{4}(1 + \log 4) - \frac{1}{4}\left(\frac{4}{9} + \frac{4}{3}\log 2\right)$$

$$= \frac{5}{36} + \frac{1}{6}\log 2$$

4. 4.54

It's known that if $X_i \overset{iid}{\sim} U_{[0,1]}$

$$-\log X_i \overset{iid}{\sim} \text{Exponential}(1) = \text{gamma}(1,1)$$

By the property of gamma distribution

$$X = \sum_{i=1}^{n} -\log X_i \sim \text{gamma}(n,1)$$

Therefore $Y = \prod_{i=1}^{n} X_i = \exp\left[\sum_{i=1}^{n} \log X_i\right]$

$$= \exp\left[-\sum_{i=1}^{n} -\log X_i\right]$$

$$= \exp\left[-X\right] \qquad \frac{dy}{dx} = e^{-x}$$

Since $f_X(x) = \frac{1}{\Gamma(n)} x^{n-1} e^{-x}$, we have

$$f_Y(y) = \left|\frac{dx}{dy}\right| f_X(x) = y^{-1} \cdot \frac{1}{\Gamma(n)} (-\log y)^{n-1} \exp[\log y]$$

$$= \frac{1}{\Gamma(n)} (-\log y)^{n-1}, \quad 0 < y < 1$$

5 (a) By the property of multivariate normal distribution,

If $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$, we have

$$CZ \sim N\left(C \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix}, C\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}C^T\right)$$

$$= N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, CC^T\right)$$

Therefore, the density of $X = CZ$ is

$$f_X(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 |CC^T|}} \exp\left[-\frac{1}{2}(x_1, x_2)(CC^T)^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]$$

So.

① we have $g(s) = \frac{1}{\sqrt{(2\pi)^2 |CC^T|}} e^{-s}, \quad s \geq 0.$

obviously, this function satisfies the requirement listed in the definition of elliptical distribution

(due to $\int_0^{+\infty} e^{-s} = 1$)

② And $A = (CC^T)^{-1}$, which is obviously symmetric and invertible. So we can write out its spectrum decomposition as $A = (CC^T)^{-1} = Q \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} Q^T$.

$\forall\ x \in \mathbb{R}^2$, $x^T A x = x^T Q \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} Q^T x =$

$$= (\Lambda^{\frac{1}{2}} Q^T x)^T (\Lambda^{\frac{1}{2}} Q^T x)$$

$$= \| \Lambda^{\frac{1}{2}} Q^T x \|_2 \geq 0$$

and $x^T A x = 0$ iff $\Lambda^{\frac{1}{2}} Q^T x = 0 \Leftrightarrow x = 0$.

This shows A is positive definite.

By ① ②, $X = CZ$ has an elliptical distribution.

And we have $\operatorname{cov}(X) = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} = \Sigma = CC^T$

so $(CC^T)^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix}$

then

$$f_X(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1}{\sigma_1}\right)^2 - 2\rho \frac{x_1}{\sigma_1} \frac{x_2}{\sigma_2} + \left(\frac{x_2}{\sigma_2}\right)^2 \right] \right\}$$

(b) "2f" part.

It's known that. $X_1, X_2 \sim N(0, \sigma^2)$ and $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = C\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$, in which $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ is a bivariate normal distribution. By the property of multivariate normal distribution, since C is non-singular, we have

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = C\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, CC^T\right)$$

We can write out the density of $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$,

$$f_X(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 |CC^T|}} \exp\left[-\frac{1}{2}(x_1, x_2)(CC^T)^{-1}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]$$

Then we have $A = (CC^T)^{-1}$.

It's given that A is a scalar matrix, i.e.

$$A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = (CC^T)^{-1}$$

so

$$f_X(x_1, x_2) = \frac{1}{\sqrt{(2\pi)^2 |A|}} \exp\left[-\frac{1}{2}(x_1, x_2)\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right]$$

$$= \frac{1}{\sqrt{(2\pi)^2 |A|}} \exp\left[-\frac{1}{2}ax_1^2 - \frac{1}{2}ax_2^2\right]$$

$$= \frac{1}{\sqrt{(2\pi)^2 |A|}} \exp\left[-\frac{1}{2}ax_1^2\right] \exp\left[-\frac{1}{2}ax_2^2\right]$$

Note that $f_X(x_1, x_2)$ can be factorized to $g(x_1)$ and $h(x_2)$, therefore, X_1 and X_2 are independent

"Only if" part

First, consider the marginal distribution of X_1, X_2.

We have $f_{X_1}(s) = \int_{-\infty}^{+\infty} g(as^2 + ax_2^2)\,dx_2$ consider $f_{X_2}(s)$

$$f_{X_2}(s) = \int_{-\infty}^{+\infty} g(ax_1^2 + as^2)\,dx_1$$

$$= \int_{-\infty}^{+\infty} g(as^2 + ax_2^2)\,dx_2 \quad \left[\begin{array}{c}\text{just substitute} \\ x_1 \text{ to } x_2\end{array}\right]$$

$$= f_{X_1}(s)$$

Therefore, X_1 and X_2 must have the same form of marginal distribution. Denote it by f and assume it to be continuous.

So, if X_1 and X_2 are independent, we have

$$g(ax_1^2 + ax_2^2) = f(x_1) f(x_2) \quad \forall x_1, x_2 \in \mathbb{R}$$

Note that we must have $f(0) > 0$. since if $f(0) = 0$, then $\forall x_1$,

$$g(ax_1^2 + a \cdot 0) = g(ax_1^2) = f(x_1) f(0) = 0.$$

note that $a > 0$, let $t = ax_1^2$, then $\forall t \geq 0$, we have $g(t) = 0$. However $f_{X_1,X_2}(x_1, x_2) = g(X^T A X)$ is a PDF. We can't let $g(t) \equiv 0 \ \forall t \geq 0$

By $f(0)>0$, we have $g(0)=f(0)\cdot f(0)>0$

Define the continuous function $g_0(u^2)=g(au^2)/g(0)$

i.e. $g_0(\tilde{u})=g(a\tilde{u})/g(0)$, $\tilde{u}\geq 0$

Then $$g_0(\tilde{u}+\tilde{v})=\frac{g(a\tilde{u}+a\tilde{v})}{g(0)}$$

$$=\frac{g(au^2+av^2)}{g(0)} \quad (\text{let } \tilde{u}=u^2,\ \tilde{v}=v^2)$$

$$=\frac{f(u)f(v)}{f(0)f(0)}=\frac{f(u)f(0)\cdot f(v)f(0)}{f(0)f(0)\cdot f(0)f(0)}$$

$$=\frac{g(au^2)}{g(0)}\cdot\frac{g(av^2)}{g(0)}=\frac{g(a\tilde{u})}{g(0)}\cdot\frac{g(a\tilde{v})}{g(0)}$$

$$=g_0(\tilde{u})g_0(\tilde{v}),\quad \forall\ \hat{u},\tilde{v}\geq 0$$

Moreover, we have $\int_0^{+\infty} g(x)\,dx<\infty$, then

$$\int_0^{+\infty} g_0(x)\,dx=\int_0^{+\infty} g(ax)/g(0)\,dx$$

$$=\frac{1}{g(0)}\int_0^{+\infty}\frac{1}{a}\cdot g(t)\,dt<\infty$$

For the equation $g_0(s+t)=g_0(s)g_0(t)$, $\forall s,t\geq 0$

$\forall$ integer $m,n\geq 1$, and all positive real number t.

we have $g_0(mt)=g_0(n\cdot\frac{m}{n}t)=[g_0(\frac{m}{n}t)]^n$

$$g_0(mt)=g_0(t+\cdots+t)=[g_0(t)]^m$$

thus $g_0(\frac{m}{n}t) = [g_0(t)]^{\frac{m}{n}}$

Note that $\frac{m}{n}$ is a positive rational number

It's known that for any real number, there exists some rational sequence that converges to it.

So, we can construct m/n as such sequence

i.e. we have $\lim_{n\to\infty} x_n = x \in \mathbb{R}^+$, $g_0(x_n t) = [g_0(t)]^{x_n}$

in which x_n is a positive rational number

and $x > 0$. Since g_0 is continuous,

we take the limit $n\to\infty$ on both sides of the equation above,

$$\lim_{n\to\infty} g_0(x_n t) = g_0(\lim_{n\to\infty} x_n t) = g_0(xt)$$
$$= \lim_{n\to\infty} [g_0(t)]^{x_n} = [g_0(t)]^{x}$$

thus $\forall x>0, t>0$, $g_0(xt) = [g_0(t)]^x$

Take $t=1$ and let $c = -\ln g_0(1)$

$$g_0(x) = [g_0(1)]^x = e^{\ln g_0(1)\cdot x} = e^{-cx}, \quad x>0$$

Note that to make $\int_0^{+\infty} g_0(x)\,dx = \int_0^{+\infty} e^{-cx}\,dx < \infty$

we must have $c > 0$

For $x=0$, we have $g_0(x) = g(0)/g(0) = 1$, which is consistent with $g_0(x) = e^{-cx}$.

So we have $g_0(x) = e^{-cx}$, $x \geq 0$

Then $g(au^2) = f(u) f(0)$

$$= g_0(u^2) \cdot g(0) \cdot f(0) \quad \forall u$$

$$\Rightarrow f(u) = g(0) \cdot \exp(-cu^2), \quad u \in \mathbb{R}$$

Since $f(u)$ is density function,

we have $\int_{-\infty}^{+\infty} f(u)\, du = \int_{-\infty}^{+\infty} g(0) e^{-cu^2} du$

$$= g(0) \cdot \sqrt{\frac{\pi}{c}} = 1.$$

Let $c = \frac{1}{2\sigma^2}$, we have

$$f(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{u^2}{2\sigma^2}\right] \sim N(0, \sigma^2)$$

So X_1, X_2 have the both $N(0, \sigma^2)$

6. (a) The joint distribution of $(X_1 \cdots X_d)$ is

$$f_{\vec{X}}(x_1 \cdots x_d) = \prod_{j=1}^{d} \frac{\lambda^{\alpha_j}}{\Gamma(\alpha_j)} x_j^{\alpha_j - 1} e^{-\lambda x_j}$$

Let $\vec{T} = (W_1 \cdots W_{d-1}, Y)$, then we have

$$f_{\vec{T}}(w_1 \cdots w_{d-1}, y) = \left|\frac{\partial \vec{x}}{\partial \vec{t}}\right| f_{\vec{X}}(x_1 \cdots x_d)$$

$$\left|\frac{\partial \vec{x}}{\partial \vec{t}}\right| = \begin{vmatrix} \frac{\partial x_1}{\partial w_1} & \frac{\partial x_1}{\partial w_2} & \cdots & \frac{\partial x_1}{\partial w_{d-1}} & \frac{\partial x_1}{\partial y} \\ \vdots & & & & \\ \frac{\partial x_n}{\partial w_1} & \frac{\partial x_n}{\partial w_2} & \cdots & \frac{\partial x_n}{\partial w_{d-1}} & \frac{\partial x_n}{\partial y} \end{vmatrix} = \begin{vmatrix} y & 0 & \cdots & 0 & w_1 \\ 0 & y & \cdots & 0 & w_2 \\ \vdots & & & \vdots & \vdots \\ & & & y & w_{d-1} \\ -y & -y & \cdots & -y & 1-w_1-\cdots-w_{d-1} \end{vmatrix}$$

$$\left(\begin{array}{l}\text{property of}\\ \text{determinant}\end{array}\right) = \begin{vmatrix} y & 0 & \cdots & 0 & w_1 \\ 0 & y & \cdots & 0 & w_2 \\ \vdots & & & \vdots & \vdots \\ & & & y & w_{d-1} \\ 0 & \cdots & \cdots & 0 & 1 \end{vmatrix} \overset{\text{(Laplace expension)}}{=} y^{d-1}$$

Thus. $f_{\vec{T}}(w_1 \cdots w_{d-1}, y) = y^{d-1} \prod_{j=1}^{d} \frac{\lambda^{\alpha_j}}{\Gamma(\alpha_j)} x_j^{\alpha_j - 1} e^{-\lambda x_j}$

$$= y^{d-1}\left[\prod_{j=1}^{d-1} \frac{\lambda^{\alpha_j}}{\Gamma(\alpha_j)} (w_j y)^{\alpha_j-1} e^{-\lambda w_j y}\right] \frac{\lambda^{\alpha_d}}{\Gamma(\alpha_d)}\left[y\left(1-\sum_{i=1}^{d-1} w_i\right)\right]^{\alpha_d-1} e^{-\lambda y\left(1-\sum_{i=1}^{d-1} w_i\right)}$$

$$= \left(\frac{\lambda^{\sum_{i=1}^{d}\alpha_j}}{\prod_{i=1}^{d}\Gamma(\alpha_j)}\left[\prod_{i=1}^{d-1} w_j^{\alpha_j-1}\right](1-w_1-\cdots-w_{d-1})^{\alpha_d-1}\right)\left(\prod_{j=1}^{d} y^{\alpha_j-1}\right) y^{d-1} e^{-\lambda y}$$

$$= \frac{\lambda^{\alpha_1+\cdots+\alpha_d}}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_d)} w_1^{\alpha_1-1} \cdots w_{d-1}^{\alpha_{d-1}-1} (1-w_1-\cdots-w_{d-1})^{\alpha_d-1} \cdot$$

$$y^{\alpha_1+\cdots+\alpha_d-1} e^{-\lambda y}, \quad \begin{array}{l} 0<w_i<1 \\ y>0 \end{array}$$

(b) To get marginal distribution of $(w_1 \cdots w_{d-1})$, we integrate over y.

$$f(w_1 \cdots w_{d-1}) = \int_0^{+\infty} f(w_1 \cdots w_{d-1}, y)\, dy$$

$$= \frac{1}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_d)} w_1^{\alpha_1 - 1} \cdots w_{d-1}^{\alpha_{d-1}-1} (1 - w_1 - \cdots w_{d-1})^{\alpha_d - 1} \cdot$$

$$\int_0^{+\infty} (\lambda y)^{\alpha_1 + \cdots + \alpha_d - 1} \lambda e^{-\lambda y} \, dy$$

$$= \frac{1}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_d)} w_1^{\alpha_1 - 1} \cdots w_{d-1}^{\alpha_{d-1}-1} (1 - w_1 - \cdots w_{d-1})^{\alpha_d - 1} \cdot$$

$$\int_0^{+\infty} t^{\alpha_1 + \cdots + \alpha_d - 1} e^{-t} \, dt$$

$$= \frac{\Gamma(\alpha_1 + \cdots + \alpha_d)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_d)} w_1^{\alpha_1 - 1} \cdots w_{d-1}^{\alpha_{d-1}-1} (1 - w_1 - \cdots - w_{d-1})^{\alpha_d - 1}$$

which is a Dirichlet distribution of parameter $(\alpha_1, \alpha_2 \cdots \alpha_d)$

To get the marginal distribution of Y, we integrate over $w_1 \cdots w_{d-1}$. Note that we can factorize $f(w_1 \cdots w_{d-1}, y)$ as

$$f(w_1 \cdots w_{d-1}, y) = f(w_1 \cdots w_{d-1}) \cdot \frac{\lambda^{\alpha_1 + \cdots + \alpha_d}}{\Gamma(\alpha_1 + \cdots + \alpha_d)} y^{\alpha_1 + \cdots \alpha_d - 1} e^{-\lambda y}$$

by the property of $f(w_1 \cdots w_{d-1})$, we have

$$f(y) = \frac{\lambda^{\alpha_1 + \cdots + \alpha_d}}{\Gamma(\alpha_1 + \cdots + \alpha_d)} y^{\alpha_1 + \cdots \alpha_d - 1} e^{-\lambda y}$$, which is gamma $(\alpha_1 + \cdots + \alpha_d, \lambda)$

From above, we have

$$f_{W_1 \cdots W_{d-1}, Y}(w_1 \cdots w_{d-1}, y) = f_{W_1 \cdots W_{d-1}}(w_1 \cdots w_{d-1}) f_Y(y)$$

Therefore $(W_1 \cdots W_{d-1}) \perp\!\!\!\perp Y$

(c) By the definition of N_i, we have

$$N_0 = 0, N_1 = n_1, \quad N_2 = n_1 + n_2, \quad N_3 = n_1 + n_2 + n_3$$

$$\cdots \quad N_k = n_1 + n_2 + \cdots + n_k = d$$

and $U_1 = \sum_{j=1}^{N_1} W_j = W_1 + \cdots + W_{n_1}$

$$U_2 = \sum_{j=n_1+1}^{n_1+n_2} W_j = W_{n_1+1} + W_{n_1+2} \cdots W_{n_1+n_2}$$

$$U_3 = \sum_{j=n_1+n_2+1}^{n_1+n_2+n_3} W_j = W_{n_1+n_2+1} + W_{n_1+n_2+2} + \cdots + W_{n_1+n_2+n_3}$$

$$\vdots$$

$$U_k = W_{n_1+\cdots+n_{k-1}+1} \cdots + W_d$$

By this, $U_1 \cdots U_k$ are just dividing $(W_1 \cdots W_d)$ to k groups, according to their subscripts, such that the j-th group contains n_j's W_i, whose subscripts are from $n_1 + \cdots + n_{j-1} + 1$ to $n_1 + \cdots + n_{j-1} + n_j$.

And denote the sum of j-th group by U_j

By the property of Dirichlet distribution:

if $X = (X_1 \cdots X_k) \sim Dir(\alpha_1 \cdots \alpha_k)$, then

$$\tilde{X} = (X_1, \cdots, X_i + X_j, \cdots X_k) \sim Dir(\alpha_1 \cdots \alpha_i + \alpha_j, \cdots \alpha_k)$$

We have

$$(U_1, U_2 \cdots U_k) \sim Dir\left(\sum_{j=1}^{N_1} \alpha_j, \sum_{j=N_1+1}^{N_2} \alpha_j \cdots \sum_{j=N_{k-1}+1}^{N_k}\right)$$

Let $\beta_i = \sum_{j=N_{i-1}+1}^{N_i} \alpha_j$, $i = 1 \cdots k$

The joint density is

$$f_{U_1 \cdots U_{k-1}}(u_1 \cdots u_{k-1}) = \frac{\Gamma(\beta_1 + \cdots + \beta_k)}{\Gamma(\beta_1) \cdots \Gamma(\beta_k)} u_1^{\beta_1 - 1} \cdots u_{k-1}^{\beta_{k-1} - 1} (1 - u_1 - \cdots - u_{k-1})^{\beta_k - 1}$$