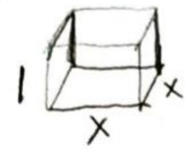
Data Science HW3

余奇安 r07921010

Problem1

1.



$V = X^2$

X is standard Gaussian distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$F(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt \quad \text{for } \mu=0, \sigma=1$$

$$F(x;0,1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$$

$$= \frac{1}{2}\left(1+\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right), \quad \left(\operatorname{erf}(x) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} e^{-t^2} dt\right)$$

$$\frac{d\operatorname{erf}(x)}{dx} = \frac{1}{\sqrt{\pi}}\left(e^{-x^2} - (-1)e^{-x^2}\right)$$

$$= \frac{2}{\sqrt{\pi}} e^{-x^2}$$

$$F_V(v) = P(V<v)$$

$$= P(X^2 < v)$$

$$= P(-\sqrt{v} \leq X \leq \sqrt{v}) =$$

$$= \frac{1}{2}\left(1+\operatorname{erf}\left(\sqrt{\tfrac{v}{2}}\right)\right) - \frac{1}{2}\left(1+\operatorname{erf}\left(-\sqrt{\tfrac{v}{2}}\right)\right)$$

$$= \frac{1}{2}\operatorname{erf}\left(\sqrt{\tfrac{v}{2}}\right) - \frac{1}{2}\operatorname{erf}\left(-\sqrt{\tfrac{v}{2}}\right) = \operatorname{erf}\left(\sqrt{\tfrac{v}{2}}\right) \quad u = \sqrt{\tfrac{v}{2}} = \left(\tfrac{v}{2}\right)^{\frac{1}{2}}$$

1)

$$\text{PDF} = \frac{dF_V(v)}{dv} = \frac{2}{\sqrt{\pi}} e^{-\frac{v}{2}} \cdot \frac{1}{2}\left(\frac{v}{2}\right)^{-\frac{1}{2}} \cdot \frac{1}{2}$$

$$= \frac{1}{2\sqrt{\pi}} e^{-\frac{v}{2}} \cdot \left(\frac{v}{2}\right)^{-\frac{1}{2}} = \frac{1}{\sqrt{2}\sqrt{\pi}} e^{-\frac{v}{2}} v^{-\frac{1}{2}}$$

2) Gamma Distribution PDF $GAM(\theta,\alpha)$

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\theta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\theta}} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

Why dependent? box constrain
100% correlation

$$\alpha - 1 = -\frac{1}{2} \Rightarrow \underline{\alpha = \frac{1}{2}}$$

$$\underline{\theta = 2}$$

Problem2

(1)

X, Y pdf

2. $f(x) = \begin{cases} e^{-x} & \text{for } x > 0 \\ 0 & \text{elsewise} \end{cases}$ $\quad Z = \max(X, Y)$ $\quad W = \min(X, Y)$

CDF $F_X(x) = F_Y(y)$

$F_X(x) = \int_0^x e^{-u}\, du = -e^{-u}\Big|_0^x = -e^{-x} - (-1) = 1 - e^{-x}$

(1)

$F_Z(z) = P(Z \le z) = P(\max(X, Y) \le z)$

$= P(X \le z)\, P(Y \le z) = F_X(z)\, F_Y(z)$

$f_Z(z) = \frac{dF_Z(z)}{dz} = f_X(z) F_Y(z) + F_X(z) f_Y(z)$

$= e^{-z} \cdot (1 - e^{-z}) + (1 - e^{-z}) e^{-z} = 2e^{-z}(1 - e^{-z})$

(2)

(2) $F_W(w) = P(W \le w) = P(\min(X, Y) \le w)$

$= 1 - P(\min(X, Y) > w)$

$f_W(w) = \frac{dF_W(w)}{dw} = \frac{-d\, P(\min(X, Y) > w)}{dw} = \frac{-d\big(P(X > w)\, P(Y > w)\big)}{dw} = \frac{-d\big((1 - P(X \le w))(1 - P(Y \le w))\big)}{dw}$

$= \frac{-d\big((1 - F_X(w))(1 - F_Y(w))\big)}{dw} = \frac{-d\big(1 - F_X(w) - F_Y(w) + F_X(w) F_Y(w)\big)}{dw}$

$= +f_X(w) + f_Y(w) - f_X(w) F_Y(w) - F_X(w) f_Y(w)$

$= e^{-w} + e^{-w} - e^{-w}(1 - e^{-w}) - (1 - e^{-w}) e^{-w}$

$= 2e^{-w} - 2e^{-w}(1 - e^{-w}) = 2e^{-w}(1 - 1 + e^{-w})$

$= 2e^{-w}(e^{-w})$

$= 2e^{-2w}$

Problem3

```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import cv2
import math as mt

# read image in grayscale
img = cv2.imread('A.jpg', 0)
img2 = cv2.imread('B.jpg',0)
print('img shape',img.shape)
print('img2 shape',img2.shape)
# obtain svd
U, S, V = np.linalg.svd(img2)

# inspect shapes of the matrices
print(U.shape, S.shape, V.shape)

size = 1936 # 638
nSingular = [size, mt.floor(size*0.8), mt.floor(size*0.5), mt.floor(size*0.2), mt.floor(size*0.1), mt.floor(size*0.05)]

plt.figure(figsize = (32, 16))
for i in range(6):
    low_rank = U[:, :nSingular[i]] @ np.diag(S[:nSingular[i]]) @ V[:nSingular[i], :]
    if(i  == 0):
        plt.subplot(2, 3, i+1), plt.imshow(low_rank, cmap = 'gray'), plt.axis('off'), plt.title("Original Image with n_Singu
    else:
        plt.subplot(2, 3, i+1), plt.imshow(low_rank, cmap = 'gray'), plt.axis('off'), plt.title("Image with n_Singulars =" +
```