Jilin Zheng // U49258796

EK381 Homework #7

Problem 7.1

$X = U+V$, $Y = U-2V$, U and V are independent, standard Gaussians (Gaussian(0,1))

(a) Yes, X and Y are linear functions of independent, standard Gaussian random variables, so they are jointly Gaussian

(b) $\mathbb{E}[X] = \mathbb{E}[U+V] = \mathbb{E}[U] + \mathbb{E}[V] = \emptyset$

$\mathbb{E}[Y] = \mathbb{E}[U-2V] = \emptyset - 2(\emptyset) = \emptyset$

(c) $Cov[X,Y] = Cov[U+V, U-2V] = 1 \cdot Var[X] + (-2)Var[Y] + (-2+1)Cov[U,V]$

$= 1 + (-2) + (-1)(\emptyset) = -1$

(d) $Var[X] = Var[U+V] = Var[U] + Var[V] + 2Cov[U,V] = 1+1+2(\emptyset) = 2$

$Var[Y] = Var[U-2V] = Var[U] + (-2)^2 Var[V] + 2(1)(-2)Cov[U,V] = 1+4 = 5$

$\rho_{X,Y} = \frac{Cov[X,Y]}{\sqrt{Var[X]Var[Y]}} = \frac{-1}{\sqrt{2 \cdot 5}} = \frac{-1}{\sqrt{10}}$

(e) $\mathbb{E}[X|Y=y] = \mu_X + \frac{Cov[X,Y]}{Var[Y]}(y - \mu_Y) = \emptyset + \frac{-1}{5}(y - \emptyset) = \frac{-y}{5}$

(f) $Var[X|Y] = (1 - \rho_{X,Y}^2)\sigma_X^2 = \left(1 - \left(\frac{-1}{\sqrt{10}}\right)^2\right)2 = \left(1 - \left(\frac{1}{10}\right)\right)2 = \frac{9}{10} = \frac{9}{5}$

Problem 7.2

(a) $\underline{\mu_Y} = \begin{bmatrix} \mathbb{E}[Y_1] \\ \mathbb{E}[Y_2] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[2X_1+1] \\ \mathbb{E}[X_1 - X_2] \end{bmatrix} = \begin{bmatrix} 2(\emptyset)+1 \\ \emptyset - \emptyset \end{bmatrix} = \begin{bmatrix} 1 \\ \emptyset \end{bmatrix}$

(b) $\underline{\Sigma_Y} = \begin{bmatrix} Cov[Y_1, Y_1] & Cov[Y_1, Y_2] \\ Cov[Y_2, Y_1] & Cov[Y_2, Y_2] \end{bmatrix} = \begin{bmatrix} 4 & 4/3 \\ 4/3 & 4/3 \end{bmatrix}$

$Var[Y_1] = Var[2X_1+1] = 4Var[X_1] = 4$

$Var[Y_2] = Var[X_1 - X_2] = Var[X_1] + Var[X_2] + 2(1)(-1)Cov[X_1, X_2] = 1 + 1 + (-2)(1/3) = 2 - 2/3 = 6/3 - 2/3 = 4/3$

$Cov[Y_1, Y_2] = Cov[2X_1+1, X_1 - X_2] = 2Var[X_1] + \emptyset + (2(-1)+\emptyset)Cov[X_1, X_2] = 2 + (-2)\frac{1}{3} = 6/3 + (-2/3) = 4/3$

$= Cov[Y_2, Y_1]$

(c) Y_1 and Y_2 are linear functions of jointly Gaussian X_1 and X_2, so Y_1 and Y_2 are also jointly Gaussian

$\rho_{Y_1, Y_2} = \frac{Cov[Y_1, Y_2]}{\sqrt{Var[Y_1]Var[Y_2]}} = \frac{4/3}{\sqrt{4 \cdot 4/3}} = \frac{4/3}{\sqrt{16/3}} = \frac{4}{3} \cdot \frac{\sqrt{3}}{4} = \frac{\sqrt{3}}{3}$

Since Y_1 and Y_2 are NOT uncorrelated ($\rho_{Y_1, Y_2} \neq \emptyset$), they are NOT independent

(d) $\mathbb{E}[Y_1 \mid Y_2 = y_2] = \mu_{Y_1} + \frac{Cov[Y_1, Y_2]}{Var[Y_2]}(y_2 - \mu_{Y_2})$

$= 1 + \frac{4/3}{4/3}(y_2 - \emptyset)$

$= 1 + y_2$

Problem 7.3

(a) $\mathbb{E}[\underline{X}] = \underline{\mu}_X = \begin{bmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_n] \end{bmatrix} = \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix}$

$$\Sigma_{\underline{X}} = \begin{bmatrix} Cov[X_1,X_1] & \cdots & Cov[X_1,X_n] \\ \vdots & & \vdots \\ Cov[X_n,X_1] & \cdots & Cov[X_n,X_n] \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2 \end{bmatrix}$$

(b) $Y = \frac{1}{n}\left(\sum_{i=1}^{n} X_i\right) = A\underline{X} = \begin{bmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$

(c) $\underset{1\times1}{\mathbb{E}[Y]} = \mathbb{E}[A\underline{X}] = \underset{1\times n}{A}\,\underset{n\times1}{\mathbb{E}[\underline{X}]} = \begin{bmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix} \begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix} = \sum_{i=1}^{n} \frac{\mu}{n} = \frac{\mu}{n}(n) = \mu$

(d) $$\Sigma_Y = A\Sigma_{\underline{X}}A^T = \underset{1\times n}{\begin{bmatrix} \frac{1}{n} & \cdots & \frac{1}{n} \end{bmatrix}} \underset{n\times n}{\begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sigma^2 \end{bmatrix}} \underset{n\times1}{\begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix}}$$

$$= \underset{1\times n}{\begin{bmatrix} \frac{\sigma^2}{n} & \cdots & \frac{\sigma^2}{n} \end{bmatrix}} \underset{n\times1}{\begin{bmatrix} 1/n \\ \vdots \\ 1/n \end{bmatrix}} = \frac{\sigma^2}{n}\left(\frac{1}{n}\right)n = \underset{1\times1}{\frac{\sigma^2}{n}}$$

$Var[Y] = \frac{\sigma^2}{n}$

(e) $$\mathbb{P}[|Y-\mathbb{E}[Y]| > \delta] = 1-\mathbb{P}[|Y-\mathbb{E}[Y]| < \delta]$$
$$= 1-\mathbb{P}[-\delta < Y-\mathbb{E}[Y] < \delta]$$
$$= 1-\mathbb{P}[-\delta+\mu < Y < \delta+\mu] \qquad \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$$
$$= 1-\left[\Phi\left(\frac{\delta+\mu-\mu}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{-\delta+\mu-\mu}{\sigma/\sqrt{n}}\right)\right]$$
$$= 1-\left[\Phi\left(\frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(\frac{-\delta\sqrt{n}}{\sigma}\right)\right]$$
$$= Q\left(\frac{\delta\sqrt{n}}{\sigma}\right) + \Phi\left(\frac{-\delta\sqrt{n}}{\sigma}\right) = Q\left(\frac{\delta}{\frac{\sigma}{\sqrt{n}}}\right) + \Phi\left(\frac{-\delta}{\frac{\sigma}{\sqrt{n}}}\right) = 2Q\left(\frac{\delta}{\frac{\sigma}{\sqrt{n}}}\right)$$

(f) $\delta = \frac{1}{10}$

$\frac{1}{1000} = 2Q\left(\frac{\frac{1}{10}}{\frac{\sigma}{\sqrt{n}}}\right) = 2Q\left(\frac{\sqrt{n}}{10\sigma}\right)$

$Q\left(\frac{\sqrt{n}}{10\sigma}\right) = \frac{1}{2000}$

$\frac{\sqrt{n}}{10\sigma} = 3.29$

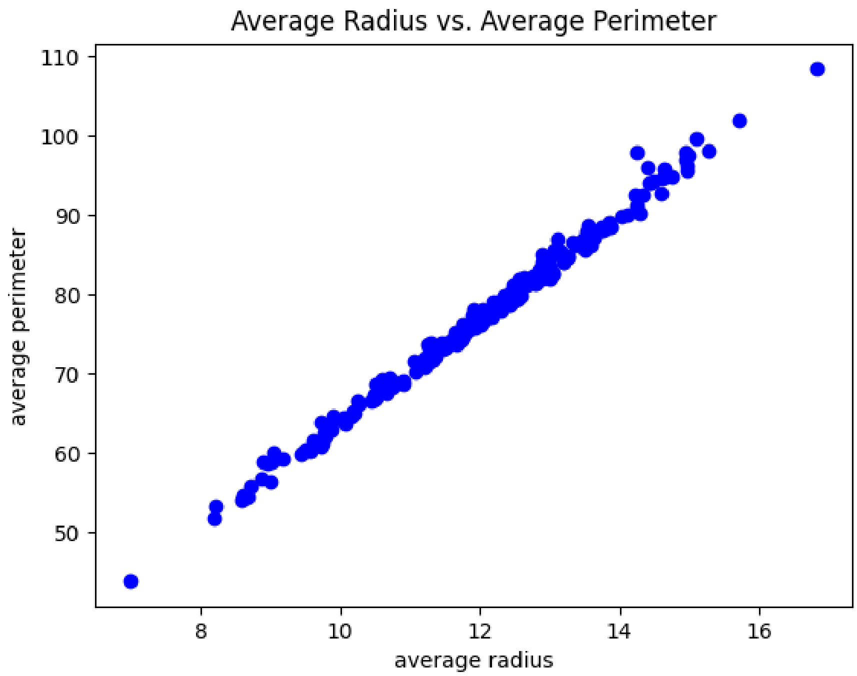
$\sqrt{n} = 32.9\sigma$

$n = 1082.41\sigma^2$

Problem 7.4

Part a.

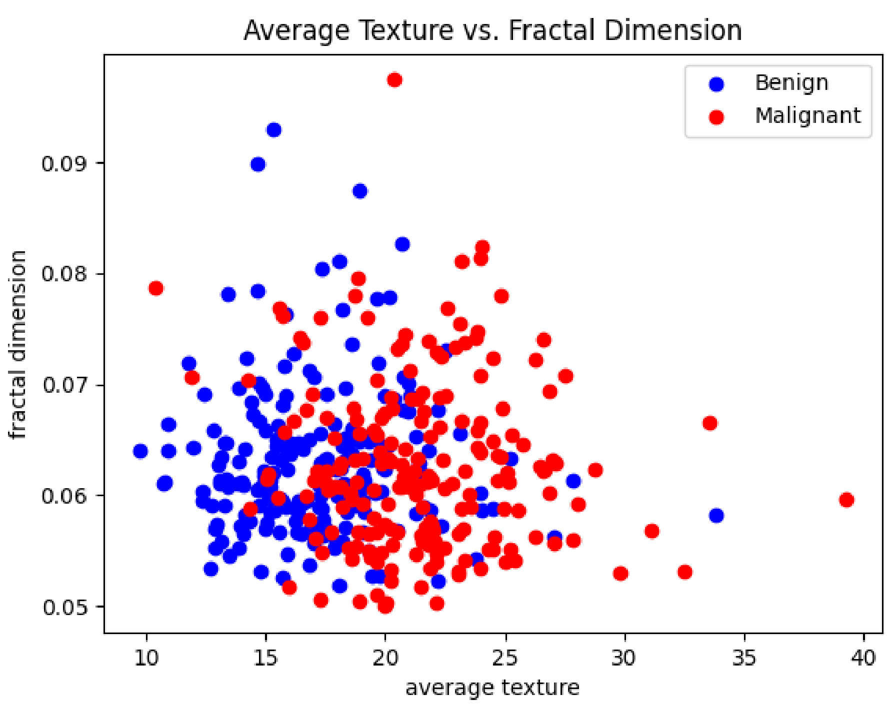
Indices of the columns:0 (average radius) and 2 (average perimeter)



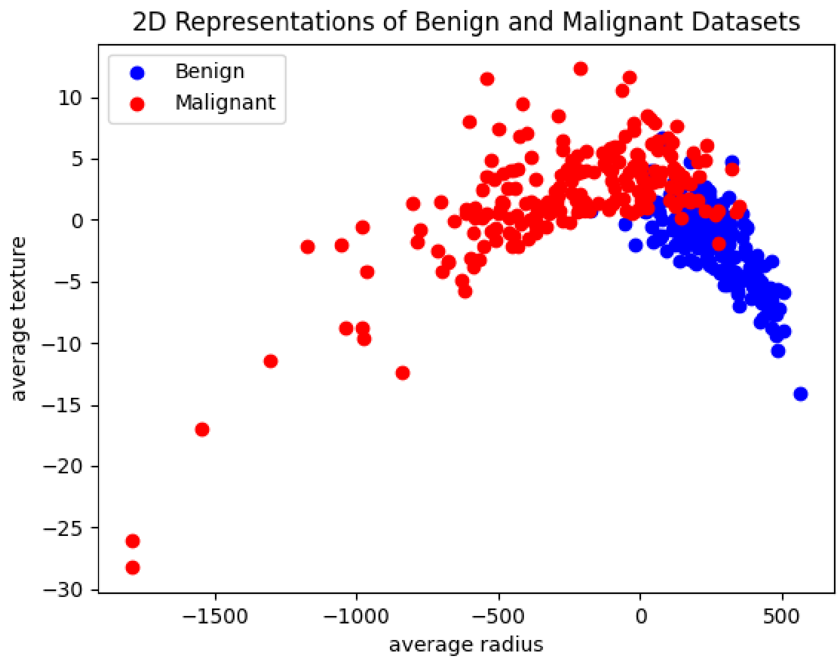
It would make sense that the two columns have a linear relationship because perimeter is a linear function of the radius, i.e. perimeter = 2 * pi * radius.

Part b.

Indices of the columns: 1 (average texture) and 9 (fractal dimension)

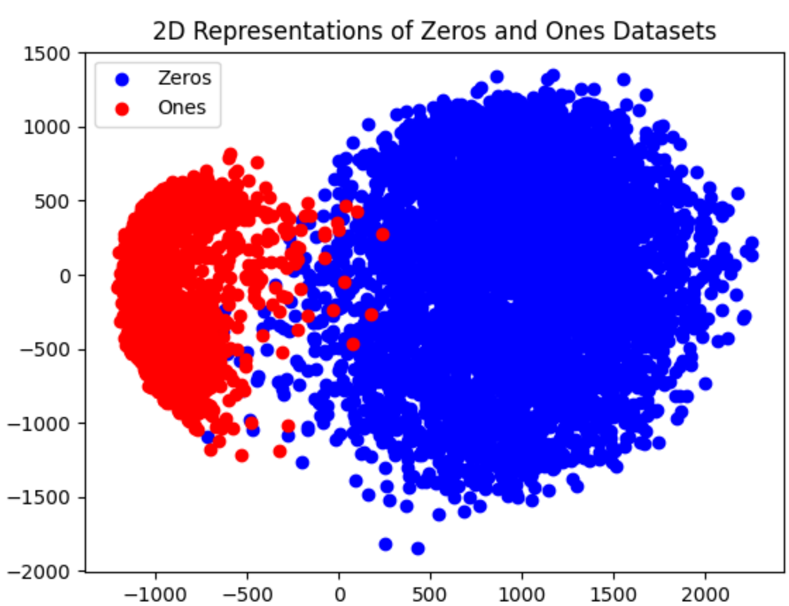


Part c.



If we did not have color labeling, it would be easier to classify the benign and malignant points in this plot rather than the plot from part b because the points have distinct ‘sides’ of the plot, i.e. the malignant dominates the left side of the plot while the benign is more towards the right side of the plot. In the plot from part b, both datasets occupy the same region.

Part d.



This dataset looks easier to classify than the previous one, as the points on the plot are even more distinctly separated: the zeros largely occupy the right side, with x-values greater than zero, and the ones occupy the left side, x-values less than zero.