Math 583 Exam 1, Fall 2012 Name___________

e 1) Consider the data set listed below.

0 8 10 12 $\underbrace{13 \quad 13}_{MED(n)=13}$ 14 18 240 460

Find the sample median absolute deviation MAD(n).

$-13, -5, -3, -1, 0, 0, 1, 5, 227, 447$

$0, 0, 1, 1, 3, 5, 5, 13, 227, 447$ $|Y - MED(n)|$

$$MAD(n) = \frac{3+5}{2} = \boxed{4}$$

7

2) Suppose $\boldsymbol{A}$ is a symmetric positive definite matrix with eigenvalue eigenvector pair $(\lambda, \boldsymbol{e})$. Then $\boldsymbol{Ae} = \lambda \boldsymbol{e}$ so $\boldsymbol{A}^2\boldsymbol{e} = \boldsymbol{AAe} = \boldsymbol{A}\lambda\boldsymbol{e}$. Find an eigenvalue eigenvector pair for $\boldsymbol{A}^2$.

$$A^2 \underline{e} = \lambda A e = \lambda \lambda \underline{e} = \lambda^2 \underline{e}$$

So $\boxed{(\lambda^2, \underline{e})}$.

7

3) Suppose $\boldsymbol{x}_1, ..., \boldsymbol{x}_n$ are iid $p \times 1$ random vectors where

$$\boldsymbol{x}_i \sim (1-\gamma) N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}) + \gamma N_p(\boldsymbol{\mu}, c\boldsymbol{\Sigma})$$

yes see back of P. 1

with $0 < \gamma < 1$ and $c > 0$. Then $E(\boldsymbol{x}_i) = \boldsymbol{\mu}$ and $\text{Cov}(\boldsymbol{x}_i) = [1 + \gamma(c-1)]\boldsymbol{\Sigma}$. Find the limiting distribution of $\sqrt{n}(\overline{\boldsymbol{x}} - \boldsymbol{c})$ for appropriate vector $\boldsymbol{c}$.

$$\sqrt{n}(\underline{\bar{x}} - \underline{\mu}) \xrightarrow{D} N_p(\underline{0}, [1 + \gamma(c-1)]\Sigma)$$

by MCLT

$\overline{7}$

4) The table $\boldsymbol{W}$ shown below represents 4 measurements on 5 people. Find a) the sample mean $\overline{\boldsymbol{x}}$ and b) the coordinatewise median MED($\boldsymbol{W}$).

	age	breadth	cephalic	circum
	19.00	149.0	78.8	547
	0.06	88.5	77.6	324
	0.21	99.0	76.4	369
	0.42	95.0	73.6	368
	0.50	104.5	78.8	385
Σ	20.19	536	385.2	1993

$$\begin{matrix} .06 & .21 & \begin{pmatrix} .42 \\ 99 \\ 77.6 \\ 369 \end{pmatrix} & .5 & 19 \\ 88.5 & 95 & & 104.5 & 149 \\ 73.6 & 76.4 & & 78.8 & 78.8 \\ 324 & 368 & & 385 & 547 \end{matrix}$$

$\underbrace{\qquad}_{\text{MED}(W)}$

$$\underline{\bar{x}} = \frac{1}{5}\begin{pmatrix} 20.19 \\ 536 \\ 385.2 \\ 1993 \end{pmatrix} = \begin{pmatrix} 4.038 \\ 107.2 \\ 77.04 \\ 398.6 \end{pmatrix}$$

$\overline{10}$

5) Let $\boldsymbol{X}$ be an $n \times p$ constant matrix and let $\boldsymbol{\beta}$ be a $p \times 1$ constant vector. Suppose $\boldsymbol{Y} \sim N_n(\boldsymbol{X\beta}, \sigma^2 \boldsymbol{I})$. Find the distribution of $\boldsymbol{HY}$ if $\boldsymbol{H}^T = \boldsymbol{H} = \boldsymbol{H}^2$ is an $n \times n$ matrix and if $\boldsymbol{HX} = \boldsymbol{X}$. Simplify.

$$HX\beta = X\beta, \quad H\sigma^2 I H^T = \sigma^2 HH^T = \sigma^2 H$$

$$\boxed{N_n(X\beta, \sigma^2 H)}$$

7

6) Recall that if $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \Sigma)$, then the conditional distribution of $\boldsymbol{X}_1$ given that $\boldsymbol{X}_2 = \boldsymbol{x}_2$ is multivariate normal with mean $\boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\boldsymbol{x}_2 - \boldsymbol{\mu}_2)$ and covariance matrix $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Let Y and X follow a bivariate normal distribution

$$\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_2\left(\begin{pmatrix} 134 \\ 96 \end{pmatrix}, \begin{pmatrix} 24.5 & 1.1 \\ 1.1 & 23.0 \end{pmatrix} \right).$$

a) Find $E(Y|X)$. $134 + 1.1 \frac{1}{23}(x - 96)$

$$= 134 - \frac{1.1(96)}{23} + \frac{1.1}{23} x = \boxed{129.4087 + 0.04783x}$$

b) Find $\text{Var}(Y|X)$.

$$24.5 - 1.1 \frac{1}{23} 1.1 = 24.5 - \frac{(1.1)^2}{23} = \boxed{24.4474}$$

12

e 7) Suppose that

Sec

M580

E 2016

$$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} \sim N_4\left(\begin{pmatrix} 1 \\ 7 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 3 & 1 \\ 1 & 0 & 1 & 5 \end{pmatrix} \right).$$

a) Find the distribution of $(X_1, X_4)^T$.

$$N_2\left[\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 5 \end{pmatrix}\right]$$

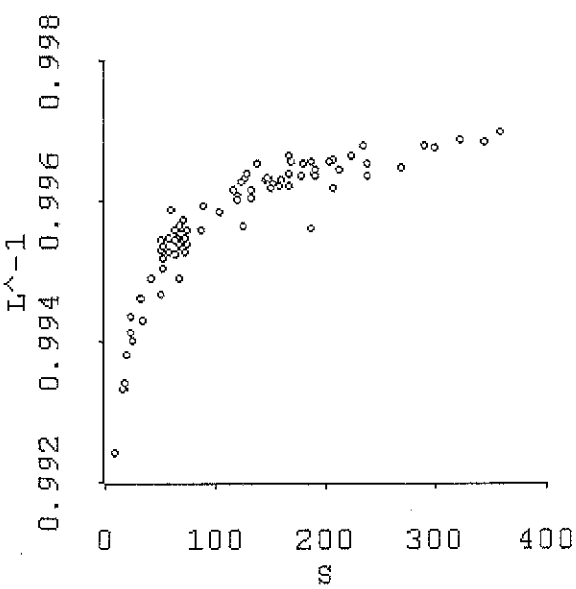
b) Which pairs of random variables X_i and X_j are independent?

$X_1 \perp\!\!\!\perp X_2$, $X_3 \perp\!\!\!\perp X_2$, $X_4 \perp\!\!\!\perp X_2$

c) Find the correlation $\rho(X_1, X_4)$. $\dfrac{\sigma_{14}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{44}}} = \dfrac{1}{\sqrt{4}\sqrt{5}} = \dfrac{1}{\sqrt{20}} = \boxed{0.2236}$

or .224

15

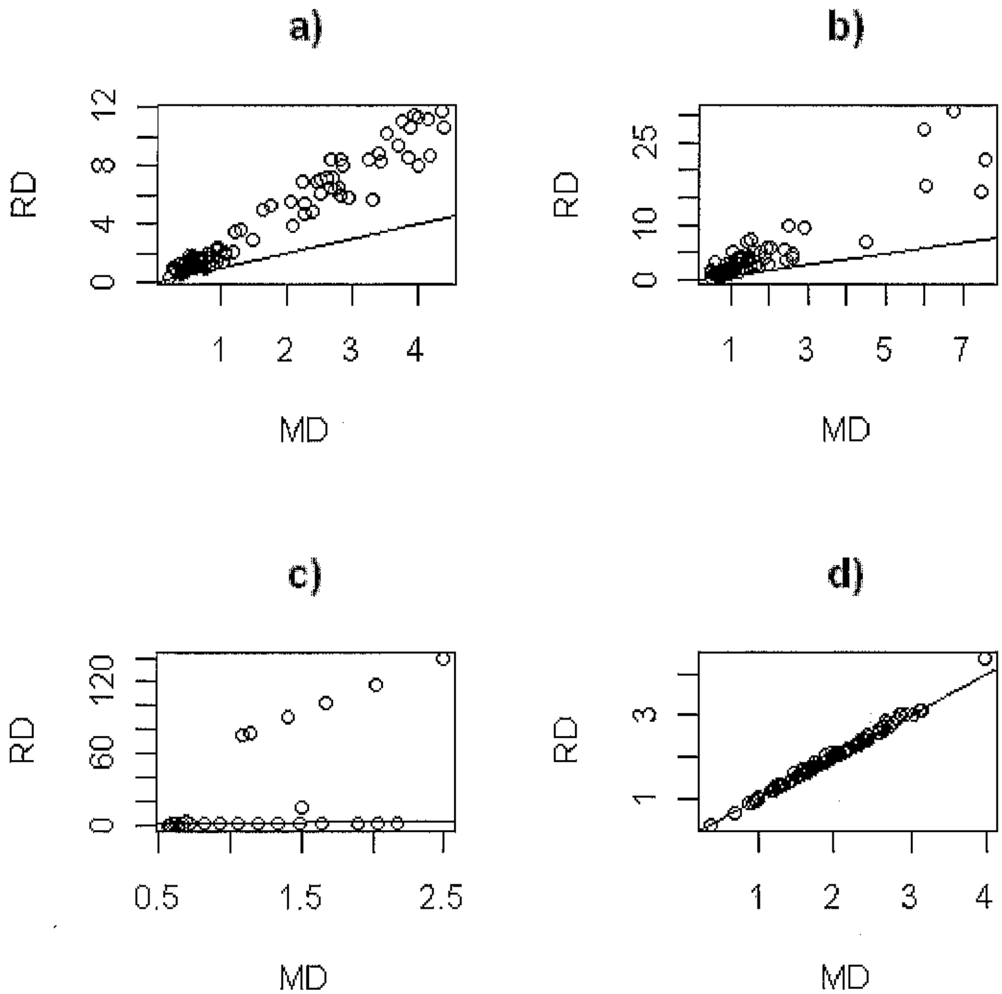


actually $L^{(-1)}$ modified transformation

8) The plot above is shell mass S versus 1/(shell length). Note that λ = 1 for S. Which transformation will increase the linearity of the plot log(S) or S^2? Explain briefly.

log(S) need to spread small values of S

5



above

9) Shown below are 4 DD plots. Classify the data distribution as multivariate normal, elliptically contoured but not multivarite normal, not elliptically contoured or outliers are present. Explain your choices briefly.

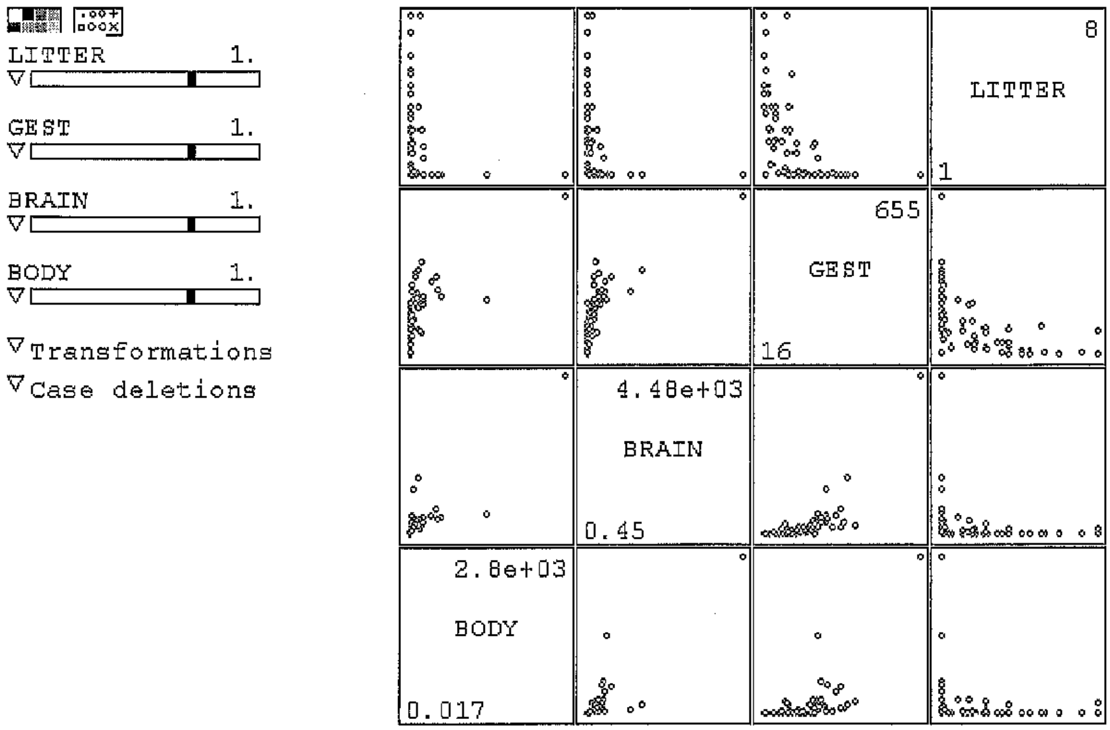
a) EC, points cluster about a line through the origin with slope > 0

b) not EC (or outliers)

c) outliers 2 lines (or not EC)

d) MVN points cluster about the identity line

20



10) Shown above is a plot where the variables are the average body weight, brain weight, gestation period , and litter size for 21 species of mammals.

a) What is the name of the above plot?

scatterplot matrix

e b) Give the transformations, if any, that you would use to remove strong nonlinearities from the variables.

log(body)
log(brain)
log(Gest)

Since $\frac{max}{min} > 10$

$\overline{}$
10