iii) $\hat{\theta}_n$ gets close to θ as $n \to \infty$ for many estimators (CLT).

not a loom very symmetric

Use simulation with large n to approx unknown quantities or as starting values if the quantity can be computed numerically.

HW1 2.23 and 2.24 $n = 10000$

$N(0,1)$ MAD(n) $\approx .669$ MAD(Y) $\approx .6745$

about 2 digits are correct for $n = 10000$ $\sqrt{n} = 100$

1 $n = 100$ $\sqrt{n} = 10$

\# correct $\propto \frac{1}{\sqrt{n}}$

EXP(1) $T_n \approx .7433 \approx \mu_T$

3.8) Simulated 95% confidence intervals

n $Y_1, \ldots, Y_n$ iid $\sim F$

nruns = # of CIs from samples of size n

$CI_i = [\hat{L}_i, \hat{U}_i]$ for $i = 1, \ldots,$ nruns

want i) % of times $\theta \in CI \approx .95$

ii) average length $\dfrac{\sum(\hat{U}_i - \hat{L}_i)}{\text{nruns}}$ to be small compared to other methods

do part of this ex

ex) table 2.6, 2.7 and problem 2.37 on HW 2

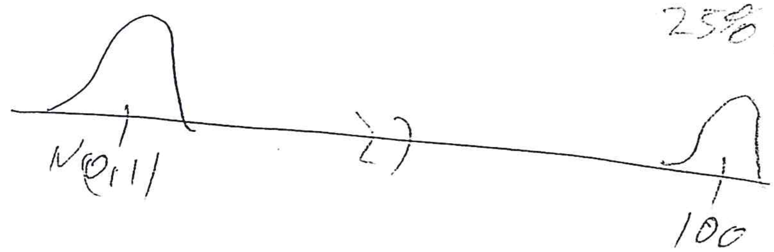
stop $Y_1, \ldots, Y_n$ are iid from one of 5 distributions

($N(0,1) \to$ type = 1 in problem 4.13

The shift or contaminated normal $\to$ type = 2 in problem 2.37 ~~4.13~~

13.9

~~Mixture~~



$$F(y) = .75\,\underbrace{\Phi(y)}_{N(0,1)\text{ cdf}} + .25\,\underbrace{\Phi(y-100)}_{N(100,1)\text{ CDF}}$$

Skip

Problem #37 computes the classical CI

$$\bar{y} \pm \underbrace{t_{n-1,.975}}_{\approx 1.96 \text{ for } n>20} \frac{s}{\sqrt{n}} = [\hat{L}_i, \hat{U}_i] \text{ for } i=1,\dots,500 = \text{nruns}$$

samples of size n. Use $n = 10, 50, 100$

$\underline{\text{COV}}$ = % of runs where $\hat{L}_i \le \theta \le \hat{U}_i$

So if 485 CI's contained θ and 15 did not,

$$\text{COV} = \frac{485}{500} = 0.97$$

$$\underline{\text{len}} = \frac{\sum_{i=1}^{500} \sqrt{n}\,(\hat{U}_i - \hat{L}_i)}{500} = \text{average scaled length}$$

If $Y_1, \dots, Y_n$ are iid with $SD = \sigma_Y$

then $\sqrt{n}\,(\hat{U}_i - \hat{L}_i) = \sqrt{n}\, 2 t_{n-1,.975} \frac{s}{\sqrt{n}} = 2 t_{n-1,.975}\, s$

$\approx 2(1.96)\,\sigma_Y = 3.92\,\sigma_Y$ for large n.

For type I, $\sigma_Y = \sigma = 1$ so len ≈ 3.92 for large n.

$E(Y) = .75(0) + .25(100) = 25$

$EY^2 = .75(1) + .25(1 + (100)^2) = 2501$ so $\sigma_Y^2 = 2501 - (25)^2 = 1876$

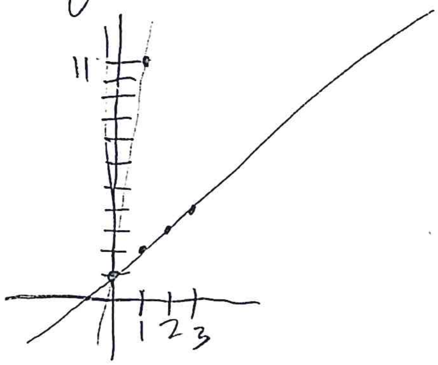
$\sigma_Y = \sqrt{1876} = 43.3128$ and len $\approx 3.92(43.3128) = 169.786$

Then $b_{J_1} = (1,10)'$ so the fitted line is $Y = 1 + 10x$

The ordered squared residuals are $0, 0, 81, (18)^2, (27)^2$

If $J_2 = \{1, 4\}$, then $(0,1)$ and $(3,4)$ are chosen, $b_{J_2} = (1,1)$ and the fitted line is $Y = 1 + X$. The ordered squared residuals are $0, 0, 0, 0, 81$

$\hat{y} = 1+1 = 2, \; Y - \hat{Y} = 11 - 2 = 9$



Let $C_n = 3$.

$$\sum_{i=1}^{C_n} r_i^2(b_{J_1}) = 0 + 0 + 81 = 81$$

$$\sum_{i=1}^{C_n} r_i^2(b_{J_2}) = 0 + 0 + 0 = 0$$

So $\hat{\beta}_{CLTS} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\hat{Y} = 1 + x$.

$k = 0 \quad K = 2$.

6) The attractors $b_{k,1}, \ldots, b_{k,3000}$ have a lot less variability than the starts $b_{0,1}, \ldots, b_{0,3000}$ if $k > 0$. Also $Q_{LTS}(b_{k,j}) \le Q_{LTS}(b_{0,j})$.

7) P 321 The CLTS estimator $\hat{\beta}_{CLTS}$ that uses $K = 3000$ elemental starts is useful for detecting certain types of outliers, but $\hat{\beta}_{CLTS}$ is not consistent.

(Using $\hat{\beta}_{OLS}$ as a start makes $\hat{\beta}_{CLTS}$ consistent.)

Ch 3) 1) p 35 A multivariate location and dispersion (MLD) model is a joint distribution $f(z \mid \mu, \Sigma)$

Co... random ... the ... a $p \times 1$ location (14.5)
vector is $\underset{\sim}{\mu}$ and the $p \times p$ symmetric positive definite dispersion matrix Σ. Hence $P(\underset{\sim}{x} \in A) = \int_A f(\underset{\sim}{z})\,d\underset{\sim}{z}$ for suitable sets A.

ex) $\underset{\sim}{x} = (x_1, x_2, x_3)^T$ where the x_i are iid $N(0, \sigma^2)$.
Then $f(\underset{\sim}{z}) = \prod_{i=1}^{3} f(z_i)$ where $f(z_i) = \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2}$.
Here $\underset{\sim}{\mu} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $\Sigma = \sigma^2 I_3$.

2) Data is $\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_n$ stored in a matrix

$$W = \begin{pmatrix} \underset{\sim}{x}_1^T \\ \vdots \\ \underset{\sim}{x}_n^T \end{pmatrix} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \\ x_{n1} & \cdots & x_{np} \end{pmatrix}, \quad \text{In } R$$

data matrix is of denoted by x.

3) Know p Let $\underset{\sim}{X} = (X_1, \ldots, X_p)^T$. The population mean of a random vector $\underset{\sim}{X}$ is

$$E(\underset{\sim}{X}) = \begin{pmatrix} E X_1 \\ \vdots \\ E X_p \end{pmatrix}$$

and the $p \times p$ pop. covariance matrix is $\Sigma_x = \mathrm{Cov}(\underset{\sim}{X}) = E(\underset{\sim}{X} - E(\underset{\sim}{X}))(\underset{\sim}{X} - E(\underset{\sim}{X}))^T = (\sigma_{ij})$. That is, the ij entry of $\mathrm{Cov}(\underset{\sim}{X})$ is $\mathrm{Cov}(X_i, X_j) = \sigma_{ij}$.

skipped of runs where CI contains θ (or PI contains Y_f).

Then $X = k W \sim bin(k, P_n)$ where hopefully

$\leftarrow$ true coverage prob

$P_n = 1 - \delta_n \approx 1 - \delta \approx 0.95.$

$V(X) = k^2 V(W) = k P_n(1-P_n)$

$V(W) = \dfrac{P_n(1-P_n)}{k}$

$\therefore SE(W) \approx \sqrt{\dfrac{\hat{P}_n(1-\hat{P}_n)}{k}} \approx 0.013$ if $\hat{P}_n \in [0.9, 0.95]$.

$0.95 \pm \underbrace{3(0.013)}_{\approx 0.04} \approx 0.91$ to 0.99 is a reasonable

simulated coverage if $P = 0.95$ = true coverage.

§2.4 40) p40 The metrically trimmed mean

M_n is the sample mean of the cases inside

$[\hat{\theta}_n - k_1 D_n, \hat{\theta}_n + k_2 D_n]$. we will use

$[MED(n) - 6\,MAD(n), MED(n) + 6\,MAD(n)] = (*)$.

If $(*) \to [a, b]$, then M_n estimates

the truncated mean $\mu_T(a,b)$.

§2.12.2 41) p57 Consider intervals $[Y_{(1)}, Y_{(c_n)}], \ldots, [Y_{(n-c_n+1)}, Y_{(n)}]$

used to compute the shorth.

a) The midpoint of the shorth interval is the

ii) Compute $(\overline{Y}_{J_i}, S^2_{J_i})$ for each interval

The MCD estimator $(\overline{Y}_{MCD}, S^2_{MCD})$ is equal to the $(\overline{Y}_{J_j}, S^2_{J_j})$ with the smallest $S^2_{J_i}$.

The LTS estimator $= \overline{Y}_{MCD}$. Let $[Y_{(d)}, Y_{(d+c_n-1)}]$ be the LTS interval.

iii) Compute the sample median M_{J_i} of the c_n cases in the ith interval. Let $Q_{LTA}(M_{J_i}) = \sum_{j \in J_i} |y_j - M_{J_i}|$. $LTA(c_n) = M_{J_j}$ with the smallest $Q_{LTA}(M_{J_i})$. Let the LTA interval be $[Y_{(a)}, Y_{(a+c_n-1)}]$.

↑ stopped

15.5

4) Let $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ be $p \times 1$ random vectors, $\underset{\sim}{a}$ a conformable constant vector and let A and B be conformable constant matrices.

Then $E(\underset{\sim}{X} + \underset{\sim}{Y}) = E(\underset{\sim}{X}) + E(\underset{\sim}{Y})$

$E(\underset{\sim}{a} + \underset{\sim}{X}) = \underset{\sim}{a} + E(\underset{\sim}{X})$

$E(A \underset{\sim}{X} B) = A\, E(\underset{\sim}{X})\, B$ and

$\text{Cov}(A\underset{\sim}{X}) = \text{cov}(\underset{\sim}{a} + A\underset{\sim}{x}) = \underset{g \times p}{A}\ \text{cov}(\underset{\sim}{X})\ \underset{p \times g}{A^T}$.

5) A $p \times 1$ random vector $\underset{\sim}{X}$ has a <u>p-dimensional multivariate normal distribution</u>

$\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, iff $\underset{\sim}{t}^T \underset{\sim}{X}$ has a univariate normal distribution for any $p \times 1$ vector $\underset{\sim}{t}$.

$E(\underset{\sim}{X}) = \underset{\sim}{\mu}$ and $\text{Cov}(\underset{\sim}{X}) = \Sigma$.

6) Usually we want Σ to be positive definite

then $\underset{\sim}{X}$ has pdf

$$f(\underset{\sim}{z}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left[-\frac{1}{2} (\underset{\sim}{z} - \underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{z} - \underset{\sim}{\mu})\right]$$

where $|\Sigma| = \det(\Sigma)$.

ex) If $p = 1$, $\Sigma = \sigma^2$ and $(\underset{\sim}{z} - \underset{\sim}{\mu})^T \Sigma^{-1} (\underset{\sim}{z} - \underset{\sim}{\mu}) = \dfrac{(z-\mu)^2}{\sigma^2}$.

7) If $\underset{\sim}{X} = (X_1, \ldots, X_p)^T$ where the X_i are ind $N(\mu_i, \sigma_i^2)$

then $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ where $\underset{\sim}{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix}$ and $\Sigma = \text{diag}(\sigma_i^2)$.

matrix, then $A\underset{\sim}{X} \sim N_q(A\underline{\mu}, A\Sigma A^T)$. (16.5)

If $\underline{a}$ is $p\times 1$ then $\underset{\sim}{a}+\underline{x} \sim N_p(\underline{\mu}+\underset{\sim}{a}, \Sigma)$.

9) p3 * If $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$, then all subsets are MVN! $(X_{k_1}, \ldots, X_{k_q})^T \sim N_q(\tilde{\mu}, \tilde{\Sigma})$

where $\tilde{\mu}_i = EX_{k_i}$ and $\tilde{\Sigma}_{ij} = \mathrm{cov}(X_{k_i}, X_{k_j})$.

$\boxed{\text{ex}}$ know for exam 1 Q3

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \sim N_3\left[\begin{pmatrix} 1 \\ 17 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right]$$

Find the dist of $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\begin{pmatrix} x_1 \\ x_3 \end{pmatrix}$ & $\begin{pmatrix} x_2 \\ x_3 \end{pmatrix}$.

Soln $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N_2\left[\begin{pmatrix} 1 \\ 17 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}\right]$

$$\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} \sim N_2\left[\begin{pmatrix} 1 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\right]$$

$$\begin{pmatrix} x_2 \\ x_3 \end{pmatrix} \sim N_2\left[\begin{pmatrix} 17 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\right]$$

10) Let $\underline{X} = \begin{pmatrix} \underline{X}_1 \\ \underline{X}_2 \end{pmatrix}$ ($\underline{X}_1$: $q\times 1$, $\underline{X}_2$: $(p-q)\times 1$), $\underline{\mu} = \begin{pmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$

p3

17) know for E1 — If $\underline{x} \sim N_p(\underline{\mu}, \Sigma)$ then $\underline{X}_1 \sim N_q(\underline{\mu}_1, \Sigma_{11})$ and $\underset{\sim}{x}_2 \sim N_{p-q}(\underline{\mu}_2, \Sigma_{22})$.

12) ... then 19

$\underline{X}_1$ and $\underline{X}_2$ are independent ($\underline{X}_1 \perp\!\!\!\perp \underline{X}_2$) iff $\Sigma_{12} = 0$

matrix of 0's

13) p356 However, if the joint dist of $\underline{X}_1$ and $\underline{X}_2$ is not multivariate normal, $\underline{X}_1$ and $\underline{X}_2$ could be uncorrelated but not independent.

14) p357 Know for E1. The pop correlation between 2 random variables X and Y

is $\rho(X,Y) = \dfrac{cov(X,Y)}{\sqrt{var(X)\ var(Y)}} = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y} = \rho(Y,X)$

if $\sigma_x > 0$, $\sigma_y > 0$.

If $\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_2(\underline{\mu}, \Sigma)$ then

$$\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_2\left[\begin{pmatrix} \mu_Y \\ \mu_X \end{pmatrix}, \begin{pmatrix} \sigma_Y^2 & \sigma_{YX} \\ \sigma_{XY} & \sigma_X^2 \end{pmatrix}\right]$$ where

$\sigma_{YX} = cov(Y,X) = cov(X,Y) = \sigma_{XY}$.

ex) $$\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3\left[\begin{pmatrix} 1 \\ 17 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right]$$

Which variables are ind? $X_1 \perp\!\!\!\perp X_3$ and $X_2 \perp\!\!\!\perp X_3$

i) $\rho(X_1, X_2) = \dfrac{1}{\sqrt{4}\sqrt{3}} = 0.2887$

ii) $\rho(X_1, X_3) = 0$

iii) $\rho(X_2, X_3) = 0$

$\underset{p\times 1}{Y} \sim N_p(\underset{\sim}{\mu}, \sigma^2 I)$

Find the distribution of $\underset{q\times p}{A}\, \underline{Y}$

Soln] $A\underset{\sim}{Y} \sim N_q(\underline{\mu}_A, \Sigma_A)$ where

$$\underline{\mu}_A = A\underset{\sim}{\mu} \quad \text{and } \Sigma_A = \sigma^2 A I A^T = \sigma^2 A A^T$$

↑ end E1 material

15] **Know for E2** If $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$

then the conditional distribution

$$X_1 \mid X_2 = x_2 \sim N_q\Big[\underset{\sim}{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underline{x}_2 - \underset{\sim}{\mu}_2),\ \underbrace{\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}_{\text{free of } \underline{X}_2 = \underline{x}_2}\Big]$$

Notation] $X_1 \mid X_2 \sim N_q\Big[\underbrace{\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}_2 - \mu_2)}_{\text{family of means depending on the value of } \underline{X}_2},\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Big]$

16] know: $\text{Cov}(\underline{x}, \underline{y}) = E[(\underline{x} - E\underline{x})(\underline{y} - E\underline{y})^T] = [\text{Cov}(\underset{\sim}{y}, \underset{\sim}{x})]^T$

ex] Let

$$\begin{pmatrix} Y \\ X_1 \\ \vdots \\ X_p \end{pmatrix} \sim N_{p+1}\left[\begin{pmatrix} E(Y) \\ E(\underline{X}) \end{pmatrix}, \begin{pmatrix} \underset{1\times 1}{\text{Var } Y} & \underset{1\times p}{\Sigma_{Y\underline{X}}} \\ \Sigma_{\underset{\sim}{X}Y} & \underset{p\times p}{\Sigma_{\underline{X}\underline{X}}} \end{pmatrix}\right]$$

where $\Sigma_{\underline{X}\underline{X}} = \text{cov}(\underline{X})$ and $\Sigma_{Y\underline{X}} = \text{Cov}(Y, \underline{X}) = E\big(\underbrace{(Y - EY)}_{\text{scalar}}\underbrace{(\underline{X} - E\underline{X})^T}_{1\times p}\big)$

Then $E(Y \mid \underline{X} = \underline{x}) = E(Y) + \Sigma_{Y\underline{X}}\Sigma_{\underline{X}\underline{X}}^{-1}(\underline{x} - E\underline{X})$

and $\text{Var}(Y \mid \underline{X} = \underline{x}) = \text{Var}(Y) - \Sigma_{Y\underline{X}}\Sigma_{\underline{X}\underline{X}}^{-1}\Sigma_{\underline{X}Y}$

both are scalars

$\ldots \Sigma_{YX} \Sigma_{XX}^{-1} E(X) + \Sigma_{YX} \Sigma_{XX}^{-1} x$

$$= \alpha + \beta^T x$$

and $Var(Y \mid X = x) \equiv \sigma^2$ is a constant free of X

ex3) bivariate normal $\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_2 \left(\begin{pmatrix} E(Y) \\ E(X) \end{pmatrix}, \begin{pmatrix} Var(Y) & cov(X,Y) \\ cov(X,Y) & Var(X) \end{pmatrix} \right)$

since $cov(X,Y) = cov(Y,X)$.

Then $Y \mid X = x \sim N_1 \left(E(Y \mid X=x), Var(Y \mid X=x) \right)$

where $E(Y \mid X = x) = EY + cov(X,Y) \frac{1}{var(X)} (x - EX)$

$$= E(Y) + \rho(X,Y) \sqrt{\frac{Var(Y)}{Var(X)}} \, (x - E(X))$$

$$\rho(X,Y) = \frac{cov(X,Y)}{\sqrt{Var(X) \, Var(Y)}}$$

and $var(Y \mid X = x) = Var(Y) - cov(X,Y) \frac{1}{var(X)} cov(X,Y)$

$$= Var(Y) - \rho(X,Y) \sqrt{\frac{Var(Y)}{Var(X)}} \; \rho(X,Y) \sqrt{Var(X)} \sqrt{Var(Y)}$$

$$= Var(Y) - [\rho(X,Y)]^2 \, Var(Y) = Var(Y) \left[1 - (\rho(X,Y))^2 \right].$$

Notice that the mean function is a line

$$E(Y \mid X = x) = E(Y) - \rho(X,Y) E(X) + \rho(X,Y) \sqrt{\frac{Var(Y)}{Var(X)}} \; x = \alpha + \beta x$$

and when $X = E(X)$, the line passes through $(E(X), E(Y))$.

then Y changes by $\rho(X,Y)\sqrt{Var(Y)}$.

Let $Z_x = \dfrac{x - E(X)}{\sqrt{Var(X)}}$ so that x is Z_x $\underbrace{\text{standard deviations}}_{SD(X)}$ away from $E(X)$. Then

$$E(Y \mid X = x) = E(Y) + \rho(X,Y)\sqrt{Var(Y)}\, Z_x$$

is $\rho(X,Y)\, Z_x$ $\underbrace{\text{standard deviations}}_{SD(Y)}$ from $E(Y)$.

17) A $p \times 1$ random vector $\underline{X}$ has an elliptically contoured (or symmetric) (EC) distribution if $\underline{X}$ has pdf $f(\underline{z}) = k_p |\Sigma|^{-\frac{1}{2}}\, g\left((\underline{z}-\underline{\mu})^T \Sigma^{-1} (\underline{z}-\underline{\mu})\right)$ where k_p is a constant and g is known. We write $\underline{X} \sim EC_p(\underline{\mu}, \Sigma, g)$.

18) If the 2nd moments exist, $E(\underline{X}) = \mu$ and $Cov(\underline{X}) = c_x \Sigma$ for some constant $c_x > 0$.

19) The population squared Mahalanobis distance

$U \equiv D^2(\underline{\mu}, \Sigma) = (\underline{X} - \underline{\mu})^T \Sigma^{-1} (\underline{X} - \underline{\mu})$ has pdf

$$h(u) = \frac{\pi^{p/2}}{\Gamma(p/2)}\, k_p\, u^{p/2 - 1}\, g(u).$$

ex) * MVN $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ is EC with $g(u) = e^{-u/2}$ and $h(u)$ is the χ^2_p pdf.

20)

i) If $E(\underline{X})$ exists and B is a constant $p \times r$ full rank matrix where $1 \le r \le p$, then $\underline{X}$ is EC iff for all such conforming matrices B,

$$E(\underline{X} \mid B^T \underline{X}) = \underline{\mu} + M_B B^T (\underline{X} - \underline{\mu}) \quad \text{for}$$

some $p \times r$ constant matrix M_B.

ii) If $\underline{X} \sim EC$ and $E(\underline{X})$ exists, any subset of $\underline{X}$ is EC

iii) If $\binom{Y}{\underline{X}}$ EC and $E\binom{Y}{\underline{X}}$ exists then (scalar)

$$E(Y \mid \underline{X}) = \alpha + \underline{\beta}^T \underline{X} \quad \text{where}$$

$$\alpha = \mu_Y - \underline{\beta}^T \underline{\mu}_X \quad \text{and} \quad \underline{\beta} = \underline{\Sigma}_{XX}^{-1} \underline{\Sigma}_{XY}.$$

Same location but scaled shape

ex) Suppose $\underline{X} \sim (1-\gamma) N_p(\underline{\mu}, \Sigma) + \gamma N_p(\underline{\mu}, c\Sigma)$ where $c > 0$ and $0 < \gamma < 1$. Show that $\underline{X}$ is EC.

Soln: Want to use 20 i). Since a MVN is EC (and $X \overset{D}{=} (1-\gamma) W_1 + \gamma W_2$ where $W_1 \sim N_p(\mu, \Sigma)$ and $W_2 \sim N_p(\underline{\mu}, c\Sigma)$)

$$E[\underline{X} \mid B^T \underline{X}] = (1-\gamma)[\underline{\mu} + M_1 B^T(\underline{X} - \underline{\mu})] + \gamma[\underline{\mu} + M_2 B^T(\underline{X} - \underline{\mu})]$$

$$= \underline{\mu} + \underbrace{[(1-\gamma) M_1 + \gamma M_2]}_{M} B^T (\underline{X} - \underline{\mu}) = \underline{\mu} + M B^T (\underline{X} - \underline{\mu})$$

and the result follows by 20 i)

→ See HW 6 3.~~4~~ (was 10.4)

where 19.5

$\underline{x}_1, \ldots, \underline{x}_n$ are the data. The ith squared (sample) Mahalanobis distance

$$D_i^2 = D_i^2(T(W), C(W)) = (\underset{\sim}{x}_i - T(W))^T C^{-1}(W) (x_i - T(W))$$

where $T(W)$ is a multivariate location estimator and $C(W)$ is a (covariance matrix) dispersion estimator.

22) $D^2_{\underline{x}}(\underline{\mu}, \Sigma) = (\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})$ is the pop squared Mahalanobis distance and

$\Sigma^{-1/2}(\underline{x} - \underline{\mu})$ is the $p$dimensional analog to the z score $z = \frac{x - \mu}{\sigma}$, so

D_i is the analog to the sample z score $z_i = \frac{x_i - \bar{x}}{\hat{\sigma}}$.

23) $D_i(T, I_p)$ is the Euclidean distance of $\underline{x}_i$ from T.

24) The classical distance MD_i uses

$$T(W) = \underset{\sim}{\bar{x}} = \frac{1}{n} \sum_{i=1}^{n} \underline{x}_i \quad \text{(sample mean) and}$$

$$C(W) = S = \frac{1}{n-1} \sum_{i=1}^{n} (\underline{x}_i - T(W))(x_i - T(W))^T$$

(sample covariance matrix).