ex] If $p=1$, $\Sigma = \sigma^2$, $(z-\mu)^T \Sigma^{-1} (z-\mu) = \frac{(z-\mu)^2}{\sigma^2}$

usual $N(\mu, \sigma^2)$ pdf

4] 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 $\Sigma = \text{diag}(\sigma_i^2)$.

5] * p30-31 If $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ and A is a $q \times p$ constant matrix, then $A\underset{\sim}{x} \sim N_q(A\underset{\sim}{\mu}, A\Sigma A^T)$.

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

6] * p 31 If $\underset{\sim}{x} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, then all subsets are MVN: $(x_{k1}, \ldots, x_{kq})^T \sim N_q(\tilde{\mu}, \tilde{\Sigma})$ where $\tilde{\mu}_i = E x_{k_i}$ and $\tilde{\Sigma}_{ij} = \text{Cov}(x_{k_i}, x_{k_j})$.

ex] <u>Know for Exam 1, Q 2</u>

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

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

Soln} $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_2\left[\begin{pmatrix} 1 \\ 17 \end{pmatrix}, \begin{pmatrix} 4 & 1 \\ 1 & 3 \end{pmatrix}\right]$

$\begin{pmatrix} X_1 \\ X_3 \end{pmatrix} \sim N_2\left[\begin{pmatrix} 1 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}\right]$

$\begin{pmatrix} X_2 \\ X_3 \end{pmatrix} \sim N_2\left[\begin{pmatrix} 17 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}\right]$

7} Let $\underset{\sim}{X} = \begin{pmatrix} \underset{\sim}{X}_1^{\,q\times 1} \\ \underset{\sim}{X}_2{}_{\,(p-q)\times 1} \end{pmatrix}$, $\underset{\sim}{\mu} = \begin{pmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$,

8} p31 <u>Know for E1, Q2</u> If $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, then

$\underset{\sim}{X}_1 \sim N_q(\underset{\sim}{\mu}_1, \Sigma_{11})$ and $\underset{\sim}{X}_2 \sim N_{p-q}(\underset{\sim}{\mu}_2, \Sigma_{22})$.

9] p31 <u>Know for E1, Q2</u> If $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$, then

$\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ are independent ($\underset{\sim}{X}_1 \perp\!\!\!\perp \underset{\sim}{X}_2$) iff (if and only if) $\Sigma_{12} = 0$ (matrix of 0s)

10} p32 However, if the joint distribution of $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ is <u>not</u> multivariate normal, $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ could be uncorrelated but not independent.

11) p32 <u>Know for E1, Q2</u> The pop correlation between 2 random variables X and Y is

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

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

If $\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_2(\underset{\sim}{\mu}, \Sigma)$, 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] $\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3\left[\begin{pmatrix} 1 \\ 17 \\ 8 \end{pmatrix}, \begin{pmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right]$

a) which variables are ind? $X_1 \perp\!\!\!\perp X_3$ and $X_2 \perp\!\!\!\perp X_3$

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

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

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

ex] <u>Know for E1</u> $\underset{p\times 1}{\underset{\sim}{Y}} \sim N_p(\underset{-}{\mu}, \sigma^2 I)$. Find the distribution of $\underset{q\times p}{A}\underset{\sim}{Y}$.

Soln: $A\underset{\sim}{Y} \sim N_q(A\underset{\sim}{\mu}, A\sigma^2 I A^T) \sim N_q(A\underset{\sim}{\mu}, \sigma^2 A A^T)$.

12] Know for E1 p31 If $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ then the conditional distribution (9.5)

I usually give these formulas.

$$\underset{\sim}{X}_1 | \underset{\sim}{X}_2 = \underset{\sim}{x}_2 \sim N_q\left[\underset{\sim}{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{x}_2 - \underset{\sim}{\mu}_2),\ \underbrace{\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}}_{\text{free of } \underset{\sim}{X}_2 = \underset{\sim}{x}_2}\right]$$

Notation] $$\underset{\sim}{X}_1 | \underset{\sim}{X}_2 \sim N_q\left[\underbrace{\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\underset{\sim}{X}_2 - \underset{\sim}{\mu}_2)}_{\text{family of means depending on } \underset{\sim}{X}_2},\ \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right]$$

ex] Let $$\begin{pmatrix} Y \\ x_1 \\ \vdots \\ x_p \end{pmatrix} \sim N_{p+1}\left[\begin{pmatrix} E(Y) \\ E(\underset{\sim}{X}) \end{pmatrix}, \begin{pmatrix} \underset{1\times 1}{\mathrm{Var}(Y)} & \underset{1\times p}{\Sigma_{Y\underset{\sim}{X}}} \\ \Sigma_{\underset{\sim}{X}Y} & \underset{p\times p}{\Sigma_{\underset{\sim}{X}\underset{\sim}{X}}} \end{pmatrix}\right]$$

where $\Sigma_{\underset{\sim}{X}\underset{\sim}{X}} \overset{= \Sigma_{\underset{\sim}{X}}}{=} \mathrm{Cov}(\underset{\sim}{X})$ and $\Sigma_{Y\underset{\sim}{X}} \overset{= \mathrm{cov}(Y, \underset{\sim}{X})}{=} E[\underbrace{(Y - EY)}_{\text{scalar}}(\underset{\sim}{X} - E\underset{\sim}{X})^T]$

Then $E(Y | \underset{\sim}{X} = \underset{\sim}{x}) = EY + \Sigma_{Y\underset{\sim}{X}}\Sigma_{\underset{\sim}{X}\underset{\sim}{X}}^{-1}(\underset{\sim}{x} - E\underset{\sim}{X})$

$= EY - \Sigma_{Y\underset{\sim}{X}}\Sigma_{\underset{\sim}{X}\underset{\sim}{X}}^{-1}E\underset{\sim}{X} + \Sigma_{Y\underset{\sim}{X}}\Sigma_{\underset{\sim}{X}\underset{\sim}{X}}^{-1}\underset{\sim}{x} = \alpha + \underset{\sim}{\beta}^T\underset{\sim}{x}$

and $V(Y | \underset{\sim}{X} = \underset{\sim}{x}) = \mathrm{Var}\, Y - \Sigma_{Y\underset{\sim}{X}}\Sigma_{\underset{\sim}{X}\underset{\sim}{X}}^{-1}\Sigma_{\underset{\sim}{X}Y} \equiv \sigma^2,$

a constant free of $\underset{\sim}{x}$. MLR

ex] bivariate normal $$\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_2\left[\begin{pmatrix} E(Y) \\ E(X) \end{pmatrix}, \begin{pmatrix} \mathrm{Var}(Y) & \mathrm{cov}(X, Y) \\ \mathrm{cov}(X, Y) & \mathrm{var}(X) \end{pmatrix}\right]$$

since $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$. Then

$Y | X = x \sim N_1[E(Y | X = x),\ V(Y | X = x)]$ where

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

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

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

$= \alpha + \beta x$, so the mean function is a line that passes through $(E(X), E(Y))$. SLR

$V(Y|X=x) = Var(Y) - Cov(X,Y) \frac{1}{var(X)} cov(X,Y) =$

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

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

§3.2

13] p33 A $p \times 1$ random vector $\underset{\sim}{X}$ has an elliptically contoured (EC) or elliptically symmetric distribution if $\underline{X}$ has pdf

$f(\underline{z}) = k_p |\Sigma|^{-\frac{1}{2}} g[(\underline{z}-\underline{\mu})^T \Sigma^{-1} (\underline{z}-\underline{\mu})]$ where

k_p is a constant and g is a known function.

Notation: $\underset{\sim}{X} \sim EC_p(\underset{\sim}{\mu}, \Sigma, g)$.

14] p33 If 2nd moments exist, $E(\underline{X}) = \underline{\mu}$

and $Cov(\underline{X}) = \Sigma_{\underline{X}} = c_X \Sigma$ for some constant $c_X > 0$.

15] p33-4 * The population squared Mahalanobis distance

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

$h(u) = \frac{\pi^{p/2}}{\Gamma(\frac{p}{2})} k_p u^{\frac{p}{2}-1} g(u)$ if $\underline{X} \sim EC_p(\underset{\sim}{\mu}, \Sigma, g)$.

ex] * MVN $\underline{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ is EC with $g(u) = e^{-u/2}$

and $h(u)$ is the χ^2_p pdf.

16] p 34-6 properties i) If $E(\underline{X})$ exists and B (10.5)
is a constant full rank matrix with $1 \le r \le p$,
then $\underline{X}$ is EC iff for all such conforming
matrices B, $E(\underline{X} | B^T \underline{X}) = \underset{\sim}{\mu} + M_B \, B^T(\underline{X} - \underline{\mu})$
where $M_B = \Sigma B (B^T \Sigma B)^{-1}$ is a $p \times r$ constant matrix.

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

iii) If $\begin{pmatrix} Y \\ \underset{\sim}{x} \end{pmatrix} \sim EC$ and $E\begin{pmatrix} Y \\ \underset{\sim}{x} \end{pmatrix}$ exists, then ← scalar

$E(Y | \underline{X}) = \alpha + \underline{\beta}^T \underline{X}$ where $\alpha = \mu_Y - \underline{\beta}^T \underline{\mu}_X$ and $\underline{\beta} = \Sigma_{XX}^{-1} \Sigma_{XY}$.

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

Soln] Since MVN is EC, by 16] i) $E(\underline{X} | B^T \underline{X})$
$= (1-\gamma)(\underset{\sim}{\mu} + M_B B^T(\underline{X} - \underline{\mu})) + \gamma (\underset{\sim}{\mu} + M_B \, B^T(X-\mu))$
$= \underline{\mu} + M_B \, B^T(\underline{X} - \underline{\mu})$ and the result follows by
16] i.

§ 3.3 17]* p38 $D^2_{\underline{x}}(T, C) = (\underline{x} - T)^T C^{-1} (x - T)$ and
$D_i^2(T, C) = D^2_{\underline{x}_i}$ is the ith squared Mahalanobis distance.
Here $(T, C) = (T(W), C(W))$;
T is a $p \times 1$ multivariate location estimator, $C > 0$ is a $p \times p$ dispersion estimator.

18] $D^2_{\underline{x}}(\underline{\mu}, \Sigma) = (\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})$ 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 $|z \text{ score}|$: $\left|\frac{x_i - \bar{x}}{\hat{\sigma}}\right|$.

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

20] The classical distance is $D_i(\bar{\underline{x}}, S)$ and will sometimes be denoted as MD_i while $D_i(T,C)$ " " RD_i.

---

§3.4 Skim Large Sample Theory, Univariate then multivariate

21] Know p 39 Central Limit Theorem (CLT) Let $Y_1, \ldots, Y_n$ be iid with $E Y_i = \mu$ and $V(Y_i) = \sigma^2$. Then $\sqrt{n}(\bar{Y} - \mu) \xrightarrow{D} N(0, \sigma^2)$.

22] If $W_n \xrightarrow{D} X$, then X is the asymptotic distribution or <u>limiting distribution</u> of W_n. X <u>does not</u> depend on n.

The approximate distribution of $\bar{Y}$ is $\bar{Y} \sim AN(\mu, \frac{\sigma^2}{n})$ which does depend on n. $\bar{Y} \approx N(\mu, \frac{\sigma^2}{n})$

23] p42 Let $\{Z_n, n = 1, 2, \ldots\}$ be a sequence of random variables with CDFs (cdfs) F_n and let X be a RV with cdf F. Then Z_n <u>converges in distribution</u> (or law) to X, written $Z_n \xrightarrow{D} X$, if $\lim_{n\to\infty} F_n(t) = F(t)$ at each continuity point of F.

24) If $X_n \xrightarrow{D} X$, then $P(a < X_n \le b) = F_n(b) - F_n(a)$ $\to F(b) - F(a) = P(a < X \le b)$ if F is continuous at a and b. So F can be used to approximate probabilities and percentiles.

Read ex's 3.8 and 3.9 carefully.

25] * p44 X_n <u>converges in probability</u> to X, $X_n \xrightarrow{P} X$, if $\lim_{n\to\infty} P(|X_n - X| \ge \varepsilon) = 0$ for all $\varepsilon > 0$

$\iff \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1 \quad \forall \varepsilon > 0$.

parameter space

26] * p44 A sequence of estimators T_n is consistent for $\tau(\theta)$ if $T_n \xrightarrow{P} \tau(\theta)$ $\forall \theta \in \Theta$. Then T_n is a consistent estimator of $\tau(\theta)$.

27] p46 ⇒ i) If $\lim_{n\to\infty} V_\theta(T_n) = 0$ $\forall \theta \in \Theta$ and $\lim_{n\to\infty} E_\theta(T_n) = \tau(\theta)$, $\forall \theta \in \Theta$, then T_n is a consistent estimator of $\tau(\theta)$.

ii) Let $0 < \delta \le 1$. If $n^\delta (T_n - \tau(\theta)) \xrightarrow{D} N(0, v(\theta))$ $\forall \theta \in \Theta$, then T_n is a consistent estimator of $\tau(\theta)$.