EC402 (2023/24)

RAGVIR'S SPEAKING NOTES FOR CLASS

Vassilis PS1 (WK 2, AT)

(Q1)

Transformations of RVs : $Y = g(X)$ where $g(\cdot)$ is a deterministic function of X.

(a) "Manual" method :

Say $P(X=x) = \begin{cases} 1/3, & x = -1, 0, 1 \\ 0, & o/w \end{cases}$

and $Y := X^2 + 2$, what is $P(Y=y)$ for all $y \in \mathbb{R}$?

x	-1	0	1
$P(Y=y)$	$1/3$	$1/3$	$1/3$
y	3	2	3

- $P(Y=2) = P(X=0) = 1/3$
- $P(Y=3) \overset{CE}{=} P\big((X=-1) \cup (X=1)\big)$

  $\overset{ME}{=} P(X=-1) + P(X=1)$

  $= 1/3 + 1/3$

  $= 2/3$
- $P\big(((Y=2) \cup (Y=3))^c\big) = 0$

$$\Rightarrow P(Y=y) = \begin{cases} 1/3, & y=2 \\ 2/3, & y=3 \\ 0, & o/w \end{cases}$$

(b) "First principles" method:

Say $f_X(x) = \begin{cases} 1/2 \,, & x \in (-1,1) \\ 0 \,, & o/w \end{cases}$ and $Y := X^2 + 2$.

Anyone feel like making a table?! Exactly! So we need to think.

$$\begin{aligned}
P(Y \le y) &= P(X^2 + 2 \le y) \\
&= P(X^2 \le y - 2) \\
&= P(|X| \le \sqrt{y-2}) \\
&= P(-\sqrt{y-2} \le X \le \sqrt{y-2}) \\
&= P(X \le \sqrt{y-2}) - P(X \le -\sqrt{y-2}) \\
&= F_X(\sqrt{y-2}) - F_X(-\sqrt{y-2}) \quad \text{--- } (1)
\end{aligned}$$

Brilliant! let's work out $F_X(x)$:

$$\int_{-1}^{x} 1/2 \, dt = t/2 \Big|_{-1}^{x} = \frac{x}{2} + \frac{1}{2} = \frac{x+1}{2}$$

$$\therefore \; F_X(x) = \begin{cases} 0, & x < -1 \\ \frac{x+1}{2}, & -1 \le x < 1 \\ 1, & o/w \end{cases} \qquad \text{— (2)}$$

Notice also if $x \in (-1,1)$, $y \in (2,3)$ — (3)

(1) & (2) & (3) $\Rightarrow$

$$F_Y(y) = \begin{cases} 0, & y < 2 \\ \frac{\left[\sqrt{y-2}\right]+1}{2} - \left(\frac{\left[-\sqrt{y-2}\right]+1}{2}\right), & 2 \le y < 3 \\ 1, & o/w \end{cases}$$

$$= \begin{cases} 0 & , y < 2 \\ \sqrt{y-2} & , 2 \le y < 3 \\ 1 & , o/w \end{cases}$$

$$\therefore \quad f_Y(y) = \begin{cases} (1/2)(y-2)^{-1/2} & , 2 \le y < 3 \\ 0 & , o/w \end{cases}$$

(Test yourself: Would "method (b)" work for the example in "method (a)"?)

Q: But, Raquir, isn't there some sort of formula?

A: Well, fine .... but be careful ... it is only valid in specific settings...

TRANSFORMATION FORMULA :

Theorem : $X \sim f_X(x)$; $Y := g(X)$ where $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$

Suppose $g(.)$ is a monotone function and that $g^{-1}(.)$ is continuously differentiable on $\mathcal{Y}$

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{\partial}{\partial y} g^{-1}(y) \right| , & y \in \mathcal{Y} \\ 0 , & o/w \end{cases}$$

Proof : let $g^{-1}(.)$ be increasing : eg: $Y := e^X$

$$F_Y(y) = P(Y \leq y) = P(g^{-1}(Y) \leq g^{-1}(y)) = P(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

So, $f_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$

let $g^{-1}(\cdot)$ be decreasing :

$$F_Y(y) = P(Y \leq y) = P\left(g^{-1}(Y) \geq g^{-1}(y)\right) = P\left(X \geq g^{-1}(y)\right) = 1 - F_X\left(g^{-1}(y)\right)$$

$$\text{So } f_Y(y) = f_X\left(g^{-1}(y)\right) \cdot -\frac{d}{dy} g^{-1}(y)$$

Putting both cases together, the result follows for any monotone $g(\cdot)$.

(c) "Formula" method:

Given $X \sim \text{Unif}(-1,1)$, ASSUME $x \in [0,1)$

- Transformation: $y := x^2 + 2$
- Inverse trans.: $x = \sqrt{y-2}$
- Range of transformed var.: $x \in (-1,1) \Rightarrow y \in (2,3)$
- $|\text{Jacobean}|$: $\left| \frac{1}{2}(y-2)^{-1/2} \right|$ → ("$|\cdot|$" is "absolute value" here.)

Now, for $x \in [-1,0)$, by symmetry, the analysis will be the same

$$\therefore \; f_Y(y) = \begin{cases} 2 \cdot \frac{1}{2} \cdot \frac{1}{2}(y-2)^{-1/2}, & 2 < y < 3 \\ 0, & \text{o/w} \end{cases}$$

(Q2)

$$E(aX+bY) := \int_y \int_x (ax+by)\cdot f_{X,Y}(x,y)\, dx\, dy$$

$$= a\int_y\int_x x\, f_{X,Y}(x,y)\, dx\, dy + b\int_y\int_x y\, f_{X,Y}(x,y)\, dx\, dy$$

$$= a\int_x x \int_y f_{X,Y}(x,y)\, dy\, dx + b\int_y y\int_x f_{X,Y}(x,y)\, dx\, dy$$

$$= a\int_x x\, f_X(x)\, dx + b\int_y y\cdot f_Y(y)\, dy$$

$$= a\,E(X) + b\,E(Y)$$

Q3 (a)

ASSUME $X \perp\!\!\!\perp Y$

THEN...

$$E[X|Y] := \int_x x \cdot f_{X|Y=y}(x)\, dx$$

$$\overset{\text{Def}^n}{=} \int_x x\, f_{X,Y}(x,y) / f_Y(y)\, dx$$

$$\overset{X \perp\!\!\!\perp Y}{=} \int_x x\, f_X(x)\, dx$$

$$= E[X]$$

KEY POINT: FULL INDEP. $\Rightarrow$ MEAN INDEP

$\not\Leftarrow$

(b) ASSUME $E[X|Y] = E[X]$ — (1)

THEN...

$$E[(X - E(X))(Y - E(Y))] \overset{\text{why?}}{=} \int_y \int_x xy\, f_{X,Y}(x,y)\, dx\, dy - \int_x x f_X(x)\, dx \int_y y f_Y(y)\, dy$$

eg: Can you prove that $Var(X) = \ldots = E[X^2] - [E(X)]^2$?

$$= \int_y \int_x xy\, f_{X|Y=y}(x) f_Y(y)\, dx\, dy - E(X)E(Y)$$

$$= \int_y y f_Y(y) \int_x x f_{X|Y=y}(x)\, dx\, dy - E(X)E(Y)$$

$$= \int_y y f_Y(y)\, E[X|Y=y]\, dy - E(X)E(Y)$$

$$\overset{\text{ASSUMP. (1)}}{=} E(X)E(Y) - E(X)E(Y) = 0$$

Again, the reverse is not true.

(Q4)

Step 1 : let's recall the univariate case ...

Say $Z \sim N(0,1)$ and $X := \mu + \sigma Z$

Then $M_Z(t) := E\left[e^{tZ}\right]$

$$= \int_{-\infty}^{\infty} e^{tz} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \, dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[z^2 - 2tz\right]} \, dz$$

$(z-t)^2 = z^2 - 2tz + t^2$

$\therefore (z-t)^2 - t^2 = z^2 - 2tz$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left[(z-t)^2 - t^2\right]} \, dz$$

$$= e^{1/2\, t^2} \left[\text{times } \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} \, dz \right]$$

which is 1 since it's the integral of a $N(t,1)$ PDF.

i.e. $M_X(t) = E[e^{tX}] = E[e^{t(\mu+\sigma Z)}]$

$$= e^{t\mu} E[e^{t\sigma Z}]$$

$$= e^{t\mu} M_Z(t\sigma)$$

$$= e^{t\mu} e^{1/2\, t^2 \sigma^2}$$

$$= e^{t\mu + \frac{1}{2}t^2\sigma^2}$$

Brilliant! Let's extend to the multivaiate case

(w/o proof because that adds no extra intuition)

ELEMENTARY MULTIVARIATE RESULTS USED IN PS 1:

(A) Remember in EC400, we did $E[\cdot]$ and $Var[\cdot]$ rules?
eg. $E[aX] = aE[X]$, and
$Var[aX] = a^2 Var[X]$, for $a \in \mathbb{R}$ and X a scalar RV.

(B) Well, vectors work the same way:
eg: let X be an $n \times 1$ vector-valued RV. Let $a \in \mathbb{R}^n$.

$$\text{Then, } E[a'X] = a' E[X], \text{ and}$$
$$Var[a'X] = a' Var[X] a$$

Note 1: $\Sigma := Var[X] = E[(X - E[X])(X - E[X])']$, an $n \times n$ matrix s.t. $\lambda_n\{\Sigma\} > 0$ and $\Sigma = \Sigma'$.

Note 2: $X \sim N(\mu, \Sigma) \iff f_X(x) = (2\pi)^{-n/2} \det\{\Sigma\}^{-1/2} \exp\left[-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right]$ for any $x \in \mathbb{R}^n$.

If additionally $\Sigma = \sigma^2 I_n$ and say $\mu_i = \mu_j = \underline{\mu}$ for $i,j = 1, \ldots, n$ then also
$X_i \overset{IID}{\sim} N(\underline{\mu}, \sigma^2)$ for $i = 1, \ldots, n$.

Step 2: Say $X_{n \times 1} \sim N(\mu_{n \times 1}, \Sigma_{n \times n})$.

Analagous to the univaiate case, we have that for some $t \in \mathbb{R}^n$,

$$M_X(t) = \exp\left[t'\mu + \frac{1}{2} t' \Sigma t \right].$$

Step 3: Define $Y := a'X$ for some $a \in \mathbb{R}^n$

(usual stuff from Q1, PS1 – i.e. we have X and $F_X(x)$!! We also have a transformation Y. Find $F_Y(y)$.)

let's see how...

$$M_Y(t) \overset{\text{Def}^n.}{=} E\left[e^{t'Y}\right] \overset{\text{Def}^n.}{=} E\left[e^{t'a'X}\right] = E\left[e^{(at)'X}\right]$$

$$\overset{\text{Def}^n.}{=} M_X(at) \overset{\text{Step 2}}{=} \exp\left[t'a'\mu + \frac{1}{2}t'a'\Sigma at\right]$$

Job done. 3 proofs in one shot! Why? Because...

By the uniqueness property of MGFs, we recognise that

$$Y \sim N\left(a'\mu, a'\Sigma a\right).$$

(5) **DEFINITION.** An $n \times n$ (square) matrix A is said to be non-singular when $Ax = 0$ if and only if $x = 0$.

PROPOSITION 5.a. If A is positive definite, then A is non-singular.

PROOF.

- let us recall that a positive definite $n \times n$ matrix A is s.t.

$$x'Ax \begin{cases} = 0 & , \text{if } x = 0 \\ > 0 & , \text{o/w} \end{cases} , \quad \text{for any } x \in \mathbb{R}^n.$$

- Looking at the above definition,

  (i) given A is positive definite, for any $x \neq 0$, $x'Ax > 0$; but then, clearly, for any $x \neq 0$, $Ax \neq 0$.

  (ii) Moreover, for $x = 0$, it is clear that $Ax = 0$.

It follows that A is non-singular. This completes the proof.

Vassilis' solution is as above but here's an alternative very useful way to think about it:

- A is P.D. $\Rightarrow \lambda_n\{A\} > 0$, where $\lambda_i\{\cdot\}$ denotes the i^{th} largest eigenvalue for $i = 1, \ldots, n$.
- But clearly then,

$$\det\{A\} = \prod_{i=1}^{n} \lambda_i\{A\} > 0.$$ The result follows.

PROPOSITION 5b. If X is an $n \times k$ matrix $(n > k)$ s.t. $\operatorname{rank}(X) = k$, then $(X'X)$ is positive definite and non-singular.

PROOF.

- Let us recall that if $\operatorname{rank}(X) = k$, we have that for any $y \in \mathbb{R}^k$,

$$Xy \begin{cases} = 0, & \text{if } y = 0 \\ \neq 0, & \text{if } y \neq 0. \end{cases}$$

it's just a "matrix way" of giving the lin. indep. condition among columns of X.

- Then, for any $y \in \mathbb{R}^k$ s.t. $y \neq 0$,

$$y'(X'X)y = (Xy)'(Xy) = \|Xy\|_2^2 > 0.$$

- Moreover, for $y = 0$, it is trivial that $Xy \overset{!}{=} 0$.

[Above, the notation "$\|\cdot\|_2$" denotes the L_2-norm]

· Given (i) the definition of positive definiteness as recapitulated in the proof of Proposition 5.a ; and subsequently,

(ii) Proposition 5.a itself,

the required results follow.

This completes the proof.

(Q6) See official solutions. I don't have much to add.

(Q7) PROPOSITION 7a. For $n \times n$ non-singular matrices A and B, it holds that:

$$(AB)^{-1} = B^{-1}A^{-1}$$

PROOF.

- Since A and B are square and non-singular, there exist matrices A^{-1} and B^{-1} s.t.

$$AA^{-1} = A^{-1}A = BB^{-1} = B^{-1}B = I_n.$$

- Next, by non-singularity of A, we have that for any $v \in \mathbb{R}^n$,

$$Av = 0 \text{ if and only if } v = 0.$$

- It must then be true that the previous statement holds for choice of $v = Bx$ for any $x \in \mathbb{R}^n$. In other words,

$$A(Bx) = 0 \text{ if and only if } Bx = 0.$$

- But by non-singularity of B, we know that

$$Bx = 0 \text{ if and only if } x = 0.$$

- We thus have that $(AB)x = 0$ if and only if $x = 0$, proving that square matrix AB is non-singular so that there exists a matrix $(AB)^{-1}$ s.t.

$$(AB)(AB)^{-1} = (AB)^{-1}(AB) = I_n.$$

- Given $(AB)(AB)^{-1} = I_n$, we have that

$$B^{-1}A^{-1}(AB)(AB)^{-1} = B^{-1}A^{-1}$$

$$\therefore (AB)^{-1} = B^{-1}A^{-1}.$$

This completes the proof.

PROPOSITION 7b. For $n \times n$ non-singular matrix A, it holds that

$$(A')^{-1} = (A^{-1})'.$$

PROOF. · Given square non-singular matrix A, we know that there exists a matrix A^{-1} s.t.

$$AA^{-1} = A^{-1}A = I_n.$$

· Then, $(AA^{-1})' = (A^{-1})' A' = I_n' = I_n$.

· Further, we know that since A^{-1} exists, $(A')^{-1}$ also exists [because $\det\{A\} = \det\{A'\} \neq 0$]

· Post multiplying $(A^{-1})' A' = I_n$ on both sides by $(A')^{-1}$, the required result follows.

· This completes the proof.

(let's also add "real" to keep things easier)

(8) PROPOSITION 8.a. For real P.D. symmetric matrix A, there exists a square N.S. matrix P s.t. $A = PP'$.

PROOF OF 8.a.

(i) Lemma 8.a.1. Since A is a real symmetric matrix, there exists a representation

$$A = V\Lambda V'$$

where $\Lambda := \text{diag}\{\lambda_1\{A\}, \ldots, \lambda_n\{A\}\}$ is a diagonal matrix of ordered eigenvalues of A, and

where $V := [v_1, \ldots, v_n]$ is an orthogonal matrix of corresponding eigenvectors.

Note: $$v_i' v_j = \begin{cases} 0, & \text{for } i \neq j \\ 1, & \text{o/w} \end{cases}$$ for $i, j = 1, \ldots, n$.

this is called "orthonormality" of vectors

Proof of 8.a.1. The stated result is exactly the spectral theorem for Hermitian matrices.

(ii) Lemma 8.a.2. $\lambda_i\{A\} > 0$, for $i = 1, \ldots, n$.

Proof of 8.a.2. The stated result is just the definition of positive definiteness.

I've defined this in Q5 already.

(iii) Lemma 8.a.3 $\lambda_i\{A\}$ for $i=1,\dots,n$ are all real.

Proof of 8.a.3

- By definition, $Av_i = \lambda_i\{A\}v_i$, where v_i denotes the i^{th} eigenvector of A for $i=1,\dots,n$.
- Then, $v_i^* A' = \overline{\lambda_i\{A\}}v_i^*$, for $i=1,\dots,n$, where $\bar{M}$ denotes the complex conjugate of some matrix M and $M^* := \bar{M}'$.
- Since $A' = A$, $v_i^* A = \overline{\lambda_i\{A\}}v_i^*$ for $i=1,\dots,n$.
- If we post-multiply both sides by v_i,

we have $v_i^* A v_i = \bar{\lambda}_i\{A\} v_i^* v_i$, or

$$\lambda_i\{A\} v_i^* v_i = \bar{\lambda}_i\{A\} v_i^* v_i,$$

for $i = 1, \dots, n$.

- This ensures that $\lambda_i\{A\} = \bar{\lambda}_i\{A\}$ for $i = 1, \dots, n$, which completes the proof of Lemma 8.a.3.

(iv) Finally, given Lemmas 8.a.1, 8.a.2, and 8.a.3, there clearly exists a diagonal matrix

$$\Lambda^{1/2} := \mathrm{diag}\left\{\sqrt{\lambda_1\{A\}}, \sqrt{\lambda_2\{A\}}, \dots, \sqrt{\lambda_n\{A\}}\right\} \quad \text{s.t.}$$

$$A = V\Lambda V' = V\Lambda^{1/2}\Lambda^{1/2}V'.$$

Defining $P = V\Lambda^{1/2}$, the result follows, and this completes (a sketch of) the proof for Proposition 8.a.

PROPOSITION 8.b. For $n \times 1$ random vector

$$X \sim N(0, \Sigma) \text{ s.t. } \operatorname{rank}\{X\} = n,$$

and $\lambda_1(\Sigma) < \infty$;

we have that $Z := X' \Sigma^{-1} X \sim \chi^2_n$.

PROOF

- Σ has to be a P.D. Symmetric matrix because... [you tell me]
- So by Lemma 8.a.1., there exist matrices P and D s.t.

  $\Sigma = PDP'$ where $D = \operatorname{diag}\{\lambda_1\{\Sigma\}, \dots, \lambda_n\{\Sigma\}\}$ and

  P is the orthogonal matrix of corresponding eigenvectors of A.
- Clearly, (i) Σ^{-1} exists, by Proposition 5.a.;

  (ii) Σ^{-1} is symmetric, since

  $$\Sigma^{-1} = (PDP')^{-1} = P'^{-1} D^{-1} P^{-1} = P D^{-1} P',$$

  and $(\Sigma^{-1})' = (PD^{-1}P')' = PD^{-1}P' = \Sigma^{-1}$;

(iii) Σ^{-1} is P.D. since $\lambda_i\{\Sigma^{-1}\} = \dfrac{1}{\lambda_{n+1-i}[\Sigma]}$ for $i=1,\dots,n$.

- It then follows from Proposition 8.a. that there exist matrices $\Sigma^{-1/2} := PD^{-1/2}$ s.t. $\Sigma^{-1} = \Sigma^{-1/2}\Sigma^{-1/2\,\prime}$.
- Consider the vector $Y := \Sigma^{-1/2\,\prime} X$.

  Clearly, $E[Y] = \Sigma^{-1/2\,\prime} E(X) = 0$, and
  $\mathrm{Var}[Y] = \Sigma^{-1/2\,\prime}\,\mathrm{Var}[X]\,\Sigma^{-1/2} = I_n$.

  See slide titled "ELEMENTARY..."
- Combining the above with the distributional result in Q4,

  $Y \sim N(0, I_n)$.
- It also follows that the typical element of Y,
  $Y_i \overset{\text{iid}}{\sim} N(0,1)$ for $i=1,\dots,n$.

- This means that $y'y = \sum_{i=1}^{n} y_i^2 \sim \chi^2_n$

you may need to review basic sampling distributions (N, t, χ^2, F) to understand this.

- The proof is completed by noting that

$$y'y = x' \Sigma^{-1/2'} \Sigma^{-1/2} x$$

$$= x' \Sigma^{-1} x$$

which means that $x' \Sigma^{-1} x \sim \chi^2_n$.

(9) PROPOSITION 9

Suppose M is an $n \times n$ symmetric, idempotent matrix with rank J. Further suppose x is an $n \times 1$ random vector of i.i.d. standard Gaussian random variates, i.e., $x \sim N(0, I_n)$. ~~Show that~~ the quadratic form $z \equiv x'Mx$ is distributed χ^2 with J degrees of freedom.

PROOF

Lemma 9.1 M has eigenvalues either 0 or 1.

Proof of 9.1

- By definition, $Mv_i = \lambda_i\{M\}v_i$ for $i = 1, \dots, n$.
- Pre-multiplying both sides by M,

$$MMv_i = M\lambda_i\{M\}v_i \quad \text{or} \quad Mv_i = \lambda_i^2\{M\}v_i$$

so that $\lambda_i\{M\}v_i = \lambda_i^2\{M\}v_i$ for $i = 1, \dots, n$.

- The result follows.

- Given Lemma 9.1 and Lemma 8.a.1, there exist matrices D, a diagonal matrix of 0's and 1's, ie eigenvalues of M; and C', an orthogonal matrix of corresponding eigenvectors s.t.

$$M = CDC'.$$

- Since $\text{rank}\{M\} = J$, it must be the case that M has J 1's and $(n-J)$ 0's as its eigenvalues.
- Then, $x'Mx = x'CDC'x = (C'x)'D(C'x)$.
- Defining $y := C'x$, we have that $y \sim N(0, I_n)$

$$\text{and } y_i \overset{\text{iid}}{\sim} N(0,1) \text{ for } i=1,\dots,n,$$

where y_i is the i^{th} element of y for $i=1,\dots,n$.

- Thus, $x'Mx = y'Dy = \sum_{i=1}^{n} y_i^2 \lambda_i\{M\} \sim \chi^2_J$, which completes the proof of Proposition 9.