TD1 SADM 2023

18/January/23

Exercise 1

U, V are independent and uniform in $[0,1]$

$X = U + V$ and $Y = U - V$

(a) $Z = \begin{bmatrix} X \\ Y \end{bmatrix}$ $\quad \mathbb{E}[Z] = \begin{bmatrix} \mathbb{E}[X] \\ \mathbb{E}[Y] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[U+V] \\ \mathbb{E}[U-V] \end{bmatrix} = \begin{bmatrix} \mathbb{E}[U] + \mathbb{E}[V] \\ \mathbb{E}[U] - \mathbb{E}[V] \end{bmatrix}$

$$\mathbb{E}[U] = \int_{\mathbb{R}} \overbrace{\mathbb{1}_{[0,1]}(u)}^{p_U(u)}\, u \,.\, du = \int_0^1 u\, du = \left[\frac{u^2}{2}\right]_0^1 = 1/2$$

$\mathbb{E}[V] = \mathbb{E}[U]$ $\quad$ Thus, $\mathbb{E}[Z] = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

independence of U and V

$$\mathrm{cov}(Z) = \begin{bmatrix} \mathrm{Var}(X) & \mathrm{cov}(X,Y) \\ \mathrm{cov}(X,Y) & \mathrm{Var}(Y) \end{bmatrix} \qquad \begin{array}{l} \mathrm{Var}(X) = \mathrm{Var}(U) + \mathrm{Var}(V) \\ \mathrm{Var}(Y) = \mathrm{Var}(U) + \mathrm{Var}(V) \end{array}$$

$$\mathrm{Var}(U) = \int_{\mathbb{R}} (u - 1/2)^2 \cdot \mathbb{1}_{[0,1]}(u)\, du = \cancel{\left[\frac{u^2}{2} - \frac{1}{2}u\right]_0^1}$$

$$= \int_0^1 \left(u^2 - u + 1/4\right) du = \left[\frac{u^3}{3} - \frac{u^2}{2} + \frac{u}{4}\right]_0^1$$

$$= \frac{1}{3} - \frac{1}{2} + \frac{1}{4} = \frac{4}{12} - \frac{6}{12} + \frac{3}{12} = \frac{1}{12}$$

~~Var(U) = Var(V)~~

$\mathrm{Var}(U) = 1/12$

Therefore, $\quad Var(X) = 1/12 + 1/12 = 1/6$

$$Var(Y) = 1/12 + 1/12 = 1/6$$

$$
\begin{aligned}
cov(X,Y) &= \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right] \\
&= \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] \\
&= \mathbb{E}[(U+V)(U-V)] - \cdot \quad 0.1 \\
&= \mathbb{E}[U^2] - \mathbb{E}[V^2] \\
&= \left(Var(U) + (\mathbb{E}[U])^2\right] - \left[Var(V) + (\mathbb{E}[V])^2\right] \\
&= \left(1/12 + 1/4\right) - \left(1/12 + 1/4\right) = 0
\end{aligned}
$$

(b) From the result in (a) we see that $cov(X,Y) = 0$

$$
\begin{aligned} X &= U+V \\ Y &= U-V \end{aligned}
\;\rightsquigarrow\;
\begin{aligned} X &= U+V \\ Y+2V &= U+V \end{aligned}
\;\rightsquigarrow\;
\begin{aligned} Y+2V &= X \\ Y &= X-2V \end{aligned}
$$

We see then that X and Y are not independent...

(A more rigourous way would be to calculate the pdf $f_{XY}(x,y)$ and show that it does not split into product of marginals, ie

$$f_{XY}(x,y) \neq f_X(x) \cdot f_Y(y) \;)$$

Exercise 2

(a) $K_X = \mathbb{E}\left[(X-\mathbb{E}(X))(X-\mathbb{E}(X))^T\right]$

$= \mathbb{E}\left[(X-\mathbb{E}(X))(X^T-\mathbb{E}[X]^T)\right]$

$= \mathbb{E}\left(XX^T - X\,\mathbb{E}[X]^T - \mathbb{E}[X]X^T + \mathbb{E}[X]\mathbb{E}[X]^T\right]$

$= \mathbb{E}[XX^T] - \mathbb{E}(X)\mathbb{E}[X]^T - \mathbb{E}[X]\mathbb{E}[X]^T + \mathbb{E}[X]\mathbb{E}[X]^T$

$= \mathbb{E}(XX^T] - \mathbb{E}(X)\mathbb{E}(X]^T$

(b) $K_{AX} = \mathbb{E}\left[(AX)(AX)^T\right] - \mathbb{E}[AX]\,\mathbb{E}[AX]^T$

$= \mathbb{E}[A XX^T A^T] - A\,\mathbb{E}[X]\,\mathbb{E}[X^T]A^T$

$= A.\left(\mathbb{E}(XX^T) - \mathbb{E}(X)\mathbb{E}(X^T)\right)A^T$

$= A K_X A^T$

(c) We have that $\begin{matrix} X = U+V \\ Y = U-V \end{matrix} \iff \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} U \\ V \end{bmatrix}$

$\mathbb{E}[Z] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\mathrm{cov}(Z) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1/12 & 0 \\ 0 & 1/12 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1/12 & 1/12 \\ 1/12 & -1/12 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

$= \begin{bmatrix} 1/6 & 0 \\ 0 & 1/6 \end{bmatrix}$

Exercise 3

$Z = \begin{bmatrix} X \\ Y \end{bmatrix}$ with $Z \sim \mathcal{N}(\mu, \Sigma)$ where $\mu = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

a)

$$f_Z(z) = f_{XY}(x,y) = \frac{1}{2\pi \cdot \det(\Sigma)} \cdot \exp\left(-\frac{1}{2} \cdot (z-\mu)^T \Sigma^{-1} (z-\mu)\right)$$

$$\det(\Sigma) = 2 - 1 = 1$$

$$\Sigma^{-1} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} (x-1) & (y-2) \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (x-1) \\ (y-2) \end{bmatrix}$$

$$= \begin{bmatrix} 2(x-1) + (y-2) & (x-1) + (y-2) \end{bmatrix} \begin{bmatrix} (x-1) \\ (y-2) \end{bmatrix}$$

$$= 2(x-1)^2 + (x-1)(y-2) + (x-1)(y-2) + (y-2)^2$$

$$= 2x^2 - 4x + 2 + (xy - 2x - y + 2) \times 2 + y^2 - 4y + 4$$

$$= 2x^2 - 8x + 10 + 2xy - 6y + y^2$$

$$f_Z(z) = f_{XY}(x,y) = \frac{1}{2\pi} \exp\left(\frac{-1}{2}\left(2x^2+y^2+2xy-8x-6y+10\right)\right)$$

(b) $$f_{Y|X=x}(y) = \frac{f_{XY}(x,y)}{f_X(x)} \qquad f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2}(x-1)^2\right)$$

$$f_{Y|X=x}(y) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2}\left(y-(3-x)\right)^2\right) = \mathcal{N}\left((3-x),1\right)$$

(c) Remember from CM1 that the best prediction of Y given $X=x$ is the conditional expectation as per:

$$\hat{Y} = \mathbb{E}\left[Y \mid X=x\right] = 3-x$$

Exercise 4

$\begin{bmatrix} X \\ Y \end{bmatrix}$ is a Gaussian vector and $Z = Y - \mathbb{E}[Y] - \frac{cov(X,Y)}{Var(X)} \cdot (X - \mathbb{E}[X])$

(a) $$\mathbb{E}[Z] = \mathbb{E}[Y] - \mathbb{E}[Y] - \frac{cov(X,Y)}{var(X)} \cdot \left(\mathbb{E}[X] - E(X)\right) = 0$$

$$Var(Z) = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 = \mathbb{E}[Z^2]$$

$$= \mathbb{E}\left[\left((Y-\mathbb{E}[Y]) - \frac{cov(X,Y)}{Var(X)} \cdot (X - \mathbb{E}[X])\right)^2\right]$$

$$\mathrm{Var}(Z) = \mathbb{E}\left[(Y-\mathbb{E}[Y])^2\right] + \frac{\mathrm{cov}(X,Y)^2}{\mathrm{Var}(X)^2} \cdot \mathbb{E}\left[(X-\mathbb{E}[X])^2\right]$$

$$- \frac{2\,\mathrm{cov}(X,Y)}{\mathrm{Var}(X)} \cdot \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right]$$

$$= \mathrm{Var}(Y) + \frac{\mathrm{cov}(X,Y)^2}{\mathrm{Var}(X)} - \frac{2\,\mathrm{cov}(X,Y)^2}{\mathrm{Var}(X)}$$

$$= \mathrm{Var}(Y) - \frac{\mathrm{cov}(X,Y)^2}{\mathrm{Var}(X)}$$

(b) $\begin{bmatrix} X \\ Y \end{bmatrix}$ is a Gaussian vector so $\begin{bmatrix} X \\ Z \end{bmatrix}$ should also be one since

$$Z = Y - \frac{\mathrm{cov}(X,Y)}{\mathrm{Var}(X)} \cdot X + \frac{\mathrm{cov}(X,Y)}{\mathrm{Var}(X)} \cdot \mathbb{E}[X] - \mathbb{E}[Y]$$

$$\underbrace{\begin{bmatrix} 1 & 0 \\ -\frac{\mathrm{cov}(X,Y)}{\mathrm{Var}(X)} & 1 \end{bmatrix}}_{M} \begin{bmatrix} X \\ Y \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ \frac{\mathrm{cov}(X,Y)}{\mathrm{Var}(X)} \cdot \mathbb{E}[X] - \mathbb{E}[Y] \end{bmatrix}}_{b}$$

$$\begin{bmatrix} X \\ Z \end{bmatrix} = M \cdot \begin{bmatrix} X \\ Y \end{bmatrix} + b$$

And since $\begin{bmatrix} X \\ Z \end{bmatrix}$ is a Gaussian vector we can just check the covariance $\text{cov}(X,Z)$ to test for independence

$$\text{cov}(X,Z) = \mathbb{E}\left[X \cdot \left(Y - \mathbb{E}[Y] - \frac{\text{cov}(X,Y)}{\text{var}(X)} \cdot (X - \mathbb{E}[X])\right)\right] - \mathbb{E}[X] \overset{0}{\cancel{\mathbb{E}[Z]}}$$

$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] - \frac{\text{cov}(X,Y)}{\text{var}(X)} \cdot \left(\mathbb{E}[X^2] - (\mathbb{E}[X])^2 \right)$$

$$= \text{cov}(X,Y) - \text{cov}(X,Y) = 0$$

(c) I don't have much information about $\begin{bmatrix} X \\ Y \end{bmatrix}$ apart from the fact that it is a Gaussian vector. Can this be enough for us to have a expression for the conditional distribution of Y given $X=x$? Yes!

$$\text{Prob}\left\{ Y \leq y \mid X = x \right\} = \text{Prob}\left\{ Z + \mathbb{E}[Y] + \frac{\text{cov}(X,Y)}{\text{Var}(X)} \cdot (x - \mathbb{E}[X]) \leq y \;\middle|\; X = x \right\}$$

since Z is independent of X

$$= \text{Prob}\left\{ Z + \mathbb{E}[Y] + \frac{\text{cov}(X,Y)}{\text{Var}(X)} \cdot (x - \mathbb{E}[X]) \leq y \right\}$$

Therefore,

$$f_{Y|X=x}(y) = \mathcal{N}\left(\mathbb{E}[Y] + \frac{\text{cov}(X,Y)}{\text{Var}(X)} \cdot (x - \mathbb{E}[X]),\ \text{Var}(Y) - \frac{\text{cov}(X,Y)^2}{\text{Var}(X)} \right)$$

$$= \mathcal{N}(\mu, \sigma^2)$$

(d) In exercise 3 we had $\mathbb{E}[X] = 1$, $cov(X,Y) = -1$

$\mathbb{E}[Y] = 2$, $var(Y) = 2$, $var(X) = 1$

so $$f_{Y|X=x}(y) = \mathcal{N}\left(2 - \frac{1}{1}\cdot(x-1),\ 2 - \frac{1}{1}\right)$$

$$= \mathcal{N}(3-x, 1)$$

Exercise 5

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon, \quad \text{with } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

(a) From the CM we have the expressions

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \qquad \text{with} \quad c_{XY} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{Y})$$

$$\hat{\beta}_1 = \frac{c_{XY}}{s_X^2} \qquad s_X^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

why do we need this? explain!

$$\mathbb{E}[\hat{\beta}_1] = \mathbb{E}_{x_1,\dots,x_N}\left[\mathbb{E}[\hat{\beta}_1 \mid X_1 = x_1, \dots, X_n = x_n]\right]$$

$$\mathbb{E}[\hat{\beta}_1 \mid X_1 \dots X_n] = \frac{1}{s_X^2}\cdot\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})\cdot(\mathbb{E}[Y_i] - \bar{Y})$$

$$\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})\cdot(\beta_0+\beta_1 x_i) = \frac{1}{N}\sum_{i=1}^{N} x_i\beta_0+\beta_1 x_i^2-\bar{x}\beta_0-\beta_1\bar{x}x_i$$

$$= \cancel{\bar{x}\beta_0} + \frac{1}{N}\sum_i x_i^2\cdot\beta_1 - \cancel{\bar{x}\beta_0} - \beta_1\cdot(\bar{x})^2$$

$$= \beta_1\cdot s_x^2$$

$$\mathbb{E}[\hat{\beta_1} \mid X_1=x_1,\dots,X_n=x_n] = \frac{1}{s_x^2}\cdot\beta_1 s_x^2 = \beta_1$$

- $\mathbb{E}[\hat{\beta_1}] = \mathbb{E}_X[\hat{\beta_1}\mid X] = \beta_1$

$$\mathbb{E}[\hat{\beta_0}] = \mathbb{E}(\bar{Y}) - \mathbb{E}[\hat{\beta_1}\cdot\bar{X}]$$

$$= \mathbb{E}\left[\frac{1}{N}\sum_i Y_i\right] - \mathbb{E}_X\left[\mathbb{E}[\hat{\beta_1}\bar{x}\mid X]\right]$$

$$Y_i = \beta_0+\beta_1 x_i+\varepsilon_i$$

$$\mathbb{E}[Y_i] = \beta_0+\beta_1\underbrace{\mathbb{E}(x_i)}_{\text{constant here!}} \qquad \frac{1}{N}\mathbb{E}\left[\sum_i Y_i\right] = \beta_0+\beta_1\bar{X}$$

$$\mathbb{E}[\hat{\beta_1}\cdot\bar{x}\mid X] = \bar{X}\beta_1 = \mathbb{E}(\hat{\beta_1}\bar{x})$$

- $\mathbb{E}[\hat{\beta_0}] = \beta_0+\beta_1\bar{X}-\beta_1\bar{X} = \beta_0$

(b) $$\hat{\beta}_1 = \frac{C_{XY}}{s_X^2} = \left(\frac{1}{N}\sum_i x_i y_i - \bar{x}\bar{y}\right) \cdot \frac{1}{s_X^2}$$

Note that $Y_i = \beta_0 + \beta_1 X_i + \varepsilon$ therefore
$\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{\varepsilon}$

$$\frac{1}{N}\sum x_i y_i = \frac{1}{N}\sum_i X_i(\beta_0 + \beta_1 X_i + \varepsilon_i)$$

$$= \beta_0 \bar{X} + \beta_1 \cdot \frac{1}{N}\sum_i X_i^2 + \frac{1}{N}\sum_i X_i \varepsilon_i$$

$$\bar{X}\bar{Y} = \beta_0 \bar{X} + \beta_1 \bar{X}\bar{X} + \bar{X}\bar{\varepsilon}$$

$$\hat{\beta}_1 = \left(\beta_1 s_X^2 + \frac{1}{N}\sum X_i \varepsilon_i - \bar{X}\bar{\varepsilon}\right) \cdot \frac{1}{s_X^2}$$

$$\hat{\beta}_1 = \beta_1 + \frac{1}{s_X^2} \cdot \frac{1}{N}\sum_i (X_i - \bar{X})\varepsilon_i$$ ← very insightful way of viewing $\hat{\beta}_1$!

- $$Var(\hat{\beta}_1) = E_X\left[Var_X(\hat{\beta}_1 | X)\right] = \frac{1}{s_X^4} \cdot \frac{1}{N^2} \cdot \sum_i (X_i - \bar{X})^2 \cdot \sigma^2$$

$$= \frac{1}{N} \cdot s_X^2 \cdot \frac{1}{s_X^4} \cdot \sigma^2 = \frac{1}{N} \cdot \frac{\sigma^2}{s_X^2}$$

6

What about $Var(\hat{\beta}_0)$?

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \frac{1}{N}\sum_i Y_i - \frac{1}{s_X^2} \cdot \left[\frac{1}{N}\sum_i x_i Y_i - \bar{X}\bar{Y}\right] \cdot \left[\frac{1}{N}\sum_i x_i\right]$$

$$\Leftrightarrow \bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{\varepsilon}$$

$$\hat{\beta}_0 = \left(\beta_0 + \cancel{\beta_1 \bar{X}} + \bar{\varepsilon}\right) - \left(\cancel{\beta_1} + \frac{1}{s_X^2} \cdot \frac{1}{N}\sum_i (x_i - \bar{x})\varepsilon_i\right)\bar{x}$$

$$= \beta_0 + \bar{\varepsilon} - \frac{1}{N}\sum_i (x_i - \bar{x})\bar{x} \cdot \varepsilon_i \cdot \frac{1}{s_X^2}$$

$$Var(\hat{\beta}_0) = \frac{\sigma^2}{N} + \frac{1}{N^2 s_X^4} \cdot \sum_i (x_i - \bar{x})^2 \cdot \bar{x}^2 \cdot \sigma^2$$

$$= \frac{\sigma^2}{N} + \frac{1}{N} \cdot \frac{1}{s_X^{\cancel{4}2}} \cdot \bar{x}^2 \cdot \cancel{s_X^2} \cdot \sigma^2$$

$$= \frac{\sigma^2}{N} + \frac{\bar{x}^2}{N} \cdot \frac{\sigma^2}{s_X^2} = \sigma^2 \cdot \left(\frac{1}{N} + \frac{\bar{x}^2}{N s_X^2}\right)$$

$$= \frac{1}{N} \cdot \sigma^2 \cdot \left(1 + \frac{\bar{x}^2}{s_X^2}\right)$$

(c) $$\hat{m}(x) = \hat{\beta}_0 + \hat{\beta}_1 x = \bar{Y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 x = \bar{Y} + \hat{\beta}_1 (x - \bar{x})$$

$$= \frac{1}{N} \sum_i Y_i + (x - \bar{x}) \cdot \left[\beta_1 + \frac{1}{N s_x^2} \cdot \sum_i (x_i - \bar{x}) \varepsilon_i \right]$$

$$= \frac{1}{N} \sum_i (\beta_0 + \beta_1 x_i + \varepsilon_i) + (x - \bar{x}) \cdot \left[\beta_1 + \frac{1}{N s_x^2} \cdot \sum_i (x_i - \bar{x}) \varepsilon_i \right]$$

$$= \beta_0 + \beta_1 \bar{x} + \bar{\varepsilon} + \beta_1 x - \beta_1 \bar{x} + \frac{(x - \bar{x})}{N s_x^2} \cdot \sum_i (x_i - \bar{x}) \varepsilon_i$$

$$= \beta_0 + \beta_1 x + \bar{\varepsilon} + \frac{(x - \bar{x})}{N s_x^2} \cdot \sum_i (x_i - \bar{x}) \varepsilon_i$$

$$\mathbb{E}[\hat{m}(x)] = \beta_0 + \beta_1 x$$

(d) $$\mathrm{Var}(\hat{m}(x)) = \mathrm{Var}(\bar{\varepsilon}) + \frac{(x - \bar{x})^2}{N^2 s_x^4} \cdot \sum_i (x_i - \bar{x})^2 \cdot \mathrm{Var}(\varepsilon_i)$$

$$= \frac{1}{N} \sigma^2 + \frac{(x - \bar{x})^2}{N} \cdot \frac{\sigma^2}{s_x^2}$$

$$= \frac{1}{N} \cdot \sigma^2 \cdot \left(1 + \frac{(x - \bar{x})^2}{s_x^2} \right)$$