# Computational Statistics

## Transformation of r.v.s

$$X \sim f_X(x)$$

$$Z = g(X)$$

Q: What is the density of Z ??

$$F_Z(z) = \text{Prob}(Z \le z)$$
$$= \text{Prob}(g(X) \le z)$$
$$= \text{Prob}(X \le g^{-1}(z))$$
$$= F_X(g^{-1}(z))$$

$$\frac{d}{dz} F_Z(z) = \frac{d}{dz} F_X(g^{-1}(z)) \quad \leftarrow \text{Apply chain rule.}$$

Examples: $g(x) = a + bX$ for $b \neq 0$

<u>Example:</u> Order statistics:

Let $X_1, \dots, X_n \overset{iid}{\sim} f_X(x)$

order statistics: $X_{(1)}, X_{(2)}, \dots, X_{(n)}$

$$X_{(1)} = \min_{1 \le i \le n} X_i \qquad X_{(n)} = \max_{1 \le i \le n} X_i$$

$\forall x \in \mathbb{R}$,

$$F_{X_{(n)}}(x) = \text{Prob}(X_{(n)} \le x) = \text{Prob}(X_1 \le x, X_2 \le x, \dots, X_n \le x)$$

$$= \text{Prob}(X_1 \le x) \cdot \text{Prob}(X_2 \le x) \cdots \text{Prob}(X_n \le x)$$

$$= [F_X(x)]^n$$

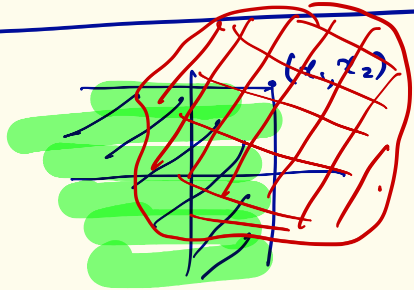
$$f_{X_{(n)}}(x) = \frac{d}{dx} F_{X_{(n)}}(x) = \frac{d}{dx}[F_X(x)]^n = n\,[F_X(x)]^{n-1} \cdot f_X(x)$$

$\forall x \in \mathbb{R}$

$$1 - F_{X_{(1)}}(x) = \text{Prob}(X_{(1)} \geq x) = \text{Prob}(X_1 \geq x, X_2 \geq x, \ldots, X_n \geq x)$$

$$= (1 - F_X(x))^n$$

---




Quick review (Joint densities)

r.variable $(\Omega \to \mathbb{R})$

random vectors $(\Omega \to \mathbb{R}^d)$

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

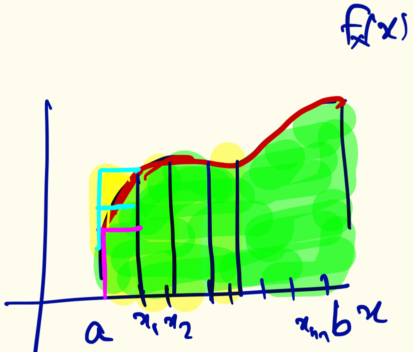
$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \text{Prob}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n)$$

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{x_1^2 + x_2^2}{2}}$$

$-\infty < x_1 < \infty$

$-\infty < x_2 < \infty$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{x_1^2+x_2^2}{2}} \, dx_1 \, dx_2$$




$$\int_a^b f_X(x)\,dx$$

discrete:

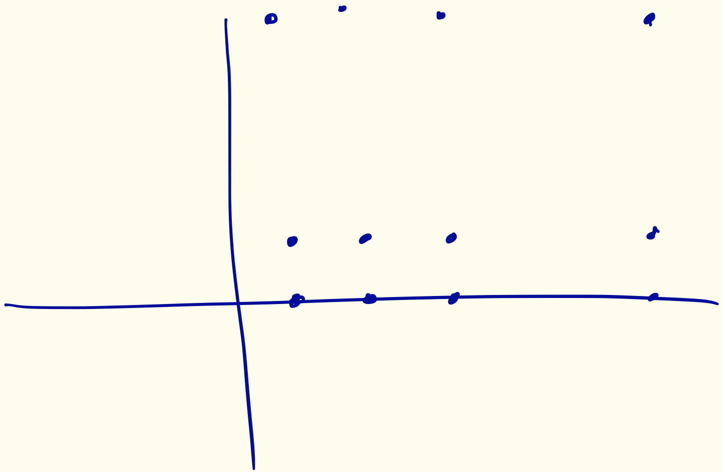
| $X_1$ | $x_1$ | $x_2$ | $x_3$ — |
|---|---|---|---|
| $P(X_1 = x)$ | $p_1$ | $p_2$ | $p_3$ |

$X_1, X_2$ : 2 discrete random variables.

| $X_2$ \ $X_1$ | $x_{11}$ | $x_{12}$ | $x_{13}$ | ... | |
|---|---|---|---|---|---|
| | | | | | ★ |
| $x_{21}$ | ★ | ★ | | | ★ |
| $x_{22}$ | | ★ | ★ | | ★ |
| | | | | | |

$X_1, X_2$ $\quad X_1 \in \{1, \ldots, 10\}$

$X_2 = \{1, \ldots, 10\}$



$$f_{X_1}(x_1) = \int_{x_2} f_{X_1, X_2}(x_1, x_2)\, dx_2$$

$\downarrow$

marginal density of $X_1$

Conditional density:

$$f_{X_2 | X_1 = x_1} = \frac{f(x_1, x_2)}{f_{X_1}(x_1)} \quad \forall\ x_2$$

Independence of r.v.s.

$X_1$ & $X_2$ are r.v-s with the joint density

$$f_{X_1, X_2}(x_1, x_2)$$

Let $f_{X_1}(x_1)$ & $f_{X_2}(x_2)$ be the marginal densities of $X_1$ and $X_2$ resp.

Then $X_1$ & $X_2$ are called as independent r.v.s.
if

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1)\, f_{X_2}(x_2)$$

$$E(h(X_1, \cdots, X_n)) = \int\int \cdots \int h(x_1, \dots, x_n)\, f_{X_1, \dots, X_n}(x_1, \dots, x_n)\, dx_1 \dots dx_n$$

For any random vector $X$

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}$$

$$\mu_i = E(X_i)$$

$$E(X) = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix} = \mu$$

$\Sigma$ = covariance matrix

$$\Sigma = E[X-\mu][X-\mu]^T$$

$$\Sigma_{ij} = E(X_i - \mu_i)(X_j - \mu_j)$$

---

Linear transformation of random vectors.

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_d \end{pmatrix}$$

$Y = AX$ where $A \in \mathbb{R}^{d \times d}$

$$\mu_x = E(X), \qquad \mu_y = E(Y)$$

$$\mu_y = E(Y) = E(AX) = A\,E(X) = A\mu_x$$

$$\Sigma_y = E(Y-\mu_y)(Y-\mu_y)^T, \qquad \Sigma_x = E(X-\mu_x)(X-\mu_x)^T$$

$$= E\left([A(X-\mu_x)][A(X-\mu_x)]^T\right)$$

$$= E\left[A(X-\mu_x)(X-\mu_x)^T A^T\right]$$

$$= A\left[E(X-\mu)(X-\mu)^T\right]A^T$$

$$\Sigma_y = A\,\Sigma_x\,A^T$$

$d=2$

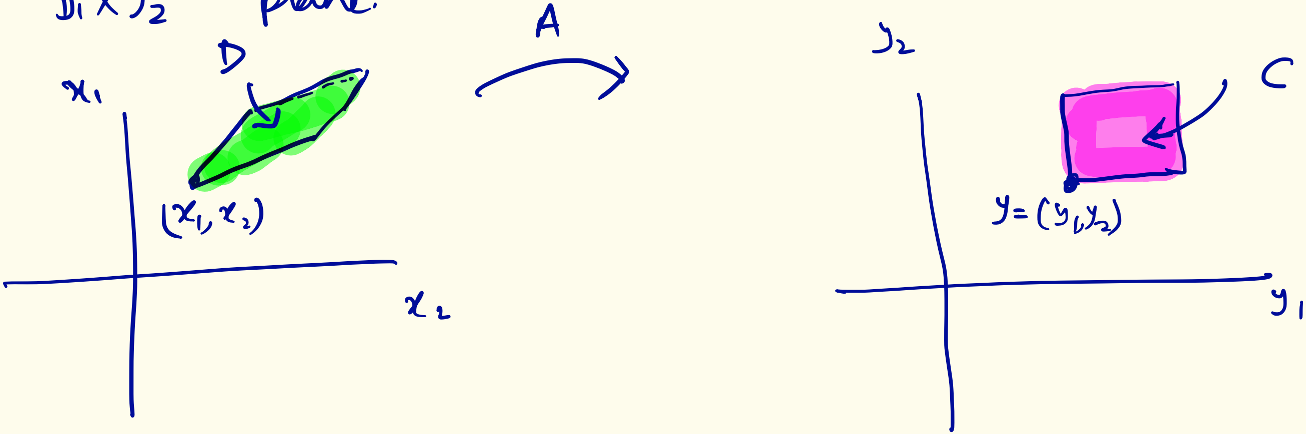
Suppose $A^{-1}$ exists.

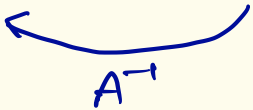
$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$

$Y = AX$

Consider the area $C = [y_1, y_1+h] \times [y_2, y_2+h]$

in $y_1 \times y_2$ plane.




$f_Y(y)\, h^2 \approx \text{Prob}(Y \in C)$

where $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$\text{Prob}(Y \in C) = \text{Prob}(X \in D) \approx h^2 f_Y(y)$$

$$= h^2 |A^{-1}| \, f_X(x)$$

$$= h^2 |A|^{-1} \, f_X(x)$$

$$\text{Prob}(Y \in C) \approx h^2 \frac{f_X(A^{-1}y)}{|A|}$$

taking $h \to 0$

$$f_Y(y) = \frac{f_X(A^{-1}(y))}{|A|} \qquad y \in \mathbb{R}^2$$

$X_1, \ldots, X_n \overset{iid}{\sim} U[0,1]$

Find the densities of $X_{(1)}$ and $X_{(n)}$