10/18/22

Part 2: 2 or more RVs (multivariate/multidimensional)

Joint distribution

discrete RVs

$X = (X_1, X_2)$ — X_1: eye color, X_2: hair color

X_1 eye \ X_2 hair	1	2	3	4
1	$p(1,1)$	$p(1,2)$		
2				
3			$p(3,3)$	
4			$p(4,3)$	

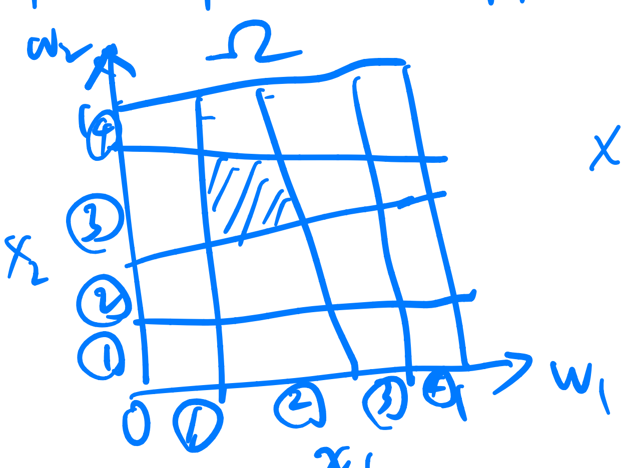
Ω = population

$\omega \sim \text{Unif}(\Omega)$

$X(\omega) = (X_1(\omega), X_2(\omega))$

probability mass function

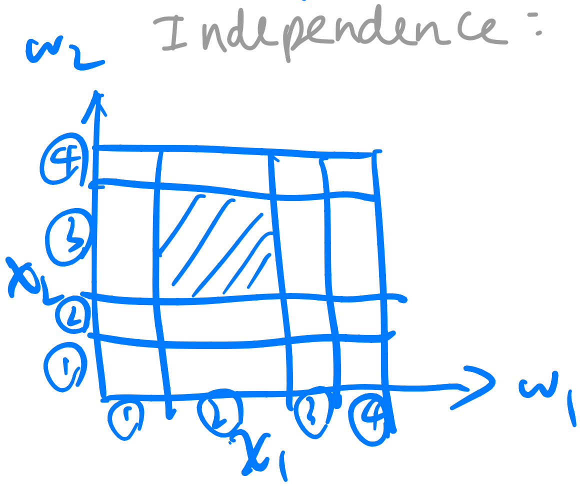
$p(x) = p(x_1, x_2)$ = population proportion



$x(\omega) = (x_1(\omega), x_2(\omega))$

$p(x) = p(x_1, x_2)$ = area of (x_1, x_2)

Independence: $X_1 \perp X_2$



$$p(x_1, x_2) = P(X_1 = x_1 \ \& \ X_2 = x_2)$$
$$= P(X_1 = x_1) \cdot P(X_2 = x_2)$$
$$= p(x_1)\, p(x_2)$$

$$E(h(X)) = \sum_x h(x)\,p(x)$$ * for discrete rvs

$$h(x_1, x_2) \qquad h(x_1,x_2)\,p(x_1,x_2)$$

Joint Distribution

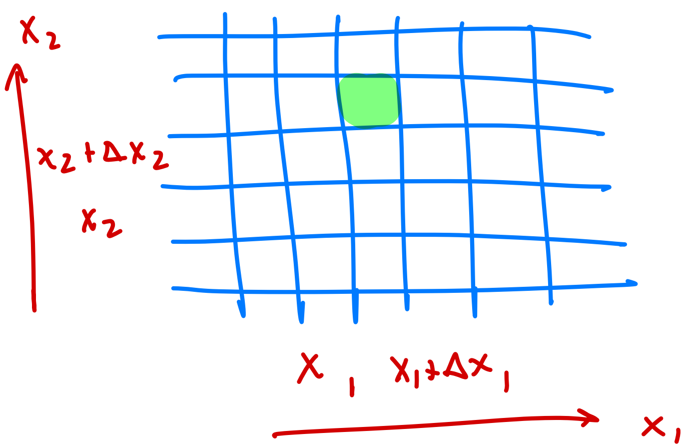
Continuous RVs

Recall univariate case $X \sim f(x)$

$$f(x) = \frac{P(X \in (x, x+\Delta x))}{\Delta x}$$

$$P(X \in (x, x+\Delta x)) = f(x)\,\Delta x$$

multivariate



$$f(x) = f(x_1, x_2) = \frac{P(X_1 \in (x_1, x_1+\Delta x) \;\&\; X_2 \in (x_2, x_2+\Delta x_2))}{\Delta x_1 \Delta x_2}$$

$$= \frac{P(X \in \overbrace{(x_1, x_1+\Delta x_2) \times (x_2, x_2+\Delta x_2)}^{\text{cell }(x_1, x_2)})}{\Delta x_1\, \Delta x_2}$$

$$P(X \in \text{cell}(x_1, x_2)) = f(x)\,\Delta x_1 \Delta x_2$$

$X_1 \perp X_2$

$f(x_1, x_2) = f(x_1) f(x_2)$

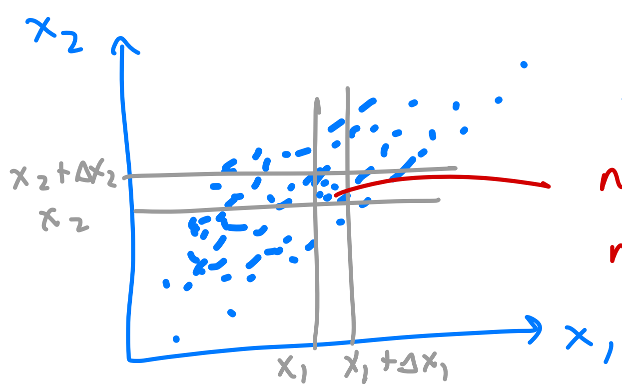
$f(x_1, x_2)\,\Delta x_1 \Delta x_2 = f(x_1)\Delta x_1 \quad f(x_2)\Delta x_2$

$$E(h(X)) = \int h(x) f(x)\,dx$$

double integral

$$= \iint h(x_1, x_2) f(x_1, x_2)\, dx_1 dx_2$$

visualize w/ a scatterplot



$n \to \infty$

$\Delta x_1, \Delta x_2 \to 0$

$n(x_1, x_2)$ = # of points in cell (x_1, x_2)

$$\frac{n(x_1, x_2)/n}{\Delta x_1 \Delta x_2} \to f(x_1, x_2)$$

generalize to 3D

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$f(x) = f(x_1, x_2, x_3)$

ex.)

electron cloud

position of electron follows pdf

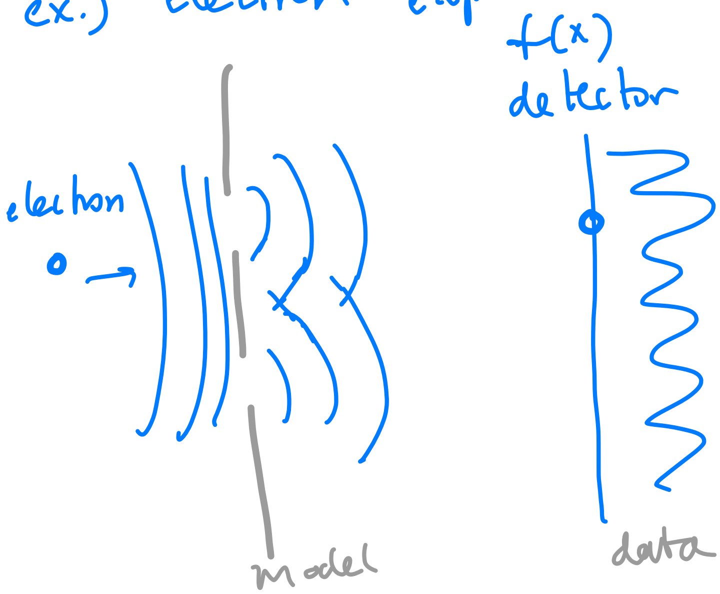
ex.) Quantum physics

$\Psi(x)$ wave function

$$f(x) = |\Psi(x)|^2$$

$$|a + bi|^2 = a^2 + b^2$$

ex.) electron experiment



System $\rightarrow$ data (observer)

$\uparrow f(x)$

model $p(x)$

observe : knowing

subjective

ϕ collapse

ex.) 2 electrons

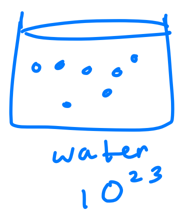
$$X = \begin{bmatrix} x^{(1)} \\ x^{(2)} \end{bmatrix} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \\ x_3^{(1)} \\ x_1^{(2)} \\ x_2^{(2)} \\ x_3^{(2)} \end{bmatrix}$$

$\Psi(x) \rightarrow f(x)$

Ψ (x, environment)

Ψ : not real.

ex.) statistical physics



$X(t)$ evolve deterministically

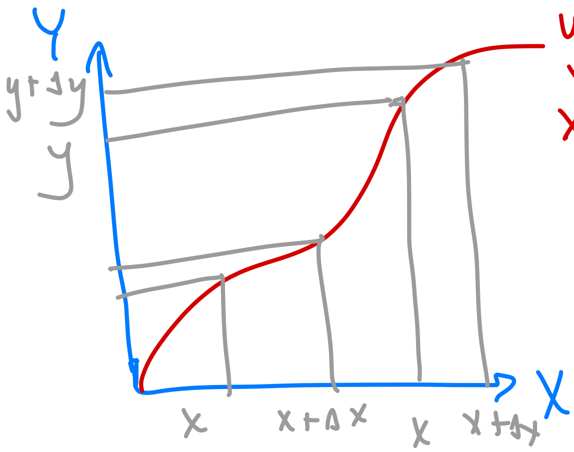
$t \in [0, T]$

$t \sim \text{Unif}[0, T]$

$X(t) \sim p(x)$

Transformation

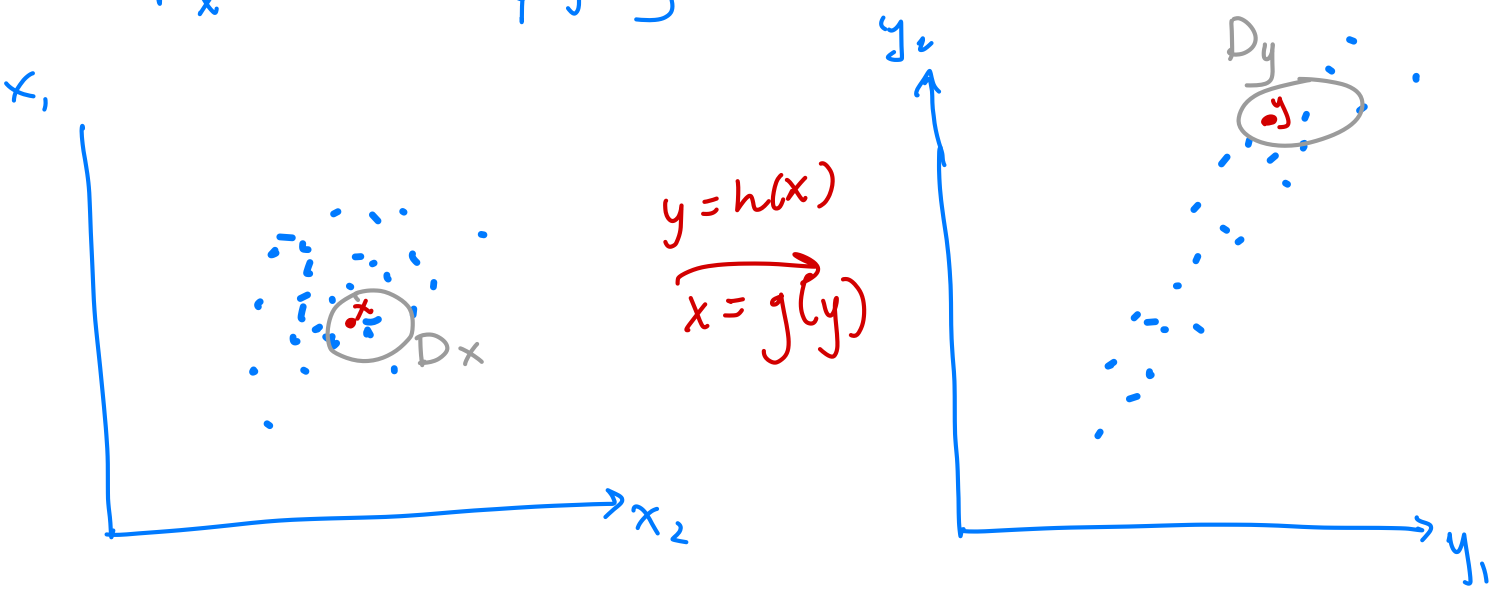
recall 1 D



$y = h(x)$

$x = h^{-1}(y) = g(y)$

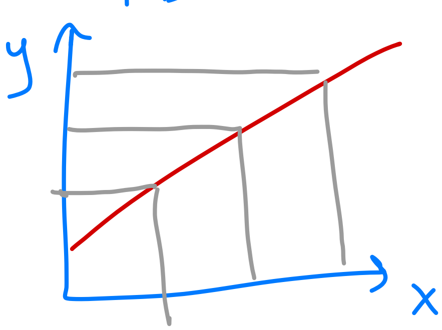
$$f_X(x)\Delta x = f_Y(y)\Delta y$$



$$f_X(x)|\underbrace{D_x}_{\text{size}}| = f_Y(y)|D_y|$$

linear case

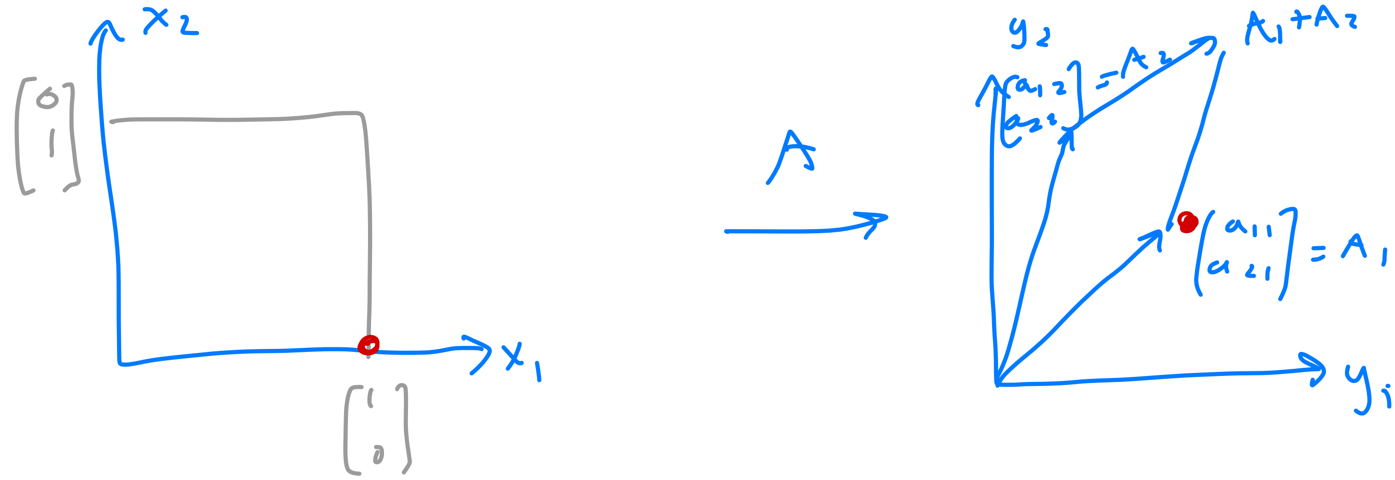
1 D



2D

$$Y = AX$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$



$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{bmatrix} A_1 & A_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = A_1$$

$|\det(A)|$ = area of parallelogram $(A_1 \; A_2)$

for 3D

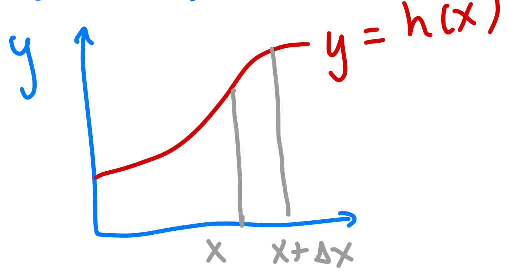
$|\det(A)|$ = volume $(A_1, A_2 A_3)$

$$\frac{|D_y|}{|D_x|} = |\det(A)|$$

<u>non-linear transformation</u>

$y = h(x)$

$$h(x + \Delta x) = h(x) + h'(x)\Delta x + o(\Delta x)$$



Multivariate Calculus

general

$$y_{m\times 1} = h(x)_{n\times 1} \qquad h'(x) = \frac{dy}{dx^T}_{m\times n}$$

$$(\text{e.g. } \underset{m\times 1}{y} = \underset{m\times n}{A}\ \underset{n\times 1}{x})$$

$$= \begin{bmatrix} d\,y_1 \\ \vdots \\ d\,y_2 \\ \vdots \\ d\,y_m \end{bmatrix} \begin{bmatrix} \frac{1}{dx_1} & \cdots & \frac{1}{dx_2} & \cdots & \frac{1}{dx_n} \end{bmatrix}$$

↳ pretend as #'s

$$y_{1\times 1} = h(x) \qquad \text{gradient } h'(x) = \frac{\partial y}{\partial x}_{n\times 1}$$

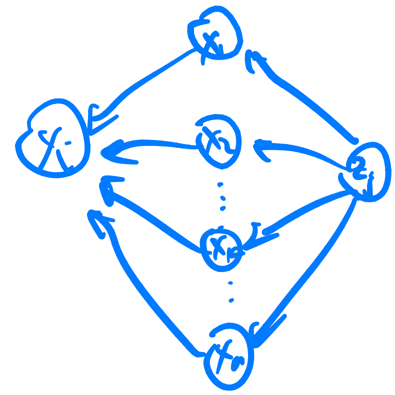
$$y = Ax \qquad y_i = \sum_k a_{ik} x_k$$

$$\frac{dy}{dx^T} = \begin{bmatrix} \frac{dy_i}{dx_j} \end{bmatrix} = \begin{bmatrix} a_{ij} \end{bmatrix} = A$$

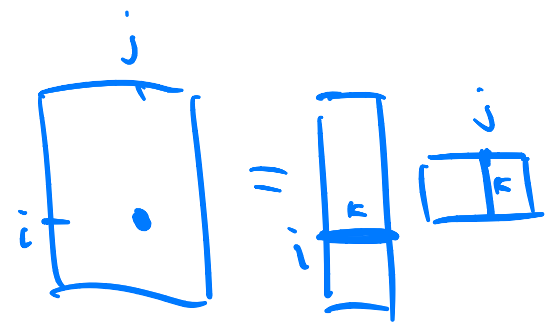
$$\underset{m\times 1}{y} = h(\underset{n\times 1}{x}) \qquad \underset{n\times 1}{x} = g(\underset{l\times 1}{z})$$

$$\frac{dy}{dz^T} = \frac{dy}{dx^T} \cdot \frac{dx}{dz^T}$$

$$\frac{\partial y_i}{\partial z_j} = \sum_k \frac{\partial y_i}{\partial x_k} \frac{\partial x_k}{\partial z_j}$$

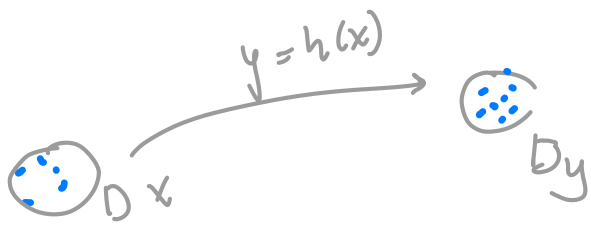


$$\underset{m\times 1}{h(x+\Delta x)} = \underset{m\times 1}{h(x)} + \underset{m\times n}{h'(x)}\,\Delta x$$



Back to density $m = n$

$$\frac{|D_y|}{|D_x|} = |\det(J)|$$



$$J = \text{Jacobian} = h'(x) = \frac{dy}{dx^T}$$

Polar method

$$X \sim f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$Y \sim f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

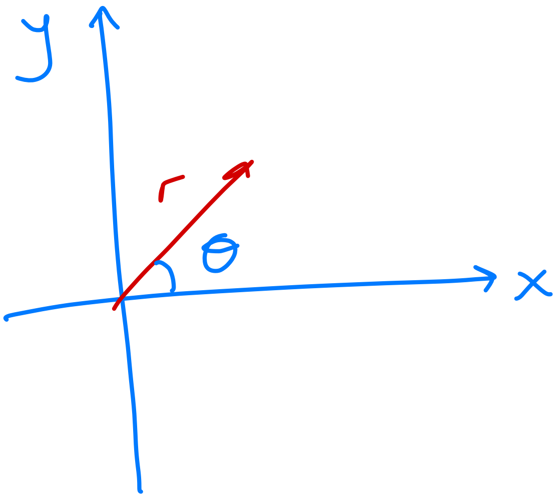
$$X \perp Y$$

$$(X, Y) \sim f(x,y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}$$

$$\begin{cases} x = y\cos\theta \\ y = r\sin\theta \end{cases}$$

$$f(r,\theta)\,\underbrace{\Delta r\,\Delta\theta}_{|D(r,\theta)|} = f(x,y)\,\underbrace{\Delta x\,\Delta y}_{|D(x,y)|}$$

local area



$$\frac{\Delta r\,\Delta\theta}{\Delta x\,\Delta y} = \det(J)$$

$$\frac{|D(x,y)|}{|D(r,\theta)|} = \det(J)$$

$$J = \begin{bmatrix} \frac{dx}{dr} & \frac{dx}{d\theta} \\ \frac{dy}{dr} & \frac{dy}{d\theta} \end{bmatrix}$$

$$t = \frac{r^2}{2}$$

$$\det(J) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r$$

$$\theta \sim \text{Unif}[0, 2\pi]$$

$$r \sim e^{-\frac{r^2}{2}}\,r\,dr$$
$$= e^{-\frac{r^2}{2}}\,d\frac{r^2}{2}$$
$$= e^{-t}\,dt$$

$$f(r,\theta) = f(x,y)\,r$$
$$= \frac{1}{2\pi}\,e^{-\frac{x^2+y^2}{2}}\,r$$
$$= \frac{1}{2\pi}\,e^{-\frac{r^2}{2}}\,r\,dr\,d\theta$$

$$T \sim \frac{R^2}{2} \sim \exp(1)$$

* generate T, then θ, then uniform dist.