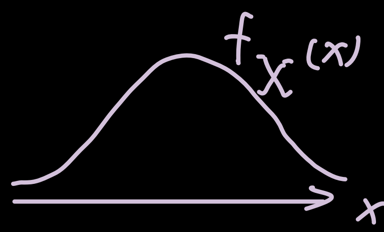
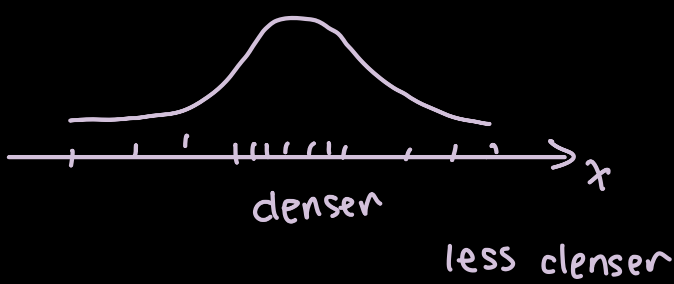
Transformation of continuous random variable

Recall $X \sim \boxed{f_X}(x)$

$\downarrow$ pdf of X



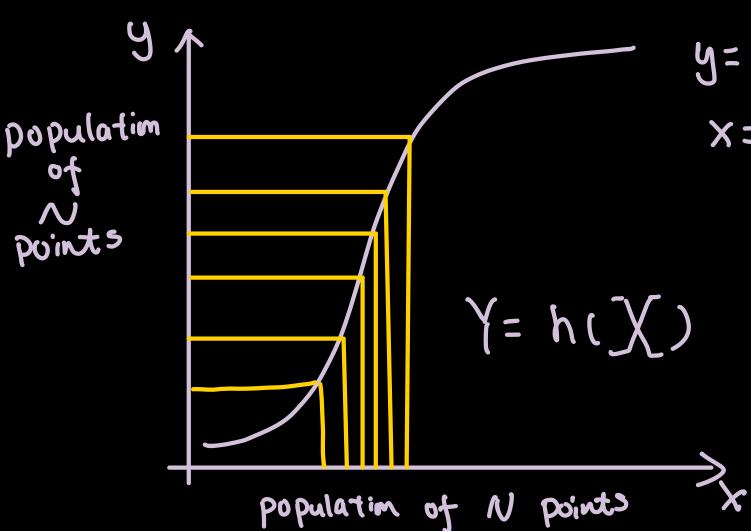
Population of N ($\approx \infty$) points $\xrightarrow[\text{Sampling}]{\text{random}}$ a point X



Number of points in $(x, x+\Delta x)$

$$N(x)/N = Pr(X \in (x, x+\Delta x)) = f(x)\,\Delta x$$

Transformation



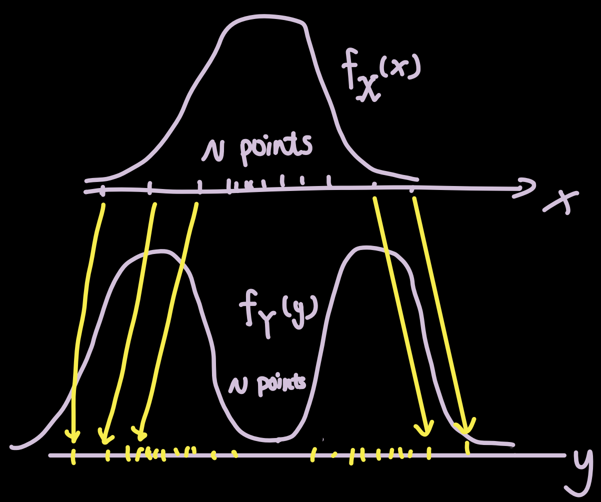
$Y = h(X)$

$y = h(x)$ (invertible)

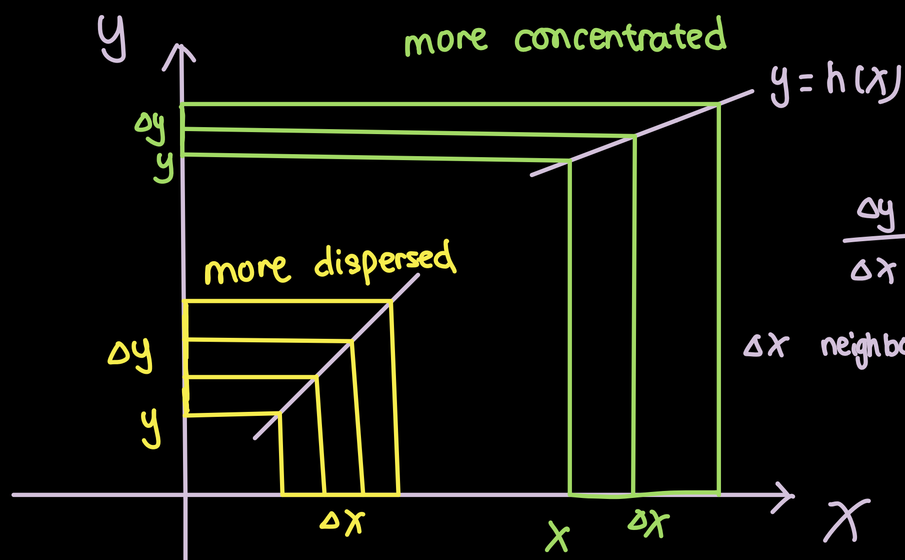
$x = g(y) = h^{-1}(y)$

monotone $\rightarrow$ order preserving mapping

density changes

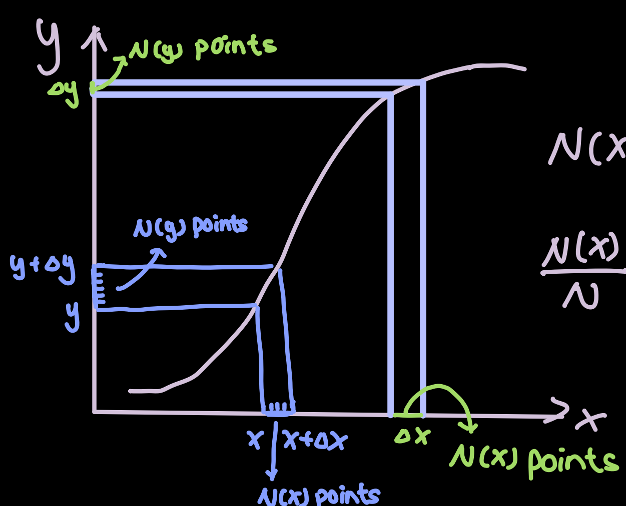


density can change greatly



$$\frac{\Delta y}{\Delta x} = h'(x)$$

Δx neighborhood is squeezed or stretched



$$N(x) = N(y)$$

and

$$\frac{N(x)}{N} = \frac{N(y)}{N}$$

prob of small bin = $\dfrac{\text{\# of points in the bin}}{\text{size}}$

$$\underbrace{P(X \in (x, x+\Delta x))}_{N(x)/N} = \underbrace{P(Y \in (y, y+\Delta y))}_{N(y)/N}$$

Equivalent

$$\Downarrow$$

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

Key equation for change of density under transformation.

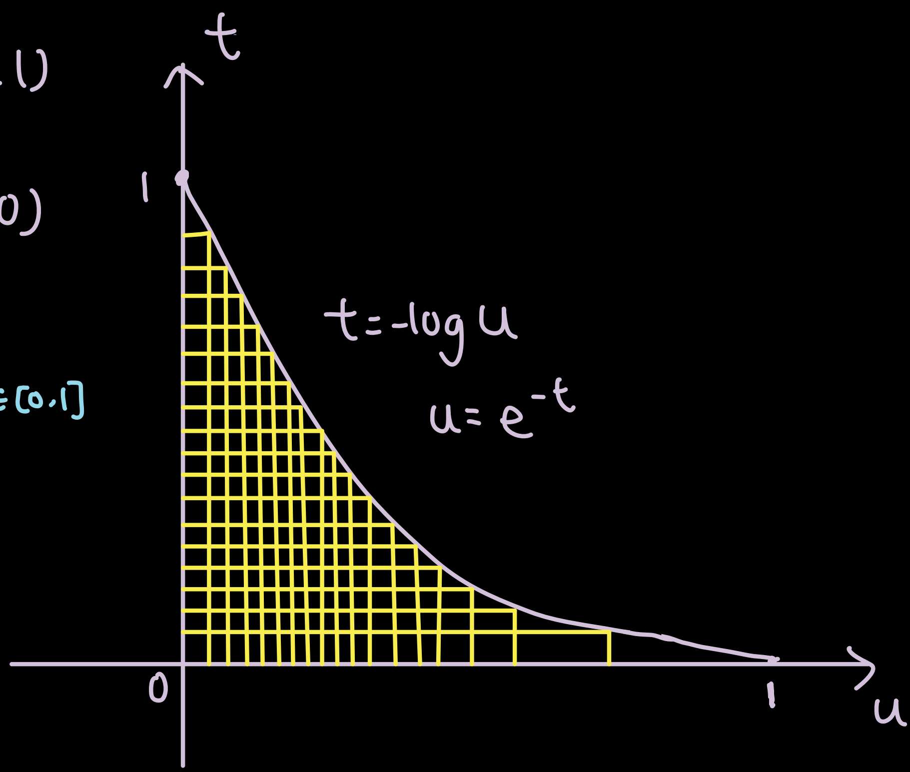
e.g.1 $T \sim \text{Exponential}(1)$

$f(t) = e^{-t} \quad (t \geq 0)$

$T = -\log U$

$f_T(t)\Delta t = \boxed{f_U(u)}\Delta u$ $\quad =1 \;\; u \in [0,1]$

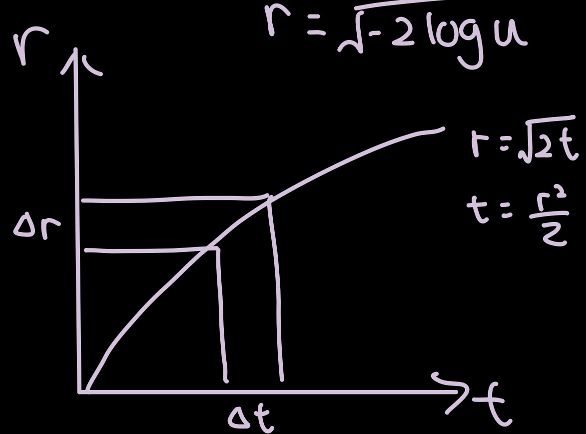
$= \Delta u$



$$f_T(t) = \left|\frac{\Delta u}{\Delta t}\right| = \left|-e^{-t}\right| = e^{-t}$$

e.g.2 $R = \sqrt{2T}$

$r = \sqrt{-2\log u}$



$f_R(1) = f_T(t)\frac{\Delta t}{\Delta r} = e^{-t}r$

$= e^{-\frac{r^2}{2}}r$

$P(U \in (u, u+\Delta t)) = \Delta u$

$\frac{\Delta t}{\Delta r} = \frac{d}{dr}\left(\frac{r^2}{2}\right) = r$

$$f_T(t)\Delta t = f_R(r)\Delta r$$

In HW, $X \sim N(0,1)$

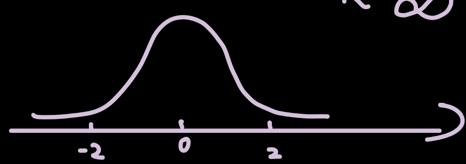
$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

another independent $Y \sim N(0,1)$

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

X and Y are independent : $X \perp Y$

$N(0,1)$ population of N points $\approx \infty$

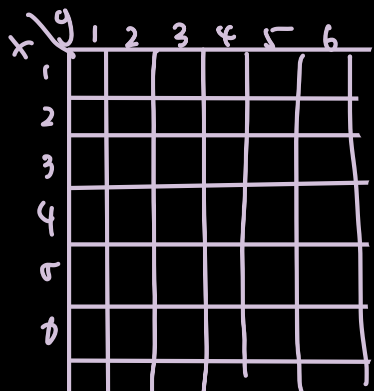


(X, Y) : a 2D random point

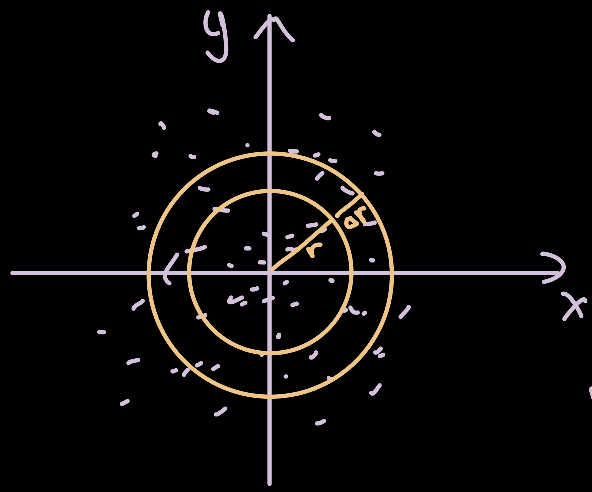
N^2 points of (x_i, y_j)

(recall rolling a die X, then roll again independently $\Rightarrow Y$)

(X, Y)



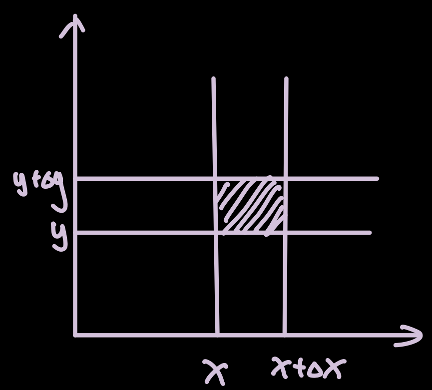
Independence = all N^2 pairs are equally likely



$(x,y) \sim f(x,y)$

$$P(X \in (x, x+\Delta x) \text{ and } Y \in (y, y+\Delta y))$$
$$= f(x,y)\,\Delta x\,\Delta y$$

Indep: $P(X \in (x, x+\Delta x),\ Y \in (y, y+\Delta y))$
$$= P(X \in (x, x+\Delta x))\,P(Y \in (y, y+\Delta y))$$



If $Y \sim N(0,1)$, then

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

density only depends on radius from $(0,0)$ to (x,y) ! has nothing to do with θ.

Polar-coordinate method

$(X, Y) \sim N(1, 0)$ independent

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

$\Theta \sim \text{unif}\,[0, 2\pi]$

$R \sim f(r)$

$$\begin{cases} r = \sqrt{x^2+y^2} \\ \theta = \arctan \frac{y}{x} \end{cases}$$

$$P(R \in (r, r+\Delta r)) = \frac{1}{\cancel{2\pi}} e^{-\frac{r^2}{2}}\, \cancel{2\pi}\, r \cdot \Delta r$$

see illustration in diagram above.

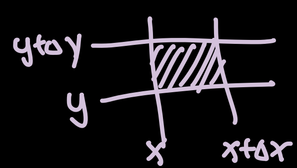
correspond to e.g. 2.

$$\begin{cases} \Theta = 2\pi U_1 \\ R = -\sqrt{2 \log U_2} \end{cases} \qquad \begin{cases} X = R\cos\Theta \\ Y = R\sin\Theta \end{cases}$$

new way to generate random #s in 2D.

Back to Independence.

$P(X \in (x, x+\Delta x) \text{ and } Y \in (y, y+\Delta y))$



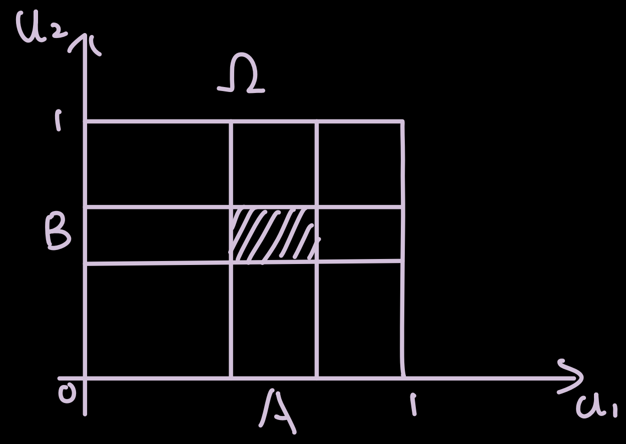
$$= \frac{\#\text{ of points in } (x, x+\Delta x)\cdot(y, y+\Delta y)}{N^2}$$

$$= \frac{\#\text{ of } x_i\text{'s in } (x, x+\Delta x)\cdot \#\text{ of } y_j \text{ in } (y, y+\Delta y)}{N^2}$$

$$= \frac{\#\text{of } x_i \text{ in} (x, x+\Delta x)}{N} \cdot \frac{\#\, y_j \text{ in} (y, y+\Delta y)}{N}$$

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

$(U_1, U_2) \sim \text{Unif}[0,1]$ indep.

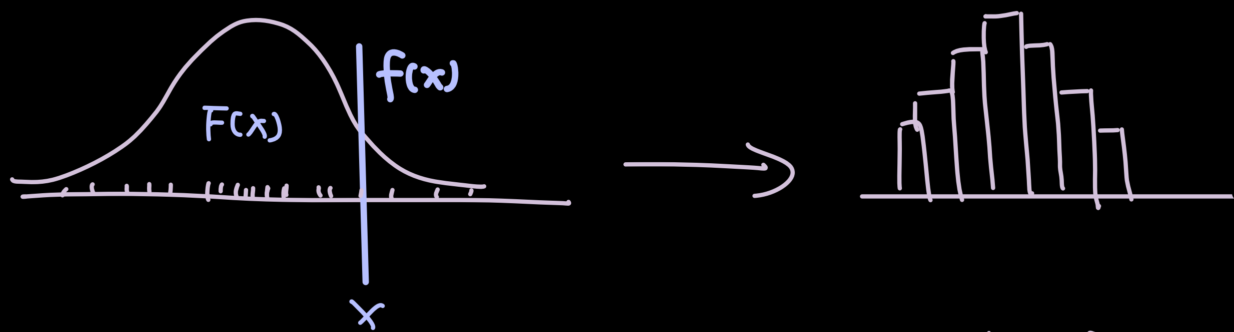


$$P(A \cap B) = \frac{|A \cap B|}{|\Omega|} = |A||B| = P(A)P(B)$$

If $A \perp B$, $P(A \cap B) = P(A)P(B)$

If $X \perp Y$, $f(x,y) = f_X(x) f_Y(y)$

Canonical Example 1



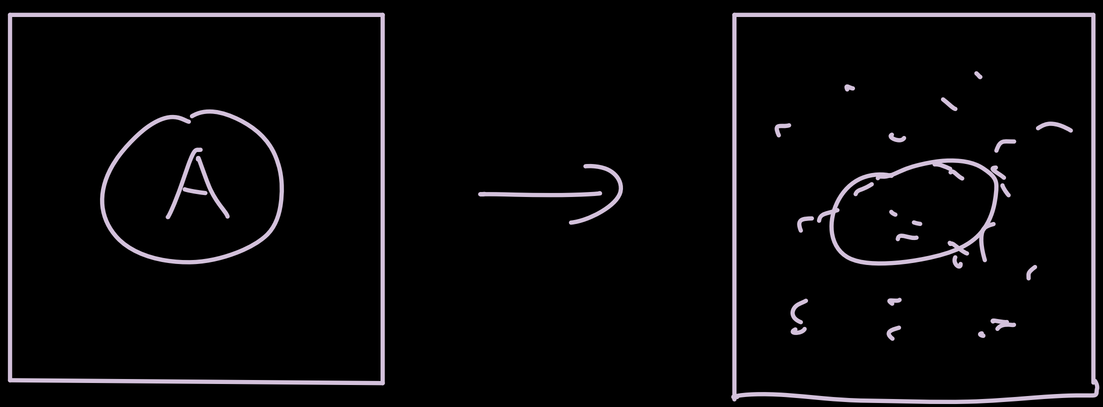
population of $N (\approx \infty)$ points
↓
300 million

deterministic

sample of n points
↓
10,000

random statistic fluctuation

Canonical Example 2



Ω n points

frequency $\frac{n(A)}{n}$

fluctuate around $P(A)$