Two Functions & 2 Random Variables

Given 2 RVs, X & Y, let $Z = g(X,Y)$, $W = h(X,Y)$

where $\left. \begin{array}{l} g: \mathbb{R}^2 \to \mathbb{R} \\ h: \mathbb{R}^2 \to \mathbb{R} \end{array} \right\}$ what is the joint density function of Z & W?

$f_{ZW}(z, \omega)$?

---

* Use the change of variables approach (approach #2)
* If Z & W represent a linear transformation of X & Y
  → treat X & Y as a vector & Z & W are then the result of matrix multiplication

---

Example: Linear Transformation of a vector (X, Y)

$$\boxed{\begin{bmatrix} Z \\ W \end{bmatrix} = A \begin{bmatrix} X \\ Y \end{bmatrix}}, \text{ where } A \text{ is a } 2\times 2 \text{ matrix}$$

If A is a sine/cosine function of some angle θ then this represents a rotation of the vector XY by the angle θ

Could also be a projection

$$\Rightarrow \left. \boxed{\begin{array}{l} g(x,y) = a_{11}x + a_{12}y \\ h(x,y) = a_{21}x + a_{22}y \end{array}} \right\} \ \& \ A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \leftarrow \text{(coefficient matrix)}$$

AFFINE TRANSFORMATION

---

* Could find f_{ZW} by starting w/ CDF: $F_{ZW}(z, \omega) = P(Z \leq z, W \leq \omega)$

$$= P((X,Y) \in D_{z\omega})$$

where $D_{z\omega} = \{ (x,y) \in \mathbb{R}^2 : g(x,y) \leq z, h(x,y) \leq \omega \}$ for every $z, \omega \in \mathbb{R}$

Integrate the joint density fn of X & Y over the set $D_{z\omega}$, to get the joint CDF of f of z & w

However, in practice it is too difficult (to find $D_{z\omega}$) $\Longrightarrow$ Instead, do change of variables approach

Change of Variables Approach (for 2 fn. & 2 RVs)

* assume $\left| \begin{array}{l} z = g(x,y) \\ w = h(x,y) \end{array} \right.$ may be solved simultaneously in order to get:

$$\longrightarrow \left| \begin{array}{l} x = g^{-1}(z,w) \\ y = h^{-1}(z,w) \end{array} \right\} \text{Simultaneous Solution}$$

(also assume that the partials of z,w w.r.t. x,y exists)

* Now $\left| \begin{array}{l} g^{-1} : \mathbb{R}^2 \to \mathbb{R} \\ h^{-1} : \mathbb{R}^2 \to \mathbb{R} \end{array} \right\}$ for the case $\begin{bmatrix} z \\ w \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow$ $\boxed{\text{this means } A^{-1} \text{ exists}}$

---

* **Result**: It may then be shown that ...

$$\boxed{f_{ZW}(z,w) = \frac{f_{XY}\left(g^{-1}(z,w), h^{-1}(z,w)\right)}{\left| \frac{\partial(z,w)}{\partial(x,y)} \right|}}$$

(determinate)

* Now find denominator term, where

$$\left| \frac{\partial(z,w)}{\partial(x,y)} \right| \equiv \left| \begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} \right|$$

(Abs. value of the determinate)

$$= \boxed{\left| \frac{\partial z}{\partial x}\frac{\partial w}{\partial y} - \frac{\partial z}{\partial y}\frac{\partial w}{\partial x} \right|} \Leftrightarrow \boxed{J \left(\text{"the Jacobian of the transformation"} \right)}$$

Jacobian Matrix

Jacobian Computation Example

(* assuming the partials of z & w w.r.t. x, y exists & the determinate is not equal to zero)

X & Y be 2 ind. RVs that are independent and identically distributed (iid or i.i.d.) Gaussian RVs

where
$$\mu_X = \mu_Y = 0$$
$$\sigma_X = \sigma_Y = \sigma$$
$$r = 0$$

let
$$R = \sqrt{X^2 + Y^2}$$
$$\Theta = \tan^{-1}(Y/X)$$
(Polar Coordinates)

(theta)

... Change of Variables formula to get...

$$f_{R\Theta}(r,\theta) = \frac{r}{2\pi\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right] u(r) \quad , \quad -\pi \le \theta \le \pi$$

$$\Longrightarrow \boxed{f_R(r) = \int_{-\infty}^{\infty} f_{R\Theta}(r,\theta)\, d\theta = \frac{r}{\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right] u(r)}$$

(which is just from minus pi to pi of the equation above)

RAYLEIGH PDF

(Using CofV for finding 1 fn of 2 RVs)

* Note that for 1 fn of 2 RVs, sometimes it is easiest to find an auxillary or dummy second function (via change-of var.) in order to approach 2 functions of 2 RVs, then integrate to get the marginal pdf of interest

(common choice is to let the 2^{nd} rv be $W = X$ or $Z = Y$)

Joint Expectation

(Expectation may be viewed as the integration of a function times the density function)

Given 2 RVs, X & Y, Let $Z = g(X,Y)$ for some $g: \mathbb{R}^2 \to \mathbb{R}$

then $E[Z] = \int_{-\infty}^{\infty} z f_Z(z)\, dz$, however, instead of this

form, often used is the result that the expected value of $g(x,y)$ is ...

("Linearity of Expectation")

JE IS A LINEAR OPERATOR

$$\boxed{E[Z] = E[g(X,Y)] = \iint_{\mathbb{R}^2} g(x,y) f_{XY}(x,y)\, dx\, dy}$$

CORRELATION

or

$$\boxed{E[Z] = \sum_{x \in R_X} \sum_{y \in R_Y} g(x,y)\, p_{XY}(x,y)}$$

Joint Expectation may take the form of a correlation btwn 2 RVs

(or correlation coefficient)

Commonly Used Forms

(Still using $E[X]$, $E[Y]$, σ_X^2, σ_Y^2)

$|r_{XY}| \leq 1$

$Corr(X,Y) \equiv E[XY]$ ← Correlation of X & Y

$Cov(X,Y) \equiv E[(X-\mu_X)(Y-\mu_Y)]$ ← Covariance of X & Y (normalized correlation)

$r_{XY} \equiv \dfrac{Cov(X,Y)}{\sigma_X \sigma_Y}$ (or "ρ_{XY}") ← Correlation Coefficient

(via subtracting the means)

* If X, Y are ind. $\Rightarrow r_{XY} = 0$ (converse is not necessarilly true)

* If $r_{XY} = 0 \Rightarrow$ X & Y are UNCORRELATED

* If $E[XY] = 0 \Rightarrow$ X & Y are ORTHOGONAL

$\therefore$ X, Y are uncorrelated IFF

$Cov(X,Y) = 0$

(or) $E[XY] = \mu_X \mu_Y$

(Show that either is true)

Cauchy-Schwartz Inequality (useful for optimization)

For RVs X & Y, the |expected value of XY| is less than or equal to sqrt of exp. val. of X^2 times exp. val. of Y^2

$$\Rightarrow \boxed{\left|E[XY]\right| \leq \sqrt{E[X^2]E[Y^2]}}$$

(abs value)

iff $\boxed{Y = a_0 X \text{ with probability } 1}$ for some constant $a_0 \in \mathbb{R}$

Consider the inequality: $0 \leq E[(aX-Y)^2] = a^2E[X^2] - 2aE[XY] + E[Y^2]$

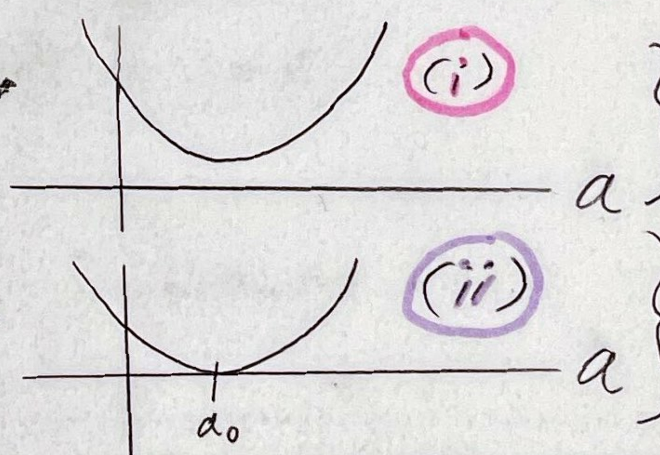
$\Rightarrow$ (quadratic fn. of the real number a) $\boxed{0 \leq E[(aX-Y)^2] = E[X^2]a^2 - 2E[XY]a + E[Y^2]}$

$a \in \mathbb{R}$

2 Cases:

(i) $E[(aX-Y)^2] > 0$

(ii) $E[(aX-Y)^2] = 0$



(i) 2 COMPLEX ROOTS

(ii) REPEATED ROOTS

(i) $4(E[XY])^2 - 4E[X^2]E[Y^2] < 0$

$$\Rightarrow \boxed{E[XY] < \sqrt{E[X^2]E[Y^2]}}$$

(ii) $\exists a_0 \in \mathbb{R} \mid E[(a_0X-Y)^2] = 0$

$\Rightarrow$ Can then be shown that if a RV X has $E[X^2]=0$, then $X(\omega)=0 \ \forall \omega \in S$ except possibly on some set A with $P(A)=0$

ie. $\exists B \in \mathcal{F}$ [illegible] $\mid X(\omega)=0 \ \forall \omega \in S$

$\boxed{Y = a_0 X \text{ with prob.}}$

Characteristic & Moment Generating Functions

* **Characteristic Function** ← Alternative representation of the prob. distributions of a RV

Representations:
- Measure P
- cdf F
- pdf f
- (pmf p) ← (discrete RVs, dne for all RVs)
- Char. fn Φ (for finding moments)
- MGF ϕ

* finding moments
* finding distributions for sums of RVs
* modeling in some cases

1 RV) Find char. fn of 1 RV X

$$\boxed{\Phi_X(\omega) = E[e^{i\omega X}], \forall \omega \in \mathbb{R} \quad i = \sqrt{-1}}$$

(Complex valued function)

$\Longleftrightarrow$

$$\boxed{\Phi_X(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f_X(x)\,dx, \forall \omega \in \mathbb{R}}$$

(Fourier Transform -esque) ← (difference is the sign is negative in the exponential)

$\Rightarrow$ Can be shown that $f_X(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \Phi_X(\omega) e^{i\omega x}\,d\omega$

CF
MGF

* **Moment Generating Function** ← of RV X

$$\boxed{\phi_X(s) \equiv E[e^{sX}], \; s \in \mathbb{R} \text{ or } s \in \mathbb{C}}$$

$\Longleftrightarrow$

$$\boxed{\phi_X(s) = \int_{-\infty}^{\infty} e^{sX} f_X(x)\,dx}$$

(Laplace Transform -esque)

* If density fn DNE, it is possible to define Φ_X & ϕ_X in terms of P or F_X
* It is possible for the char fn, Φ_X, to exist but the MGF, ϕ_X, DNE (E.g., Cauchy Random Variable function)

The Moment Generating Theorem

Given RV X w/ mgf ϕ_X, the <u>n^{th} moment</u> of X may be computed as

(Expected value of $[X^n]$)

$$\boxed{E[X^n] = \phi^{(n)}(0) \equiv \frac{d^n}{ds^n}\phi_X(s)\Big|_{s=0}}$$

("evaluated at zero") (n^{th} order derivative w.r.t. s)

$$\frac{d^n}{ds^n}\phi_X(s) = \frac{d^n}{ds^n}E[e^{sX}] = E\left[\frac{d^n}{ds^n}e^{sX}\right] = E\left[X^n e^{sX}\right]$$

$$\Rightarrow \boxed{\phi^n(0) = E[X^n]}$$

Similarlly, the result may be written in terms of the charecteristic function

---

2RV) The joint characteristic function of RVs X & Y:

$$\Phi_{XY}(\omega_1, \omega_2) \equiv E\left[e^{i(\omega_1 X + \omega_2 Y)}\right], \text{ for } \omega_1, \omega_2 \in \mathbb{R}$$

$$\Rightarrow \Phi_{XY}(\omega_1, \omega_2) = \iint_{\mathbb{R}^2} e^{i(\omega_1 x + \omega_2 y)} f_{XY}(x,y)\,dx\,dy$$

$$= \sum\sum xy\, P_{XY}(x,y) \text{ for } \underline{\text{discrete RVs}}$$

Joint MGF:

$$\boxed{\phi_{XY}(s_1, s_2) \equiv E\left[e^{s_1 X + s_2 Y}\right],\ s_1, s_2 \in \mathbb{C}}$$

Joint Moments:

$$\boxed{\mu_{jk} = E[X^j Y^k] = \frac{\partial^j \partial^k}{\partial s_1^j \partial s_2^k}\phi_{XY}(s_1, s_2)\Big|_{s_1=0,\, s_2=0}}$$

---

* $\Phi_X(\omega) = \Phi_{XY}(\omega, 0)$, $\forall \omega \in \mathbb{R}$ & $\Phi_Y(\omega) = \Phi_{XY}(0, \omega)$, $\forall \omega \in \mathbb{R}$
* If $Z = aX + bY$ for $a, b \in \mathbb{R}$, then $\Phi_Z(\omega) = \Phi_{XY}(a\omega, b\omega)$
* If X, Y are ind. then $\Phi_{XY}(\omega_1, \omega_2) = \Phi_X(\omega_1)\Phi_Y(\omega_2)$
* If $Z = X + Y$ & X, Y are ind. then $\Phi_Z(\omega) = \Phi_X(\omega)\Phi_Y(\omega)$, $\forall \omega \in \mathbb{R}$

$\therefore$ density fn of $Z \Leftrightarrow$ Convolution of density fn of X & density fn of Y

*Iff $Z = X + Y$ & X, Y are i.i.d $\Phi_Z(\omega) = (\Phi(\omega))^2$

Conditional Distributions for 2 RVs

* Recall: Conditional CDF of RV X, given that event B occurs, is

$$F_X(x|B) \equiv P(X \le x | B), \text{ if } P(B) > 0$$

* Consider: $B = \{y_1 < Y \le y_2\}$, for RV Y & $y_1, y_2 \in \mathbb{R}$, $y_1 < y_2$

* Derive: $f_X(x|Y=y)$, for $y \in \mathbb{R}$

$$\left(\begin{matrix}\text{Conditional}\\\text{Probability}\end{matrix}\right) \Rightarrow \underbrace{F_X(x|y_1 < Y \le y_2)}_{\text{(Conditional CDF)}} = \frac{P(X \le x, y_1 < Y \le y_2)}{P(y_1 < Y \le y_2)} = \underbrace{\frac{F_{XY}(x,y_2) - F_{XY}(x,y_1)}{F_Y(y_2) - F_Y(y_1)}}_{\text{(differentiate w.r.t. "x" to get cond. PDF)}}$$

* **Integral Form:**

$$f_X(x|y_1 < Y \le y_2) = \frac{\frac{\partial}{\partial x}\left[\int_{-\infty}^{x}\int_{-\infty}^{y_2} f_{XY}(x',y)\,dx'dy - \int_{-\infty}^{x}\int_{-\infty}^{y_1} f_{XY}(x',y)\,dx'dy\right]}{F_Y(y_2) - F_Y(y_1)}$$

$$= \frac{\int_{-\infty}^{y_2} f_{XY}(x,y)\,dy - \int_{-\infty}^{y_1} f_{XY}(x,y)\,dy}{F_Y(y_2) - F_Y(y_1)} = \frac{\int_{y_1}^{y_2} f_{XY}(x,y)\,dy}{F_Y(y_2) - F_Y(y_1)}$$

* **Want to condition on $\{Y=y\}$, $y \in \mathbb{R}$?**

$\Rightarrow$ Define:

$$f_X(x|Y=y) \equiv \lim_{\Delta y \to 0}\left[f_X(x|y < Y \le y + \Delta y)\right]$$

* Let $y_1 = y$, $y_2 = y + \Delta y$

$$\Rightarrow f_X(x|y < Y \le y + \Delta y) = \frac{\int_y^{y+\Delta y} f_{XY}(x,y')\,dy'}{F_Y(y+\Delta y) - F_Y(y)}$$

* Multiply by $\frac{(1/\Delta y)}{(1/\Delta y)}$ & then take the limit of the numerator & denominator as $y \to 0$

$$\Rightarrow \boxed{f_X(x|Y=y) = \frac{f_{XY}(x,y)}{f_Y(y)}}$$

* **Notation:** $f_{X|Y}(x|y) \equiv f_X(x|Y=y)$, often written as $f(x|y)$ in practice

$$\boxed{f_{X|Y}(x|y) = \frac{f_{XY}(x,y)}{f_Y(y)}} \quad \& \& \quad \boxed{f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}}$$

$\Rightarrow$

Bayes' Theorem:

$$\boxed{f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}, \quad \forall (x,y) \in \mathbb{R}}$$

Total Probability Law:

$$\boxed{f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_{-\infty}^{\infty} f_{Y|X}(y|x) f_X(x)\,dx}$$

3 Forms of Bayes' Theorem (Summary)

* Note for X & Y as ind. RVs,

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{f_X(x)\,f_Y(y)}{f_X(x)} = f_Y(y) \quad \Bigg\} \text{(i.e., } f_{Y|X}(y|x) \text{ does not depend on } x\text{)}$$

(1) $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad A, B \in \mathcal{F}$$ (for when X & Y are DISCRETE)

where $A = \{X = x\}$, $B = \{Y = y\}$ $\Rightarrow$ $$P_{Y|X}(y|x) = \frac{P_{X|Y}(x|y)P_Y(y)}{P_X(x)}$$

where $P_{Y|X}(y|x) \equiv P(Y=y|X=x)$ && $P_{X|Y}(x|y) \equiv P(X=x|Y=y)$

$\forall x \in R_X, y \in R_Y$

(Conditional pmfs)

(2) $$P(A|Y=y) = \frac{f_Y(y|A)\,P(A)}{f_Y(y)}, \quad A \in \mathcal{F}$$ (for when Y is CONTINUOUS, X is DISCRETE)

where $A = \{X = x\}$, $x \in R_X$ $\Rightarrow$ $$P_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)\,P_X(x)}{f_Y(y)}$$

(3) $$f_{Y|X}(y|x) = \frac{f_{X|Y}(x|y)\,f_Y(y)}{f_X(x)}$$ (for when X & Y are CONTINUOUS)

Conditional Expectation (3rd form of expectation)

* Recall: Expectation is the integral of a function (g) times the density function

* In this case, the density function is a conditional density function

* Let X, Y be 2 RVs on $(\mathcal{S}, \mathcal{F}, P)$, & $g: \mathbb{R}^2 \to \mathbb{R}$

$\Rightarrow$ What is the expected value of $g(X,Y)$ given that $Y = y$

$$E[g(X,Y)|Y=y] = \iint_{\mathbb{R}^2} g(x,y')\underbrace{f_{XY}(x,y'|Y=y)}\,dx\,dy'$$

(Conditioning on a continuous RV equalling a particular real value with the event probability 0)

$\Rightarrow$ Conditioning on $y < Y \leq y + \Delta y$ & letting $\Delta y \to 0$

* Can then show that:

$$E[g(X,Y)|Y=y] = \int_{-\infty}^{\infty} g(x,y)\,f_{XY}(x|y)\,dx$$

Note that it can be shown:

$E[g(X,Y)|Y=y] =$

$E[g(X,y)|Y=y]$

(Now X is the only random variable)

* Often have $g(X,Y) = g(X)$

$$\Rightarrow E[g(X)|Y=y] = \int_{-\infty}^{\infty} g(x)\,f_{X|Y}(x|y)\,dx \leftarrow \text{Special Case}$$

* Another special case: $g(X) = x$

$$\Rightarrow E[X|Y=y] = \int_{-\infty}^{\infty} x\,f_{X|Y}(x|y)\,dx \leftarrow \text{Used in iterated expectation}$$

Conditional Expectation: Iterated Expectation

* Sometimes it is easiest to find $E[g(X,Y)]$ via $f_{Y|X}(y|x)$ & $f_X(x)$ instead of the joint density $f_{XY}(x,y)$

$$\Rightarrow E[g(X,Y)] = \iint_{\mathbb{R}^2} g(x,y) f_{XY}(x,y)\,dx\,dy = \iint_{\mathbb{R}^2} f_X(x)\, g(x,y)\, f_{Y|X}(y|x)\,dy\,dx$$

(joint density function)

(conditional density × marginal density)

$$E[g(X,Y)] = \int_{\mathbb{R}} f_X(x)\, \underbrace{E[g(X,Y)|X=x]}_{h(x)}\,dx$$

* Note that $E[g(X,Y)|X=x]$ is a function of $x \in \mathbb{R}$, this function is called h

$$\Rightarrow \boxed{h(x) \equiv E[g(X,Y)|X=x]} \quad h: \mathbb{R} \to \mathbb{R}$$

(thus, we can create a RV $h(X)$ via making it depend on X)

$$\therefore \boxed{h(X) \equiv E[g(X,Y)|X]}$$ (DIFFERENT FUNCTIONS)

$$\Rightarrow \boxed{E[g(X,Y)] = \int_{\mathbb{R}} E[g(X,Y)|X=x]\, f_X(x)\,dx = \int_{\mathbb{R}} h(x) f_X(x)\,dx = E[h(X)]}$$

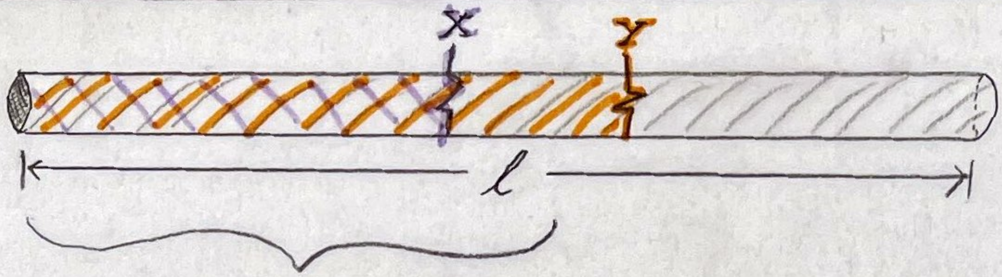
* Alternative Notation: $\boxed{E[g(X,Y)] = E[E[g(X,Y)|X]]}$

"Iterated Expectation"

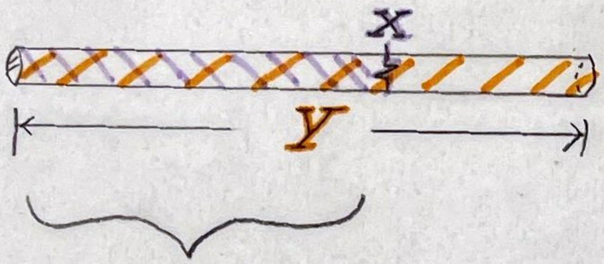
* Important Case: $g(X,Y) = Y$ $\Rightarrow$ $E[Y] = E[E[Y|X]]$

Iterated Expectation Example

* let l be the length of a stick, & then suppose breaking the stick at a uniformly chosen point Y, & then again at a uniformly chosen point X. Find $E[X]$

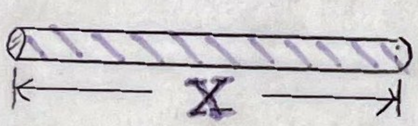


$\Downarrow$



$$\left(\underbrace{\iint_{\mathbb{R}^2} x f_{X|Y}(x|y)\, dx\, f_Y(y)\, dy} \right)$$

$\Downarrow$



$$E[X] = \iint_{\mathbb{R}^2} x f_{XY}(x,y)\, dx\, dy = \int_{\mathbb{R}} x f_X(x)\, dx$$

* Do not know f_{XY} or f_X

* Do know $f_{X|Y}$ and f_Y $\rightarrow$

$$\left. \begin{array}{ll} f_Y(y) = 1/l, & 0 \le y \le l \\ f_{X|Y}(x|y) = 1/y, & 0 \le x \le y \end{array} \right\}$$

now use $E[X] = E[E[X|Y]]$

$\Rightarrow$ $E[X|Y]$ $\leftarrow$ Find from $E[X|Y=y] = \int_{x=0}^{x=y} x f_{X|Y}(x|y)\, dx = \int_{x=0}^{x=y} x \left(\frac{1}{y}\right) dx = \left. \frac{x^2}{2y} \right|_{x=0}^{x=y}$

$\Rightarrow$ $\boxed{E[X|Y=y] = \frac{y}{2} \equiv h(y)}$ $\therefore$ $\boxed{h(Y) = \frac{Y}{2}}$

(Since X is Uniform on $[0, y]$ given $Y=y$)

$\Rightarrow$ $E[X|Y] \equiv h(Y) = \frac{Y}{2}$

$\therefore$ $\boxed{E[X] = E[E[X|Y]] = E\left[\frac{Y}{2}\right] = \frac{1}{2} E[Y] = \frac{l}{4}}$

$\because$ Y is a uniform RV on $[0, l]$