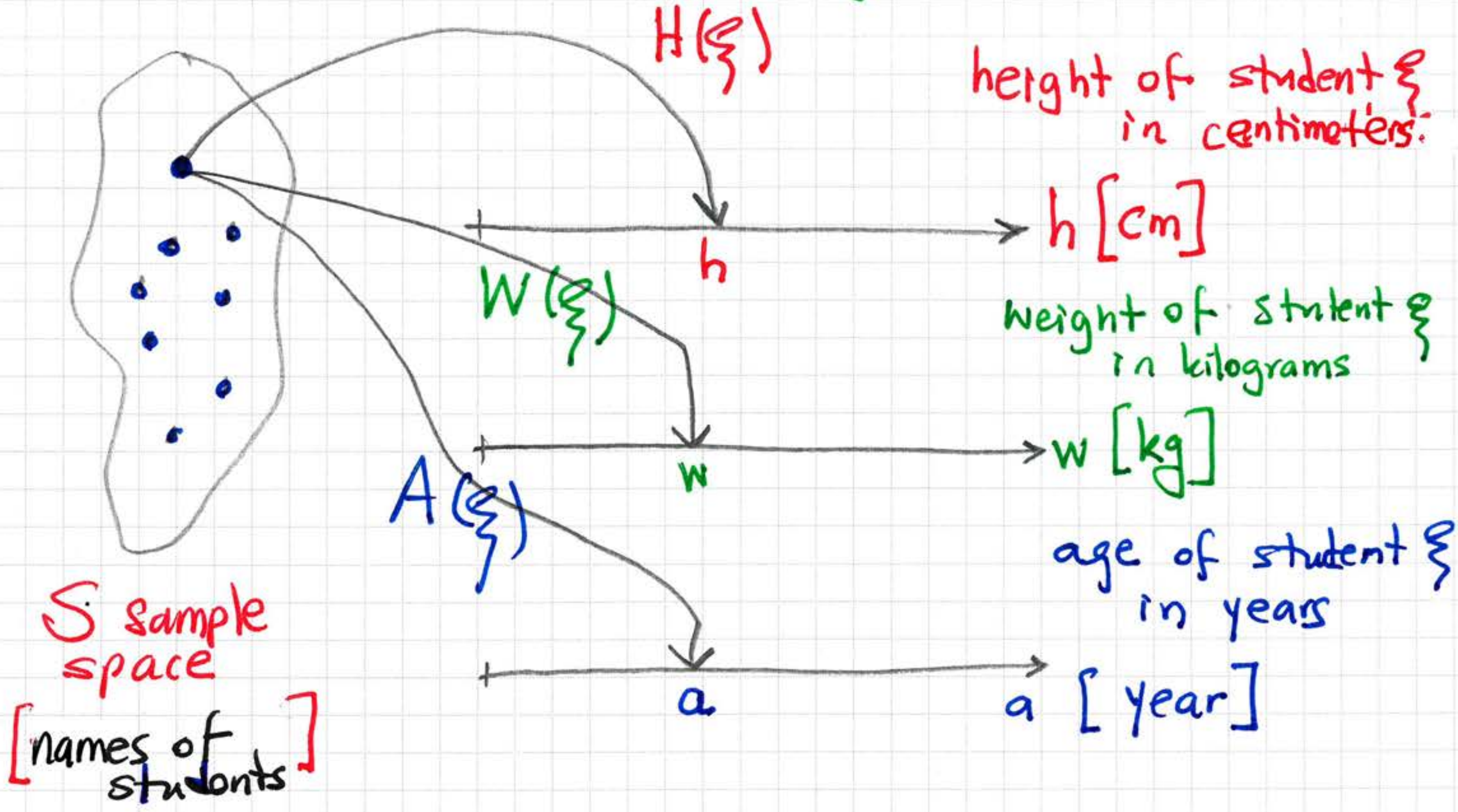
Chapter 4

# MULTIPLE RANDOM VARIABLES

- So far, in Chapter 3 of our texbook, we have developed techniques for calculating the probabilities of events involving a single random variable in isolation.
- Now, at this chapter, we develop methods for calculating the probabilities of events that involve the joint behavior of two or more random variables.
- We are interested in determining when a set of random variables are independent, as well as in quantifying their degree of "Correlation" when they are not independent.
- For this purpose, we present some fundametal notions about multiple random variables. We discuss the case of two random variables in detail. Because in this case, at 2-Dimensional (2-D) case, we can draw on our geometric intuition.
- Finally, we will generalize our 2-dimentional results to the general case of multiple n-dimensional random variables ($n > 2$).

## Example:

Let a random experiment consists of selecting a student's name from a box. Let $\xi$ denote the outcome of this experiment, we define the following 3 functions:




The vector,

$$\underline{X} = \begin{bmatrix} H(\xi) \\ W(\xi) \\ A(\xi) \end{bmatrix}$$

is a three-dimontional random vector.

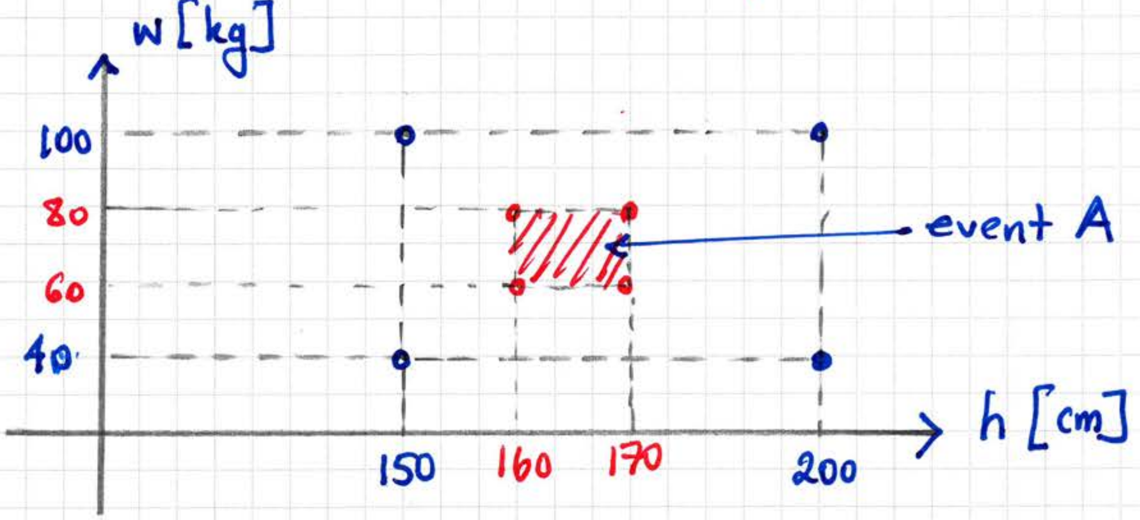
- For example, we might have a population of college students, all heights and weights lie in the interval

$$150 \leq H \leq 200 \text{ cms} \quad \text{and}$$
$$40 \leq W \leq 100 \text{ kgs.}$$

- Therefore, the continuous random variables $(H, W)$ would take on values in the 2-D sample space;

$$R_{H,W} = \{(h,w);\ 150 \leq h \leq 200,\ 40 \leq w \leq 100\}$$

which is a subset of the plane, $\mathbb{R}^2$.




- We might wish to determine the probability of the even A,

$$\Pr\{A\} = \Pr\{160 \leq H \leq 170;\ 60 \leq W \leq 80\}.$$

- In order to compute such a probability, we will defina a joint pdf (or pmf) for the continuous (or discrete) random variables H and W. $f_{HW}(h,w)$ (or $P_{HW}(h,w)$) will be a 2-D function of h and w.

## A: CONTINUOUS CASE

- In the case of a continuous single random variable, we needed to find the area under the pdf as the desired probability.

- Now, the integration of the joint pdf, which is a function of two variables, will produce probability.
- It is interesting to note that, we will now determining the volume under the joint pdf.

- Indeed, all our concepts for a single random variable will be extended to the 2-dimentional case.
- For $\underline{X} = \begin{bmatrix} X \\ Y \end{bmatrix}$ or $(X,Y)$ random variables, we define a joint pdf

$f_{XY}(x,y)$ which satisfy the following properties:

(1) $f_{XY}(x,y) \geq 0 \quad \forall (x,y) \in R_{XY}$ (range)

(2) $$\iint_{R_{XY}} f_{XY}(x,y)\,dx\,dy = 1$$

(3) The probability of an event A in the range $R_{XY}$ is defined as

$$\Pr\{(X,Y) \in A\} = \iint_{A} f_{XY}(x,y)\,dx\,dy.$$

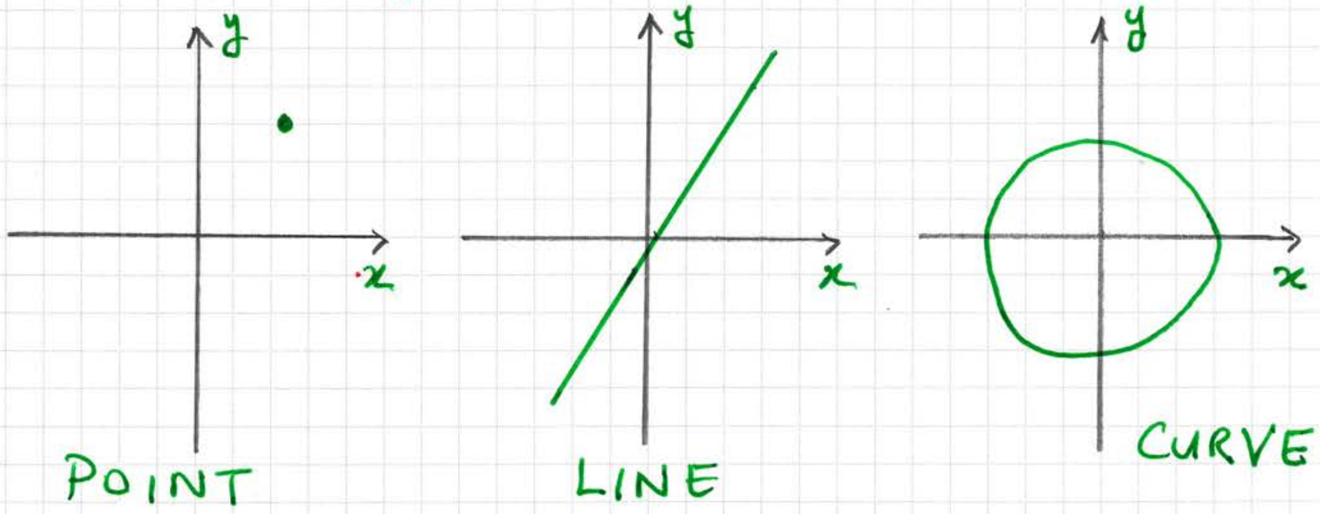
Important Remark:

- Remember that for a single continuous random variable, the probability of X attaining any valu is zero, namely,

$$\Pr\{X = x\} = 0.$$

This is because the area under the pdf $f_X(x)$ is zero for any zero length interval.

- Similarly, for jointly continuous random variable X and Y, the probability of an event defined on the xy-plane will be zero if the region of the event in the plane has zero area. Then the volume under the joint pdf will be zero. Some examples are show here.




All events have zero area. The probabilities of these events are zero.

B: <u>DISCRETE CASE</u>

In the case of X and Y random variables, the joint probability mass functions of (X, Y) specifies the probablities of the product-form event

$$P_{XY}(x_j, y_k) = \Pr\{X = x_j \text{ and } Y = y_k\}$$
for all $x_j, y_k$.

We observe the following properties, which are direct extensions for the single-random-variable case:

(1) $p_{XY}(x_j, y_k) \leq 1$

(2) $\sum_j \sum_k p_{XY}(x_j, y_k) = 1$

(3) $Pr\{(X,Y) \in A\} = \sum\sum_{(x_j, y_k) \in A} p_{XY}(x_j, y_k)$

Example: [T.T. Soong]

In the structural reliability studies, the resistance Y of a structural element and the force X applied to it are generally regarded as random variables.

Suppose that the joint pdf of X and Y is specified to be

$$f_{XY}(x,y) = \begin{cases} ab\,e^{-(ax+by)}, & \text{for } (x,y) > 0 \\ 0, & \text{for } (x,y) \leq 0, \end{cases}$$

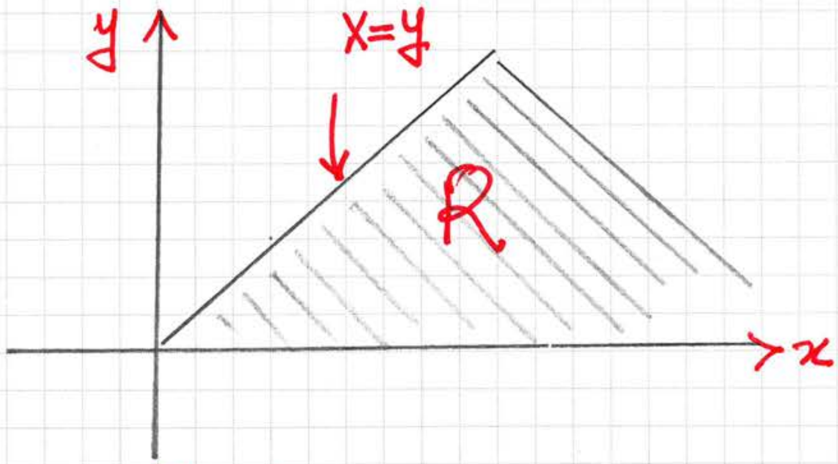
where a and b are known positive constants.

- The probability of failure $P_f$, is defined by $Pr\{Y \le X\}$, namely,

$$P_F = Pr\{Y \le X\}$$

$$= \iint_R f_{XY}(x,y)\,dx\,dy$$

where $R$ is the region satisfying $Y \le X$. Since $X$ and $Y$ take only positive values, the region $R$ is shown as follows;




$$P_f = \int_0^{\infty}\int_y^{\infty} ab\,e^{-(ax+by)}\,dx\,dy$$

$$= \frac{b}{a+b}$$ • The Probability of failure!

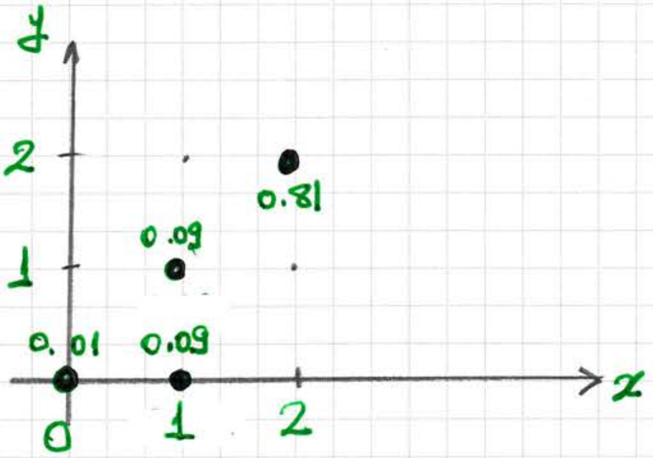
## The Representation of the joint pmf:

The joint pmf can be given as a set of labeled points in the $(x,y)$-plane where each point is a possible value (probability $>0$) of the pair $(x,y)$:

as a simple list,

$$P_{X,Y}(x,y) = \begin{cases} 0.81 &, \quad x=2,\ y=2 \\ 0.09 &, \quad x=1,\ y=1 \\ 0.09 &, \quad x=1,\ y=0 \\ 0.01 &, \quad x=0,\ y=0 \\ 0 &, \quad \text{otherwise.} \end{cases}$$

OR




or a third representation of $P_{X,Y}(x,y)$ is the matrix:

| $P_{XY}(x,y)$ | $y=0$ | $y=1$ | $y=2$ |
|---|---|---|---|
| $x=0$ | 0.01 | 0 | 0 |
| $x=1$ | 0.09 | 0.09 | 0 |
| $x=2$ | 0 | 0 | 0.81 |

- Note that all the probabilities add up to 1. This reflects the second axiom of probability that states $\Pr\{S\}=1$. We write this

$$\sum\sum_{(x,y)\in R_{XY}} p_{XY}(x,y) = 1.$$

Example:

For the joint pdf given above, find the probability of the event B, where $B=\{X=Y\}$, namely,

$$B=\{(0,0),\ (1,1),\ (2,2)\} \quad \text{and}$$

$$\Pr\{B\} = p_{XY}(0,0) + p_{XY}(1,1) + p_{XY}(2,2)$$

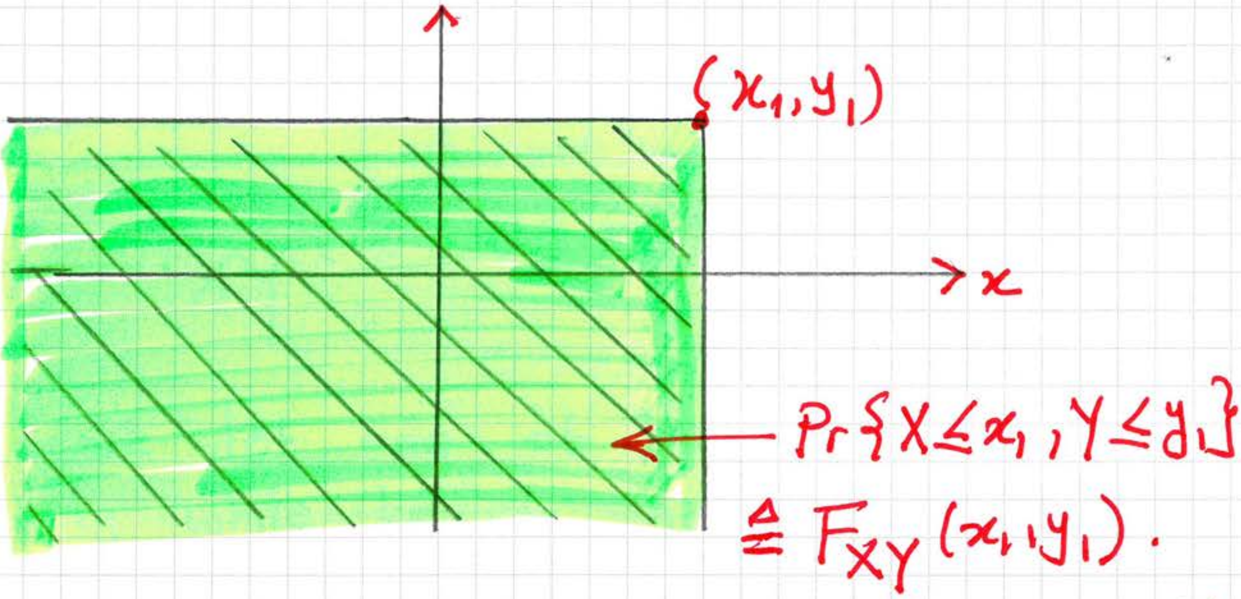
$$= 0.01 + 0.09 + 0.81 = 0.91.$$

Joint Cumulative Distribution Function of (X,Y)

[or Joint Distibution Function of (X,Y)]

It is defined as

$$F_{XY}(x,y) = \Pr\{X\le x \text{ and } Y\le y\}$$
$$= \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(u,t)\,du\,dt.$$




The joint cdf is defined as the probability of the semi-infinite rectangle by $(x_1, y_1)$.

## Properties of the joint cdf:

(1) $F_{XY}(x, y)$ is a nondecreasing function in the "northeast" direction, that is,

If $x_1 \le x_2$ and $y_1 \le y_2$, then

$$F_{XY}(x_1, y_1) \le F_{XY}(x_2, y_2).$$

(2) It is impossible for either $X$ or $Y$ to assume a value less than $-\infty$. Therefore,

$$F_{XY}(-\infty, y) = F_{XY}(x, -\infty) = 0.$$

(3) It is certain that X and Y assume less than infinity. Therefore,

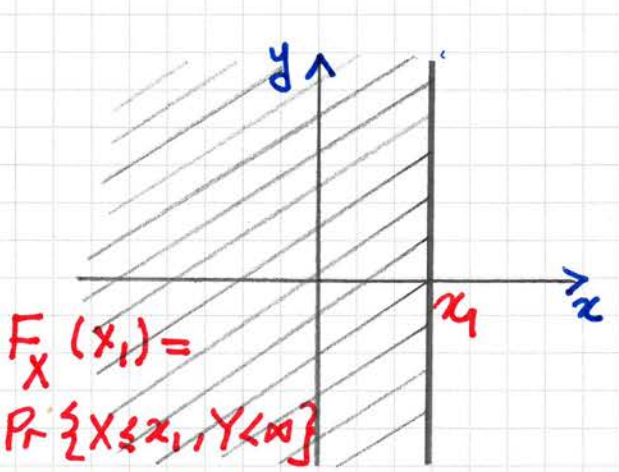
$$F_{XY}(\infty, \infty) = 1.$$

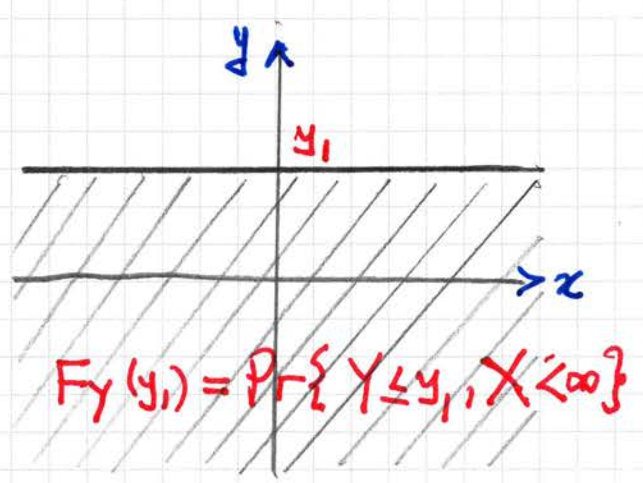
(4) If we let one of the variables approach infinity while keeping the other fixed, we obtain the <u>marginal</u> cumulative distribution functions.

$$\begin{aligned} F_X(x) &= F_{XY}(x, \infty) \\ &= Pr\{X \leq x, Y < \infty\} \\ &= Pr\{X \leq x\}, \end{aligned}$$

and similarly,

$$\begin{aligned} F_Y(y) &= F_{XY}(\infty, y) \\ &= Pr\{X < \infty, Y \leq y\} = Pr\{Y \leq y\}. \end{aligned}$$



(5) If X and Y jointly continuous, joint pdf can be obtained by taking the derivative,

$$f_{XY}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) .$$

At this expression, we assume the existance of the derivative.

## MARGINAL PROBABILITY DISTRIBUTIONS

In an experiment that produces two-random variables X and Y, it is always possible to consider one of the random variables, X, and ignor the other one, Y.

In order to determine the marginal pdf $f_X(x)$ from the given joint pdf $f_{XY}(x,y)$, we consider the following event:

$$A = \{ (x,y) : a \leq x \leq b , -\infty < y < \infty \}$$

whose probability must be the same as

$$A_X = \{ x : a \leq x \leq b \} .$$

There fore,

$$\Pr\{a \le X \le b\} = \Pr\{A_X\} = \Pr\{A\}$$

$$= \iint_A f_{XY}(x,y)\,dx\,dy$$

$$= \int_{y=-\infty}^{\infty} \int_{x=a}^{b} f_{XY}(x,y)\,dx\,dy$$

$$= \int_{x=a}^{b} \underbrace{\left[ \int_{y=-\infty}^{\infty} f_{XY}(x,y)\,dy \right]}_{f_X(x)} dx \ ,$$

and clearly we must have,

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

$f_X(x)$ is the marginal pdf of $X$.

Similarly, the marginal pdf of $Y$ is written as

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx.$$

- If we have a joint probability mass function (pmf) for discrete random variables $X$ and $Y$, we can obtain the marginal pmfs from the joint pmf:

$$P_X(x_j) = \Pr\{X = x_j\} = \sum_{\text{all } k} P_{XY}(x_j, y_k)$$

and

$$P_Y(y_k) = \Pr\{Y = y_k\} = \sum_{\text{all } j} P_{XY}(x_j, y_k).$$

Example:

Let us consider the following joint pmf;

| | $Y=0$ | $Y=1$ | $Y=2$ | $P_X(x)$ |
|---|---|---|---|---|
| $X=0$ | 0.01 | 0 | 0 | 0.01 |
| $X=1$ | 0.09 | 0.09 | 0 | 0.18 |
| $X=2$ | 0 | 0 | 0.81 | 0.81 |
| $P_Y(y)$ | 0.1 | 0.09 | 0.81 | |

To find the marginal pmfs $P_X(x)$ and $P_Y(y)$, we note that X and Y have the same range

$$R_X = R_Y = \{0, 1, 2\}.$$

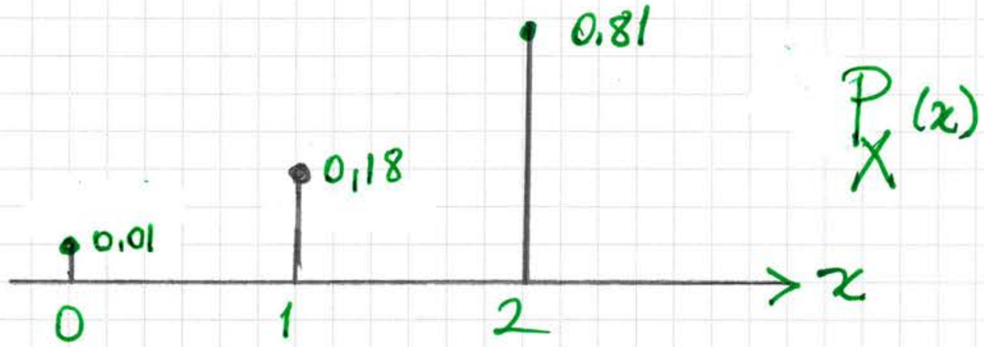
Therefore,

$$P_X(0) = \sum_{y=0}^{2} P_{XY}(0,y) = 0.1$$

$$P_X(1) = \sum_{y=0}^{2} P_{XY}(1,y) = 0.09 + 0.09 = 0.18$$

$$P_X(2) = \sum_{y=0}^{2} P_{XY}(2,y) = 0.81$$

$$P_X(x) = 0 \quad , \quad x \neq 0, 1, 2.$$



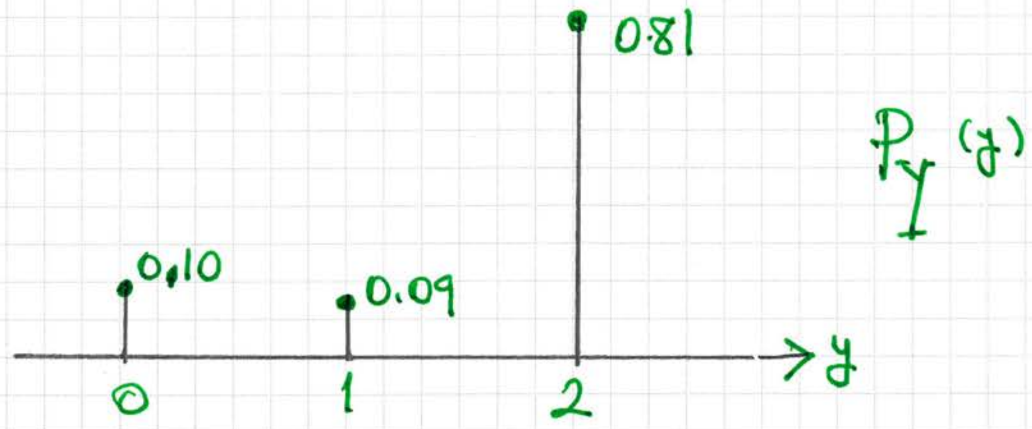

Similarly, for the marginal pmf of Y, we obtain,

$$P_Y(0) = \sum_{x=0}^{2} P_{XY}(x,0) = 0.01 + 0.09 = 0.10$$

$$P_Y(1) = \sum_{x=0}^{2} P_{XY}(x,1) = 0.09$$

$$P_Y(2) = \sum_{x=0}^{2} P_{XY}(x,2) = 0.81$$

$$P_Y(y) = 0 \quad , \text{ for } y \neq 0, 1, 2.$$




EXPECTED VALUES AND VARIANCIAS:

- We can find the mean and the variance of X and Y, from directly the joint pmf or joint pdf.
- If X and Y are discrete random veriables,

$$\mu_X = \sum_{\text{all } j} x_j \underbrace{\left[ \sum_{\text{all } k} P_{XY}(x_j, y_k) \right]}_{P_X(x_j)}$$

$$= \sum_{\text{all } j} x_j \, P_X(x_j),$$

and

$$\sigma_X^2 = \sum_{\text{all } j} (x_j - \mu_X)^2 \underbrace{\sum_{\text{all } k} P_{XY}(x_j, y_k)}_{P_X(x_j)}$$

$$= \sum_{\text{all } j} (x_j - \mu_X)^2 P_X(x_j).$$

## CONDITIONING BY A RANDOM VARIABLE

- **Conditional pmfs:**

For any event $Y = y_k$, such that $P_Y(y_k) > 0$, the conditional pmf of $X$ given $Y = y_k$ is,

$$P_{X|Y}(x_j \mid y_k) = \Pr\{X = x_j \mid Y = y_k\}$$

and is defined as follows,

$$P_{X|Y}(x_j|y_k) = \frac{P_{XY}(x_j, y_k)}{P_Y(y_k)}.$$

joint pmf of X and Y

Conditional pmf of X given $Y = y_k$

Marginal pmf of Y

- Conditional pdf:

If X and Y are continuous random variables, the conditional pdf is defined as,

- For y such that $f_Y(y) > 0$, the conditional pdf of X given $\{Y = y\}$ is,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

joint pdf of (X,Y)

conditional pdf of X

marginal pdf of Y

- This definition implies,

Joint pdf of (X,Y)

$$f_{Y|X}(y|x) = \frac{f_{XY}(x,y)}{f_X(x)}.$$

conditional pdf of Y

marginal pdf of X

# INDEPENDENT RANDOM VARIABLES

- From the chapter 1, we know the concept of independent events, events A an B are independent if and only if

$$Pr\{A \cap B\} = Pr\{AB\} = Pr\{A\} Pr\{B\}.$$

In other words, the probability of the intersection is the product of the individual probabilities.

- Applying the idea of independence to random variables, we can say that X and Y are independent random variables if the events $\{X=x\}$ and $\{Y=y\}$ are independent for all $x \in R_X$ and all $y \in R_Y$.

- Depending upon the discrete or the continuous random variables, we have the following definitions in terms of the pmfs and the pdfs, respectively,

Discrete Case:

$$P_{X,Y}(x_j, y_k) = P_X(x_j) P_Y(y_k), \text{ for all } x_j, \text{ all } y_k.$$

## Continuous Case:

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$$

for all $x$, all $y$.

- If $X$ and $Y$ are independent discrete random variables,

$$P_{X|Y}(x_j \mid y_k) = P_X(x_j) \quad , \text{ for all } y_k$$

$$P_{Y|X}(y_k \mid x_j) = P_Y(y_k) \quad , \text{ for all } x_j.$$

- If $X$ and $Y$ are independent continuous random variables, then we can write that

$$f_{X|Y}(x \mid y) = f_X(x) \quad , \text{ for all } y$$

$$f_{Y|X}(y \mid x) = f_Y(y) \quad , \text{ for all } x.$$

- For independent X and Y random variables, we show that they satify the following properties:

(1) $VAR[X+Y] = VAR[X] + VAR[Y]$

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$

(2) $r_{XY} \triangleq$ korelasyon between X and Y

$$r_{XY} = E[XY] = E[X]E[Y].$$

(3)$_a$ $\sigma_{XY} = Cov[X,Y] =$ kovaryans between X and Y

$$\sigma_{XY} = Cov[X,Y] = 0 \, .$$

(3)$_b$ $\rho_{XY} =$ Correlation Coefficient

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = 0$$

(4) $E[X \mid Y=y] = E[X]$, for all $y \in R_Y$

$E[Y \mid X=x] = E[Y]$, for all $x \in R_X$.

(5)

$$E[h(X)g(Y)] = E[h(X)]\,E[g(Y)].$$

Example:

If $X$ and $Y$ independent random variables, we can compute the pdf of

$$Z = X+Y$$

In terms of $f_X(x)$ and $f_Y(y)$ as follows:

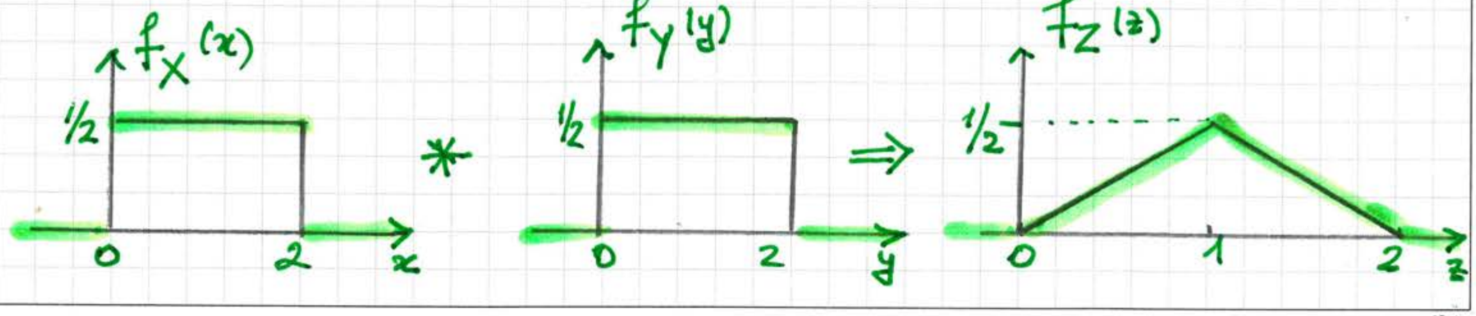
$$f_Z(z) = f_X(x) * f_Y(y)$$

$$= \int_{-\infty}^{\infty} f_X(z-y)\, f_Y(y)\, dy$$

$$= \int_{-\infty}^{\infty} f_X(x)\, f_Y(z-x)\, dx$$

Ex. 1:

If $X$ and $Y$ are uniformly distributed between 0 and 2, we have



Ex. 2: If $X \sim \mathcal{N}(0,1)$ and $Y \sim \mathcal{N}(0,1)$ are independent random variables, then $Z = X+Y$ is also a Gaussian random varible, $Z \sim \mathcal{N}(0,2)$.

## CONDITIONAL EXPECTATIONS [T.T. Soong]

- $E[X|Y]$ is a function of random variable $Y$ for which the value at $Y = y_i$ is $E[X|Y=y_i]$. Therefore, $E[X|Y]$ is itself a random variable.

- One of its very useful properties is that

$$E[X] = E\big[E[X|Y]\big].$$

If $Y$ is a discrete random variable taking on values $y_1, y_2, \ldots$, the above equation states that

$$E[X] = \sum_k E[X|Y=y_k]\, Pr\{Y=y_k\},$$

and

$$E[X] = \int_{-\infty}^{\infty} E[X|y]\, f_Y(y)\, dy$$

if $Y$ is continuous.

Proof:

Let us show this equation is true when both X and Y are discrete. Starting from the righ-hand side of equation, we have

$$\sum_k E[X \mid Y=y_k] \Pr\{Y=y_k\}$$
$$= \sum_k \sum_j x_j \Pr\{X=x_j \mid Y=y_k\} \Pr\{Y=y_k\}$$

However, we know that,

$$\Pr\{X=x_j \mid Y=y_k\} = \frac{\Pr\{X=x_j \cap Y=y_k\}}{\Pr\{Y=y_k\}},$$

we have

$$\sum_k E[X \mid Y=y_k] \Pr\{Y=y_k\} = \sum_k \sum_j x_j P_{XY}(x_j, y_k)$$
$$= \sum_j x_j \sum_k P_{XY}(x_j, y_k)$$
$$= \sum_j x_j P_X(x_j)$$
$$= E[X].$$

**Example:** The survival of a motorist stranded in a (mahsur kalmış) snowstorm depends on which of the three directions the motorist chooses to walk.

- The first road leads to safety after one hour of travel,
- The second leads to safety after three hours of travel,
- but the third will circle back to the original spot after two hours.

Determine the average time to safety if the motorist is equally likely to choose any one of the roads.

**Answer:**

Let $Y = 1, 2,$ and $3$ be the events that the motorist chooses the first, second and third road, respectively. Then,

$$\Pr\{Y=k\} = \frac{1}{3}, \quad \text{for } k = 1, 2, 3.$$

Let $X$ be the time to safety, in hours.

We have,

$$E[X] = \sum_{k=1}^{3} E[X \mid Y=k] \Pr\{Y=k\}$$
$$= \sum_{k=1}^{3} \frac{1}{3} E[X \mid Y=k].$$

Now, we can write

$$E[X \mid Y=1] = 1,$$

$$E[X \mid Y=2] = 3,$$

$$E[X \mid Y=3] = 2 + E[X].$$

if the motorist chooses the third road, then it takes 2 hours to find that he is back to the starting point and the problem is as before.

Hence

$$E[X] = \frac{1}{3}\left[1 + 3 + 2 + E[X]\right]$$

OR

$$E[X] = 3 \text{ hours.}$$

EXPECTATION OF A FUNCTION OF TWO RANDOM VARIABLES

$Z = H(X,Y)$ is a function of two random variables. The expected value of $Z$ can be written as follows

$$\mu_Z = \int_{-\infty}^{\infty} z\, f_Z(z)\, dz.$$

However, without calculating the pdf of $Z$, $f_Z(z)$, it is possible to find $\mu_Z$ from the joint pmf or pdf of the $X$ and $Y$ random variables.

$$\mu_Z = E[Z]$$
$$= E[H(X,Y)]$$
$$= \begin{cases} \iint H(x,y)\, f_{XY}(x,y)\, dx\, dy & \text{, Continuous case} \\ \sum_j \sum_k H(x_j, y_k)\, P_{X,Y}(x_j, y_k) & \text{, Discrete case} \end{cases}$$

COVARIANCE, CORRELATION AND CORRELATION COEFFICIENT

(1) Covariance ($\sigma_{XY}$)
between $X$ and $Y$ random variables is defined.

$$\sigma_{XY} = Cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)]$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x,y)\, dx\, dy - \mu_X \mu_Y .$$