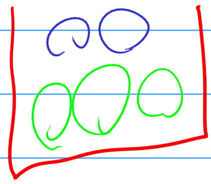
Discrete Joint Distribution



x	0	1	2
$f(x)$	$\frac{{}_2C_0\,{}_3C_2}{{}_5C_2}$	$\frac{{}_2C_1\,{}_3C_1}{{}_5C_2}$	$\frac{{}_2C_2\,{}_3C_0}{{}_5C_2}$

Y = no. of green
$\in \{0,1,2,3\}$

X = no. of blue.
$\in \{0,1,2\}$

y	0	1	2
$f(y)$	?	?	?

$$f_{XY}(x,y) = P\{X=x, Y=y\} \Leftrightarrow f_{XY}(x,y) \geq 0 \quad \forall (x,y)$$

$$\sum_x \sum_y f_{XY}(x,y) = 1$$

Over some region A:

$$P\{(x,y) \in A\} = \sum_A\sum f_{XY}(x,y)$$

PMF of X:

$$f_X(x) = P\{X=x\} = \sum_y f_{XY}(x,y)$$

PMF of Y:

$$f_Y(y) = P\{Y=y\} = \sum_x f_{XY}(x,y)$$

For example: $f_{XYZ}(x,y,z) \Rightarrow f_X(x) = \sum_y \sum_z f_{XYZ}(x,y,z)$

Independent Random Variables: $f_{XY}(x,y) = f_X(x) f_Y(y) \quad \forall (x,y) \in R_X \times R_Y$

Conditional Probability: $\boxed{f_{Y|X} = \dfrac{f_{XY}(x,y)}{f_X(x)}}$ if independent $\Rightarrow f_{Y|X} = f_Y(y)$

Example

Two balls are selected at random from a box that contains 3 blue balls, 2 red balls and 3 green balls. If X is the number of blue balls in the sample drawn and Y is the number of red balls selected. Find

- (a) The joint PMF of X and Y. $x \in \{0,1,2\}$, $y \in \{0,1,2\}$
- (b) The probability that sum of the number of red balls and blue balls is at most equal to 1 in the sample.
- (c) The marginal PMFs of X and Y and determine whether they are independent or not.
- (d) The expected number of blue balls in sample.

a)

Y \ X	0	1	2
0	$\dfrac{3C0 \times 2C0 \times 3C2}{8C2}$	$\dfrac{3C1 \times 2C0 \times 3C1}{8C2}$	$\dfrac{3C2 \times 2C0 \times 3C0}{8C2}$
1	$\dfrac{3C0 \times 2C1 \times 3C1}{8C2}$ b	$\dfrac{3C1 \times 2C1 \times 3C0}{8C2}$	0
2	$\dfrac{3C0 \times 2C2 \times 3C0}{8C2}$	0	0

b) $P\{Y+X \leq 1\} = \sum_{x+y\leq 1}\sum f(x,y)$, $(x,y) \in \{(0,1),(1,0),(0,0)\}$

c) $P\{X = x\} = \sum_{y} f(x,y)$, $P\{Y=y\} = \sum_{x} f(x,y)$

d) $\mathcal{E}[x] = \sum_x x f_X(x) = \sum_x x \sum_y f_{XY}(x,y) = \underbrace{\sum_x \sum_y x f(x,y)}$

$\mathcal{E}[xy] = \sum_y y \sum_x x \; f(x,y) = \sum_y \sum_x xy\, f(x,y)$

if independent $\Rightarrow$ $\mathcal{E}[xy] = \sum_y \sum_x xy\, f(x) f(y)$

$= \sum_y y f(y) \times \sum_x x f(x)$

$= \boxed{\mathcal{E}[y] \times \mathcal{E}[x]}$

Example

Two discrete R.V.s X and Y have the following joint PMF

$$f_{XY}(x,y) = k\left(\frac{1}{3}\right)^{x+y}, \quad x,y = 0,1,2,\cdots$$

- (a) Find the constant k.
- (b) Find $f_{Y|X}(y)$ and hence determine whether X and Y are independent or not.

$K \sum_{y=0}^{\infty} \sum_{x=0}^{\infty} \left(\frac{1}{3}\right)^x \left(\frac{1}{3}\right)^y = 1 \Rightarrow K = ??$

geometric series

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

Example

A manufacturing company uses two inspecting devices. The first inspection monitor is able to detect 90% of the defective items it receives, whereas the second is able to do so in 95% of the cases. Assume that three defective items are produced and sent out for inspection. Let X denote the number of items that are identified as defective by device 1 and let Y denote the number of items that are identified as defective by device 2, respectively. Assume that the devices are independent. Determine the joint probability distribution of X and Y.

2

Binomial

Bernolli

$$f_X(x) = 3C_x\,(0.9)^x\,(0.1)^{3-x} \quad , \quad f_Y(y) = 3C_y\,(0.95)^y\,(0.05)^{3-y}$$

$$f_{XY}(x,y) = f_X(x)\,f_Y(y) \quad , \quad x, y \in \{0, 1, 2, 3\}$$

Joint PDF

$$f_{XY}(x,y) \geq 0 \quad \forall (x,y)$$

$$\int_x \int_y f_{XY}(x,y)\,dx\,dy = 1 \quad ,$$

For any region A:

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

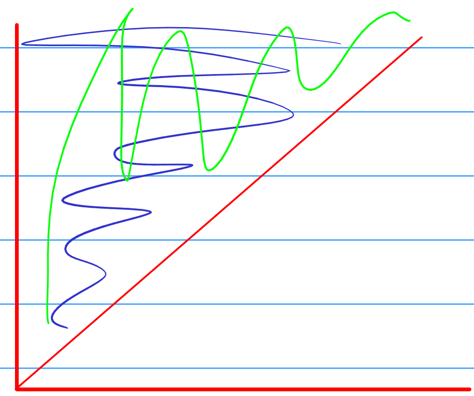
$$f_X(x) = \int_y f_{XY}(x,y)\,dy \quad \Leftrightarrow \quad f_Y(y) = \int_x f_{XY}(x,y)\,dx$$

Example

Let the random variable X denote the time until a computer server connects to your machine (in seconds) and let Y denote the time until the server authorizes you as a valid user (in seconds). Generally, it is known that the former is less than the later. Assume that the joint pdf for X and Y is given by

$$f_{XY}(x,y) = 6\exp[-x-2y] \quad x < y$$

- (a) Verify that $f_{XY}(x,y)$ is a valid PDF.
- (b) If a customer is satisfied if the time until a computer server connects to his machine is less than 1 second and that until the server authorizes him as a valid user is less than 2 seconds, what is the probability that the customer's requirements are satisfied?
- (c) Find the marginal distributions of X and Y and determine whether they are independent or not.



integral limits

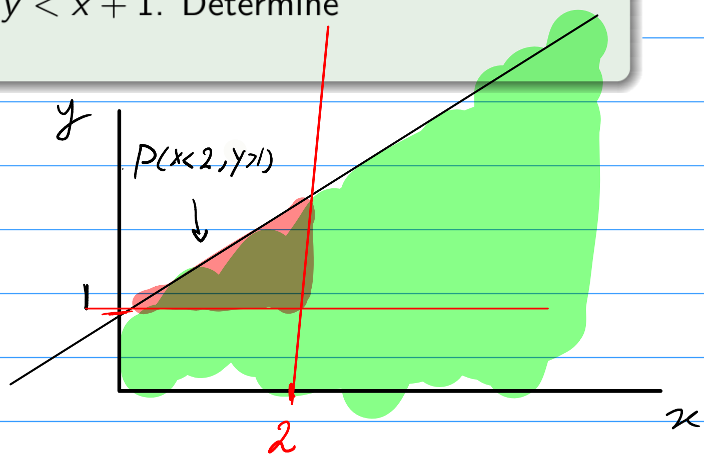
Example

Two random variables X and Y are uniformly distributed over the region determined by $0 < x < 4$ and $0 < y < x + 1$. Determine $P\{X < 2, Y > 1\}$.

$f(x,y) = k$

$\int_0^4 \int_0^{x+1} k \, dy \, dx = 1$

$P(X<2, Y>1)$



Expectations

$$\mathcal{E}[\phi(X,Y)] = \sum_x \sum_y \phi(X,Y) f(x,y)$$

$$\mathcal{E}[\phi(X,Y)] = \int_x \int_y \phi(X,Y) f(x,y) \, dy \, dx$$

$\mathcal{E}[XY] = \mathcal{E}[X]\,\mathcal{E}[Y] \iff$ independent

$$\mathrm{Var}[X+Y] = \mathcal{E}[(X+Y)^2] - (\mathcal{E}[X+Y])^2$$

$$= \mathcal{E}[X^2] + 2\mathcal{E}[XY] + \mathcal{E}[Y^2] - (\mu_x + \mu_y)^2$$

$$= \mathcal{E}[\cancel{X^2}] + 2\mathcal{E}[XY] + \mathcal{E}[\cancel{Y^2}] - \cancel{\mu_x^2} - 2\mu_x \mu_y - \cancel{\mu_y^2}$$

$=$ Continue this.

Covariance: measure of association.

$\mathrm{Cov}[X,Y] = 0 \Rightarrow \mathcal{E}[XY] = \mu_x \mu_y$

Example

Suppose that $X \sim \mathcal{U}(-1,1)$ and $Y = X^2$, find the covariance of X and Y.

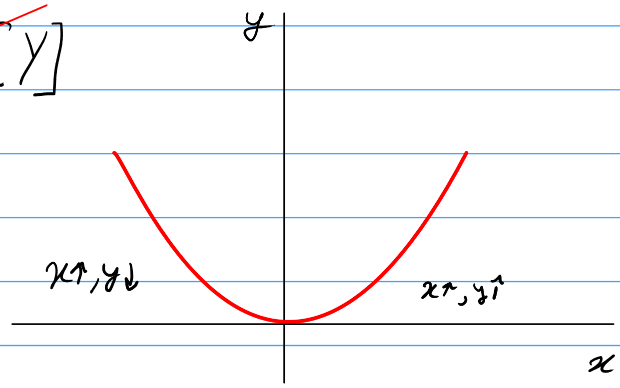
$X \sim \mathcal{U}(-1,1)$, $Y = X^2$

$$Cov[X,Y] = \mathcal{E}[XY] - \cancel{\mathcal{E}[X]\mathcal{E}[Y]}$$

$$= \mathcal{E}[X^3]$$

$$= k\int_{-1}^{1} x^3 dx = 0$$

$\Rightarrow$



$\boxed{\text{dependent but uncorrelated}}$

Covariance and Correlation Coefficient

Correlation coefficient is a dimensionless version of the covariance. Its magnitude conveys the strength of the relationship between two random variables.

$$\rho = \frac{\mathcal{COV}\{X,Y\}}{\sigma_x \sigma_y}$$

If X and Y are linearly dependent, then $\rho = \pm 1$. Also, if $\rho = \pm 1$, this is an indicator that the relation between X and Y is linear.

Example

Given a random variable Z with zero mean and variance unity, find the correlation coefficient of the random variable $X = Z - 1$ and $Y = Z + 1$