DISCRETE RANDOM VARIABLES

CHAPTER-3

CS6015-LINEAR ALGEBRA AND RANDOM PROCESSES

Probability mass function

- A random variable X is *discrete* if it takes values only in some countable set $\{x_1, x_2, \ldots\}$
- Its distribution function is represented as
$$F(x) = P(X \leq x)$$
- The **(probability) mass function** of a discrete random variable X is the function $f: \mathbb{R} \rightarrow [0,1]$ given by $\boldsymbol{f(x) = P(X = x)}$.
- The distribution and mass functions are related by
$$F(x) = \sum_{i: x_i \leq x} f(x_i)$$

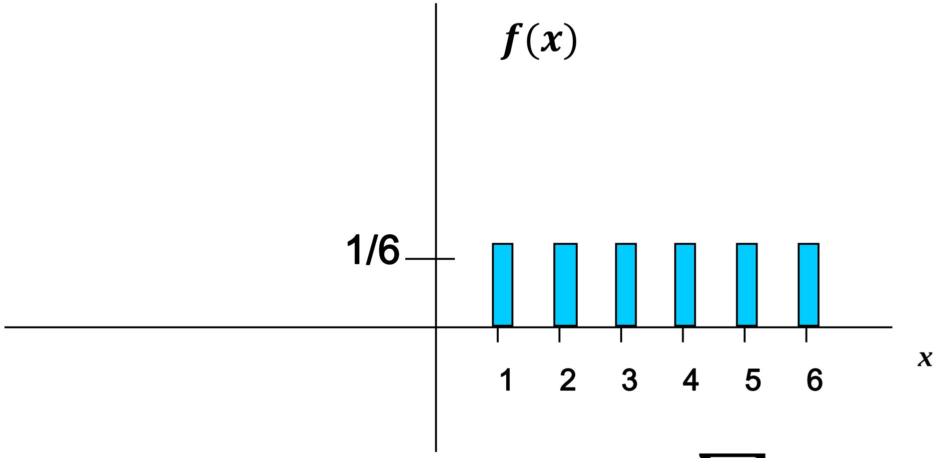
- The probability mass function $f: \mathbb{R} \to [0,1]$ satisfies:

  (a) the set of x *such that* $f(x) \neq 0$ is *countable,*

  (b) $\sum_i f(x_i) = 1$ where $x_1, x_2, \ldots$ are the values of x *such that*

  $$f(x) \neq 0.$$

Discrete example: roll of a die



$$\sum_{\text{all } x} f(x) = 1$$

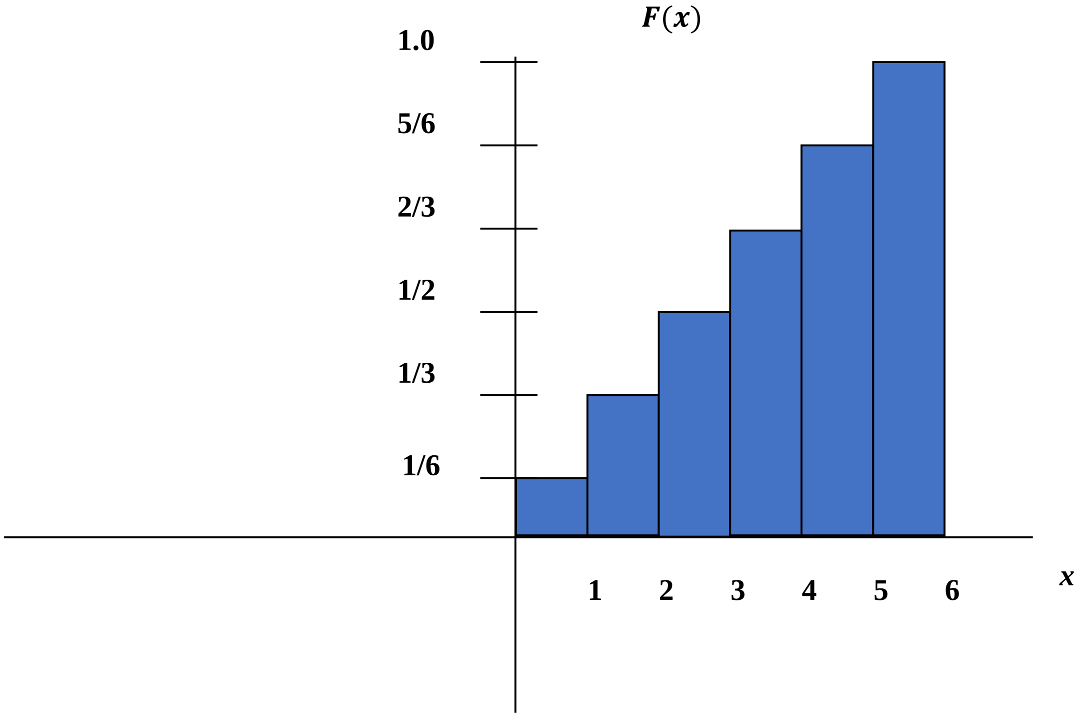
https://view.officeapps.live.com/op/view.aspx?src=https://web.stanford.edu/~kcobb/hrp259/lecture4.ppt

Probability mass function (pmf)

x	$f(x)$
1	$P(X = 1) = 1/6$
2	$P(X = 2) = 1/6$
3	$P(X = 3) = 1/6$
4	$P(X = 4) = 1/6$
5	$P(X = 5) = 1/6$
6	$P(X = 6) = 1/6$

1.0

Cumulative distribution function (CDF)



Cumulative distribution function (CDF)

x	$P(X \leq x)$
1	$P(X \leq 1) = 1/6$
2	$P(X \leq 2) = 2/6$
3	$P(X \leq 3) = 3/6$
4	$P(X \leq 4) = 4/6$
5	$P(X \leq 5) = 5/6$
6	$P(X \leq 6) = 6/6$ **=1**

Examples

1. What's the probability that you roll a 3 or less?

$$P(x \leq 3) = 1/2$$

2. What's the probability that you roll a 5 or higher?

$$P(x \geq 5) = 1 - P(x \leq 4) = 1 - 2/3 = 1/3$$

Practice Problem

Which of the following are probability functions?

a. $f(x) = .25\ for\ x = 9, 10, 11, 12$

b. $f(x) = \frac{3-x}{2}\ for\ \ x = 1, 2, 3, 4$

c. $f(x) = \frac{x^2+x+1}{25} for\ x = 0, 1, 2, 3$

Answer (a)

a. $f(x) = .25 \ \ for \ \ x = 9,10,11,12$

x	***f(x)***
9	**.25**
10	**.25**
11	**.25**
12	**<u>.25</u>**
	1.0

Yes, probability function!

Answer (b)

b. $f(x) = \frac{3-x}{2}\ for\ x = 1,2,3,4$

x	***f(x)***
1	**(3-1)/2=1.0**
2	**(3-2)/2=.5**
3	**(3-3)/2=0**
4	**(3-4)/2=-.5**

Though this sums to 1, you can't have a negative probability; therefore, it's not a probability function.

Answer (c)

c. $f(x) = \frac{x^2+x+1}{25}\ for\ x = 0,1,2,3$

x	f(x)
0	1/25
1	3/25
2	7/25
3	13/25

24/25

Doesn't sum to 1. Thus, it's not a probability function.

Practice Problem

- The number of ships to arrive at a harbor on any given day is a random variable represented by x. The probability distribution for x is:

x	10	11	12	13	14
P(x)	.4	.2	.2	.1	.1

Find the probability that on a given day:

a. exactly 14 ships arrive $\mathbf{P}(\boldsymbol{x} = \mathbf{14}) = .\mathbf{1}$

b. At least 12 ships arrive $\boldsymbol{P}(\boldsymbol{x} \geq \mathbf{12}) = (.\mathbf{2} + .\mathbf{1} + .\mathbf{1}) = .\mathbf{4}$

c. At most 11 ships arrive $\boldsymbol{P}(\boldsymbol{x} \leq \mathbf{11}) = (.\mathbf{4} + .\mathbf{2}) = .\mathbf{6}$

Practice Problem

You are lecturing to a group of 1000 students. You ask each of them to randomly pick an integer between 1 and 10. Assuming, their picks are truly random:

- What's your best guess for how many students picked the number 9?

  Since $\mathrm{P}(x = 9) = 1/10$, we'd expect about 1/10th of the 1000 students to pick 9.

  Answer: 100 students.

- What percentage of the students would you expect, picked a number less than or equal to 6?

  Since $\mathrm{P}(x \leq 6) = \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = .6$

  Answer: $\mathbf{60}\%$

Binomial Distribution

- A coin is tossed n times, and a head turns up each time with probability $p(= 1 - q)$.
- Then $\Omega = \{H, T\}^n$.
- The total number $\boldsymbol{X}$ of heads takes values in the set $\{0, 1, 2, \dots, n\}$ and is a discrete random variable.
- Its probability mass function $f(x) = P(X = x)$ satisfies $f(x) = 0 \; if \; x \notin \{0,1,2, \dots, n\}$

Binomial Distribution

- Let $0 \leq k \leq n$ and consider $f(k)$. Exactly $\binom{n}{k}$ points in Ω give a total of k heads; each of these points occurs with probability $p^k q^{n-k}$ and so

$$\boldsymbol{f(k) = \binom{n}{k} p^k q^{n-k} \text{ if } 0 \leq k \leq n}$$

- The random variable X is said to have the *binomial distribution* with parameters n and p, written as: bin(n, p). It is the sum $X = Y_1 + Y_2 + \ldots + Y_n$ of n Bernoulli variables.

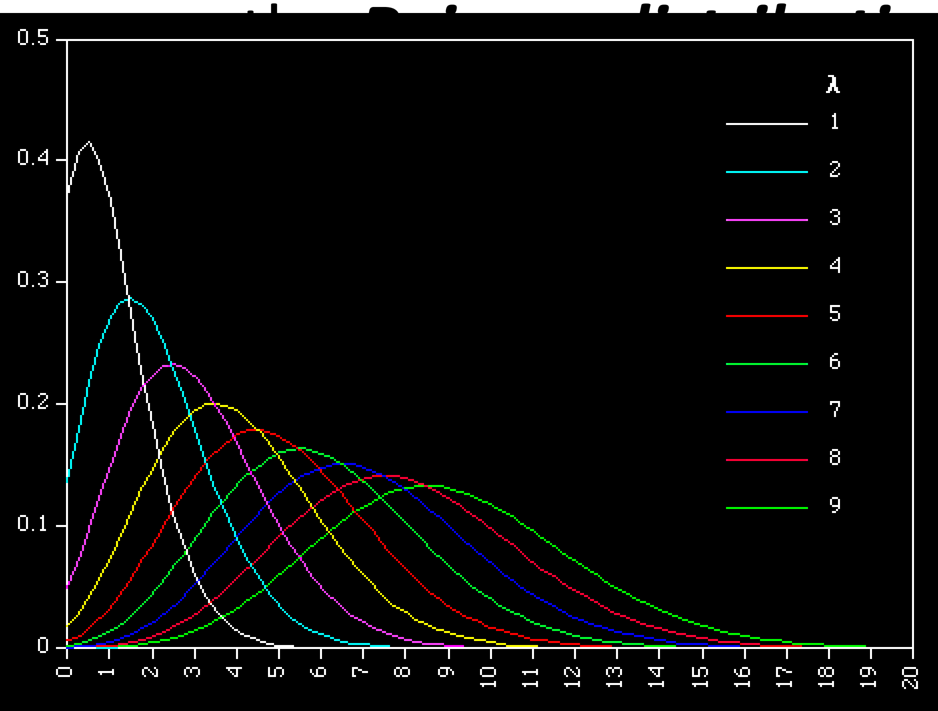
Poisson Distribution

- If a random variable $\boldsymbol{X}$ takes values in the set {0, 1, 2, ... } with mass function

$$\boldsymbol{f(k) = \frac{\lambda^k}{k!}e^{-\lambda}}$$

$\boldsymbol{k = 0, 1, 2, \dots}$ where $\lambda > 0$, then X is said to have

[illegible] with parameter λ.



Joint probability mass function

- Remember that for a discrete R.V. X, we define the PMF as $\mathrm{f(x)} = \mathrm{P_X(x)} = P(X = x)$.

- Now if we have two R.V.s X and Y and we would like to study them jointly, then we define the joint probability mass function as follows:

$$\boldsymbol{P_{XY}(x, y) = P(X = x, Y = y)}$$

- $P_X(x)$ and $P_Y(y)$ are called marginal PMFs.

Joint Distribution Function

- The joint distribution function $F : \mathbb{R}^2 \to [0,1]$ of X and Y , where X and Y are discrete variables is given by :

$$\boldsymbol{F(x, y) = P(X \leq x \; and \; Y \leq y)}$$

- The discrete random variables X and Y are independent if and only if

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

Independence

- Remember that events A and B are called **'independent'** if the occurrence of A does not change the subsequent probability of B occurring.

- More rigorously, A and B are independent if and only if $\boldsymbol{P(A \cap B) = P(A)P(B)}$.

- Discrete variables X and Y are independent if the events $\{X = x\}$ and $\{Y = y\}$ are independent for all x and y.

Independence

- Suppose X takes values in the set $\{x_1, x_2, \dots\}$ and Y takes values in the set $\{y_1, y_2, \dots\}$.

- Let $A_i = \{X = x_i\}$ and $B_j = \{Y = y_j\}$

The random variables X and Y are independent if and only if A_i and B_j are independent for all pairs i, j. A similar definition holds for collections $\{X_1, X_2, \dots, X_n\}$ of discrete variables.

Example: Poisson Flips

- A coin is tossed once and heads turns up with probability $p = 1 - q$. Let X and Y be the numbers of heads and tails respectively. It is no surprise that X and Y are not independent. After all,

$$P(X = Y = 1) = 0\,;\ \mathrm{P}(\mathrm{X} = 1)P(\mathrm{Y} = 1) = \mathrm{p}(1 - \mathrm{p})$$

- Suppose now that the coin is tossed a random N number of times, where N has the Poisson distribution with parameter λ.
- It is a remarkable fact that the resulting numbers X and Y of heads and tails *are* independent, since

$$P(X = x, Y = y)$$

$$= P(X = x, Y = y | N = x + y) P(N = x + y)$$

$$= \binom{x + y}{x} p^x q^y \cdot \frac{\lambda^{x+y}}{(x+y)!} e^{-\lambda}$$

$$= \frac{(x+y)!}{x!y!} p^x q^y \frac{\lambda^{x+y}}{(x+y)!} e^{-\lambda}$$

$$= \frac{(\boldsymbol{\lambda p})^{\boldsymbol{x}} (\boldsymbol{\lambda q})^{\boldsymbol{y}}}{\boldsymbol{x}!\boldsymbol{y}!} \boldsymbol{e}^{-\boldsymbol{\lambda}}$$

- $P(X = x) \quad = \sum_{n \geq x}^{\infty} P(X = x|N = n)P(N = n)$

$$= \sum_{n \geq x} P(X = x|N = n)P(N = n)$$

$$= \sum_{n \geq x} \binom{n}{x} p^x q^{n-x} \frac{\lambda^n e^{-\lambda}}{n!}$$

$$= \sum_{n \geq x} \frac{n!}{x!(n-x)!} \frac{(\lambda p)^x (\lambda q)^{n-x}}{n!} e^{-\lambda}$$

$$= \frac{(\lambda p)^x e^{-\lambda}}{x!} \sum_{n \geq x} \frac{(\lambda q)^{n-x}}{(n-x)!} = \frac{(\lambda p)^x e^{-\lambda} e^{\lambda q}}{x!}$$

$$= \boldsymbol{\frac{(\lambda p)^x e^{-\lambda p}}{x!}}$$

- Similarly $P(Y = y) = \boldsymbol{\frac{(\lambda q)^y e^{-\lambda q}}{y!}}$

- $P(X = x, Y = y) = \frac{(\lambda p)^x (\lambda q)^y}{x! y!} e^{-\lambda}$

$$= \frac{(\lambda p)^x e^{-\lambda p}}{x!} \frac{(\lambda q)^y e^{-\lambda q}}{y!}$$

$$= P(X = x)P(Y = y) = \frac{(\lambda p)^x (\lambda q)^y}{x! y!} e^{-\lambda}$$

This shows that X and Y are independent.

- If X and Y are independent and $g, h : \mathbb{R} \to \mathbb{R}$, then $g(X)$ and $h(Y)$ are independent also.

- $\{X_i : i \in I\}$ is an independent family if and only if

$$P(X_i = x_i \; for \; all \; i \in J) = \prod_{i \in J} P(X_i = x_i)$$

Example

- Consider two random variables X and Y with joint PMF as given in the table:

	Y=2	Y=4	Y=5
X=1	1/12	1/24	1/24
X=2	1/6	1/12	1/8
X=3	1/4	1/8	1/12

a. Find $P(X \leq 2, Y \leq 4)$
b. Find the marginal PMFs of X and Y
c. Are X and Y independent?

https://www.probabilitycourse.com/chapter5/5_1_6_solved_prob.php

a. To find $P(X \leq 2, Y \leq 4)$, we can write :

$$P(X \leq 2, Y \leq 4) = P_{XY}(1,2) + P_{XY}(1,4) + P_{XY}(2,2) + P_{XY}(2,4)$$
$$= \frac{1}{12} + \frac{1}{24} + \frac{1}{6} + \frac{1}{12} = \mathbf{\frac{3}{8}}$$

b. $$P_X(x) = \begin{cases} \frac{1}{6} & x = 1 \\ \frac{3}{8} & x = 2 \\ \frac{11}{24} & x = 3 \\ 0 & o/w \end{cases}$$

$$P_Y(y) = \begin{cases} \frac{1}{2} & y = 2 \\ \frac{1}{4} & y = 4 \\ \frac{1}{4} & y = 5 \\ 0 & o/w \end{cases}$$

c. To check if X and Y are independent, we need to check if $P(X = x, Y = y) = P(X = x)P(Y = y)$.

Looking at the table and results from the previous parts, we get:

$$P(X = 2, Y = 2) = \frac{1}{6} \neq P(X = 2)P(Y = 2) = \frac{3}{16}$$

Thus **X and Y are not independent.**

Expectation

- The **mean value, or expectation, or expected value** of the random variable X with mass function f is defined to be:

$$E(X) = \sum_{x:f(x)>0} xf(x)$$

- Example: $x \rightarrow -2, -1, 1, 3$

$$f(x) \rightarrow \frac{1}{4}, \quad \frac{1}{8}, \quad \frac{1}{4}, \quad \frac{3}{8}$$

$$E(X) = -2\left(\frac{1}{4}\right) + (-1)\left(\frac{1}{8}\right) + 1\left(\frac{1}{4}\right) + 3\left(\frac{3}{8}\right) = \frac{3}{4}$$

- The random variable $Y = X^2$ takes values 1,4,9 with probabilities $\frac{3}{8}, \frac{1}{4}, \frac{3}{8}$ respectively

- So, $\boldsymbol{E}(\boldsymbol{Y}) = 1\left(\frac{3}{8}\right) + 4\left(\frac{1}{4}\right) + 9\left(\frac{3}{8}\right) = \frac{19}{4}$

- Alternatively, $\boldsymbol{E}\left(\boldsymbol{X}^{\mathbf{2}}\right) = \sum_x x^2 P(X = x)$

$$= 4\left(\frac{1}{4}\right) + 1\left(\frac{1}{8}\right) + 1\left(\frac{1}{4}\right) + 9\left(\frac{3}{8}\right) = \frac{19}{4}$$

- Lemma: If X has mass function f and $g: \mathbb{R} \to \mathbb{R}$, then

$$E\big(g(X)\big) = \sum_x g(x) f(x)$$

- If k is a positive integer, the k^{th} moment m_k of X is defined to be $\boldsymbol{m_k = E(X^k)}$.

- The k^{th} central moment $\sigma_k = E((X - m_1)^k)$

- The two moments of most use are $\boldsymbol{m_1 = E(X) = \mu}$ and $\boldsymbol{\sigma_2 = E\left((X - EX)^2\right) = var(X)}$ called the **mean** (or expectation) and **variance** respectively.

- $\sigma = \sqrt{var(X)}$ is called the **standard deviation**.

- $\sigma_2 = E\left((X - m_1)^2\right)$
  $= \sum_x (x - m_1)^2 f(x)$
  $= \sum_x x^2 f(x) - 2m_1 \sum_x x f(x) + m_1^2 \sum_x f(x)$
  $= m_2 - m_1^2$

- Thus $\boldsymbol{var(X) = E\left((X - EX)^2\right) = E\left(X^2\right) - (EX)^2}$

Indicator variable

- Expected value of an indicator variable is just the probability of that event. (Remember that a random variable I_A is the indicator random variable for event A, if $I_A = 1$ when A occurs and $I_A = 0$ otherwise.)

- i.e., $\boldsymbol{E(I_A) = P(A)}$

- **Proof**: $E(I_A) = 1. P(I_A = 1) + 0. P(I_A = 0)$
$$= P(I_A = 1) = P(A)$$

Bernoulli Variable

- Let X be a Bernoulli variable taking the value 1 with probability $p(= 1 - q)$.

- Then $E(X) = \sum_x xf(x) = 0.q + 1.p = p$

- $E(X^2) = \sum_x x^2 f(x) = 0.q + 1.p = p$

- So $\boldsymbol{var}(\boldsymbol{X}) = E(X^2) - (EX)^2 = p - p^2$
$$= p(1 - p) = \boldsymbol{pq}$$

Binomial Variable

- $\boldsymbol{EX} \ = \sum_{k=0}^{n} kf(k)$
  $= \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k}$
  $= np\sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!(k-1)!} p^{k-1} q^{n-k}$
  $= np\sum_{l=0}^{n-1} \binom{n-1}{l} p^l q^{n-1-l} = \mathbf{np}$

- Likewise, $\boldsymbol{var(X) = npq}$

Theorem

- The expectation operator E has the following properties :

a. If $X \geq 0$ then $E(X) \geq 0$

b. If $a, b \in \mathbb{R}$ then
$$E(aX + bY) = aE(X) + bE(Y)$$

c. The random variable 1, taking the value 1 always, has expectation $E(1) = 1$

If $a, b \in \mathbb{R}$ then $E(aX + bY) = aE(X) + bE(Y)$

- **Proof** :

Let $A_x = \{X = x\}, B_y = \{Y = y\}$. Then

$$aX + bY = \sum_{x,y} (aX + bY) I_{A_x \cap B_y}$$

$E(aX + bY) = \sum_{x,y} (aX + bY) P(A_x \cap B_y)$ Since $\boldsymbol{E(I_A) = P(A)}$

$$\text{However, } \sum_y P(A_x \cap B_y) = P\left(A_x \cap \left(\bigcup_y B_y\right)\right)$$
$$= P(A_x \cap \Omega) = P(A_x)$$

Likewise, $\sum_x P(A_x \cap B_y) = P(B_y)$

- This give

$$E(aX + bY)$$
$$= \sum_x ax \sum_y P(A_x \cap B_y) + \sum_y by \sum_x P(A_x \cap B_y)$$

$$= a \sum_x xP(A_x) + b \sum_y yP(B_y)$$

$$= \boldsymbol{aE(X) + bE(Y)}$$

Remark: It is not in general true that $E(XY)$ is the same as $E(X)E(Y)$.

Lemma

- If X and Y are independent then $\boldsymbol{E(XY) = E(X)E(Y)}$.

Proof :

$$XY = \sum_{x,y} xy\, I_{A_x \cap B_y}$$

$E(XY) = \sum_{x,y} xyP(A_x)P(B_y)$ **by independence**

$$= \sum_x xP(A_x) \sum_y yP(B_y)$$

$$= E(X)E(Y)$$

- ***X*** and ***Y*** are called **uncorrelated** if $E(XY) = E(X)E(Y)$.
- Independent variables are uncorrelated. The converse is not true.
- **Theorem :**

a. $var(aX) = a^2 var(X)\ for\ a \in \mathbb{R}$

b. $var(X+Y) = var(X) + var(Y)\ if\ X\ and\ Y\ are\ uncorrelated$

Proof :

a. $var(aX) = E\left((aX - EaX)^2\right) = E(a^2(X - EX)^2)$

$$= a^2 E\left((X - EX)^2\right) = a^2 var(X)$$

b. $var(X+Y) = E\{\left(X + Y - E(X+Y)\right)^2\}$

$$= E\left[(X - EX)^2 + 2\left(XY - E(X)E(Y)\right) + (Y - EY)^2\right]$$

$$= var(X) + 2\left(E(XY) - E(X)E(Y)\right) + var(Y)$$

$$= var(X) + var(Y)$$

{Since $E(XY) = E(X)E(Y)$ if X and Y are uncorrelated}

Covariance

- The covariance of X and Y is
$$cov(X, Y) = E[(X - EX)(Y - EY)]$$
- The correlation coefficient of X and Y is
$$\rho(X, Y) = \frac{cov(X,Y)}{\sqrt{var(X).var(Y)}}$$

As long as the variances are non-zero.

- $cov(X, X) = var(X)$
- $cov(X, Y) = E(XY) - E(X)E(Y)$
- X and Y are uncorrelated if $cov(X, Y) = 0$
- Also, independent variables are always uncorrelated, although the converse is not true.

Cauchy-Schwarz inequality

- For random variables X and Y,

$$\boldsymbol{E}\{(\boldsymbol{XY})\}^{\boldsymbol{2}} \leq \boldsymbol{E}\left(\boldsymbol{X}^{\boldsymbol{2}}\right)\boldsymbol{E}(\boldsymbol{Y}^{\boldsymbol{2}})$$

with equality if and only if $\mathrm{P(aX = bY)} = 1$ for some real a and b , at least one of which is non-zero.

- $\rho = +1$ if and only if Y *increases* linearly with *X* and $\rho = -1$ if and only if Y *decreases* linearly as X increases.

Example

- Let ***X*** and ***Y*** take values in { 1, 2, 3} and {-1, 0, 2} respectively, with joint mass function ***f*** where *f(x, y)* is the appropriate entry in Table :

	Y=-1	Y=0	Y=2	f_x
X=1	1/18	3/18	2/18	6/18
X=2	2/18	0	3/18	5/18
X=3	0	4/18	3/18	7/18
f_y	3/18	7/18	8/18	

$$E(XY) = \sum_{x,y} xyf(x,y) = \qquad ; \ E(X) = \sum_{x} xf(x) =$$

$$E(Y) =$$

- $E(XY) = \sum_{x,y} xyf(x,y) = 29/18$
- $E(X) = \sum_x xf(x) = 37/18$
- $E(Y) = 13/18$

- $var(X) = E(X^2) - E(X)^2 = 233/324$
- $var(Y) = 461/324$
- $cov(X,Y) = \frac{41}{324}$
- $\rho(X,Y) = 41/\sqrt{107413}$

Conditional Distribution Function

- The conditional distribution function of Y given $X = x$, written $\boldsymbol{F}_{(\boldsymbol{Y}|\boldsymbol{X})}(.\,|\mathbf{x})$ is defined by:

$$\boldsymbol{F}_{(\boldsymbol{Y}|\boldsymbol{X})}(\boldsymbol{y}|\boldsymbol{x}) = \boldsymbol{P}(\boldsymbol{Y} \leq \boldsymbol{y}|\boldsymbol{X} = \boldsymbol{x})$$

for any x such that $P(X = x) > 0$

Conditional (probability) mass function

- The conditional (probability) mass function of Y given $X = x$, written $f_{(Y|X)}(.|x)$, is defined by

$$\boldsymbol{f_{(Y|X)}(y|x) = P(Y = y|X = x)}$$

- $\boldsymbol{f_{(Y|X)} = \frac{f_{X,Y}}{f_X}}$

for any x such that $P(X = x) > 0$

- Clearly $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent if and only if

$$f_{(Y|X)} = f_Y$$

Conditional Expectation

- Let $\psi(x) = E(Y|X = x)$. Then $\psi(X)$ is called the **conditional expectation** of Y given X, written as $\boldsymbol{E(Y|X)}$.

- Although 'conditional expectation' sounds like a number, **it is actually a random variable**.

- $\psi(X) = \sum_y y f_{(Y|X)}(y|x)$

- **Theorem** : The conditional expectation $\psi(X) = E(Y|X)$ satisfies

$$\boldsymbol{E\big(\psi(X)\big) = E(Y)}$$

Proof :

$$E\big(\psi(X)\big) = \sum_{x} \psi(x) f_X(x) = \sum_{x,y} y f_{(Y|X)}(y|x)\, f_X(x)$$

$$= \sum_{x,y} y f_{X,Y}(x,y) = \sum_{y} y\, f_Y(y) = \boldsymbol{E(Y)}$$

- Thus, $\boldsymbol{E(Y) = \sum_x E(Y|X = x)\, P(X = x)}$

Theorem

- The conditional expectation $\psi(X) = E(Y|X)$ satisfies

$$E\big(\psi(X)g(X)\big) = E(Yg(X))$$

For any function g for which both expectations exist.

Proof : $E\big(\psi(X)g(X)\big) = \sum_x \psi(x)g(x)\, f_X(x)$

$$= \sum_{x,y} yg(x)\, f_{(Y|X)}(y|x) f_X(x)$$

$$= \sum_{x,y} yg(x)\, f_{X,Y}(x,y)$$

$$= \boldsymbol{E(Yg(X))}$$

Example problem

- To transmit message i using an optical communication system.
- When light of intensity λ_i strikes the photodetector, the number of photoelectrons generated is a Poisson (λ_i) random variable.
- Find the conditional probability that the number of photoelectrons observed at the photodetector is less than 2 given that message i was sent.

https://www.cs.ccu.edu.tw/~wylin/probability/prob_ch3.pdf

Solution

- Let X denote the message to be sent, and let Y denote the number of photoelectrons generated by the photodetector.

- The problem statement is telling us that

$$P(Y = n|X = i) = \frac{\lambda_i^n e^{-\lambda_i}}{n!}, n = 0,1,2, \dots$$

- The conditional probability to be calculated is

$$\begin{aligned} P(Y < 2|X = i) &= P(Y = 0 \; or \; Y = 1|X = i) \\ &= P(Y = 0|X = i) + P(Y = 1|X = i) \\ &= \boldsymbol{e^{-\lambda_i} + \lambda_i e^{-\lambda_i}} \end{aligned}$$

Sums of Random Variables

- $Z = X + Y$

- Given joint pmf $f(x, y)$, we want to find pmf of Z.

- $\{X + Y = Z\} = \bigcup_x(\{X = x\} \cap \{Y = z - x\})$

$$
\begin{aligned}
f_{X+Y}(z) &= P(X + Y = z) \\
&= P(\bigcup_x \{X = x\} \cap \{Y = z - x\}) \\
&= \sum_x P(X = x, \mathrm{Y} = \mathrm{z} - \mathrm{x}) \\
&= \sum_x f(x, z - x)
\end{aligned}
$$

- If X and Y are independent, then

$$P(X+Y=z) = f_{X+Y}(z)$$
$$= \sum_x f_X(x) f_Y(z-x)$$
$$= \sum_y f_X(z-y) f_Y(y)$$

- The mass function of $X+Y$ is called the ***convolution*** of the mass functions of X and Y, and is written

$$f_{X+Y} = f_X * f_Y$$