Two discrete random variables

Recall for a discrete random variable X, the pmf for $X = \{x_1, x_2, \cdots, x_n\}$ satisfies

- $f(x_i) \geq 0$
- $\sum_{i=1}^{n} f(x_i) = 1$
- $f(x_i) = P(X = x_i)$

Now consider two random variables, X, Y, their joint pmf satisfies the following conditions

- $f_{X,Y}(x, y) \geq 0$ for all x, y.
- $\sum_x \sum_y f_{X,Y}(x, y) = 1.$
- $$f_{X,Y}(x, y) = P(X = x \text{ and } Y = y) = P(X = x, Y = y).$$

Example 1: Suppose you toss a coin 2 times. Let random variable X

$$X = \begin{cases} 1 & \text{if first toss is a head} \\ 0 & \text{if first toss is a tail} \end{cases}$$

and let Y be the random variable with number of heads in the 2 tosses.

- possible values of $X = \{0, 1\}$, $Y = \{0, 1, 2\}$

		Y		
		0	1	2
X	0	f(0,0)	f(0,1)	f(0,2)
	1	f(1,0)	f(1,1)	f(1,2)

$f(0,0) = P(X = 0, Y = 0) =$
$P(\text{1st toss is a tail, no heads in two tosses}) = P(TT) = 1/4$

- $f(0,0) = P(X = 0, Y = 0) =$
  $P(\text{1st toss is a tail, no heads in two tosses}) = P(TT) = 1/4$
- $f(0,1) = P(X = 0, Y = 1) =$
  $P(\text{1st toss is a tail, 1 heads in two tosses}) = P(TH) = 1/4$
- $f(0,2) = P(X = 0, Y = 2) =$
  $P(\text{1st toss is a tail, 2 heads in two tosses}) = 0$
- $f(1,0) = P(X = 1, Y = 0) =$
  $P(\text{1st toss is a head, no heads in two tosses}) = 0$
- $f(1,1) = P(X = 1, Y = 1) =$
  $P(\text{1st toss is a head, 1 heads in two tosses}) = P(HT) = 1/4$
- $f(1,2) = P(X = 1, Y = 2) =$
  $P(\text{1st toss is a head, 2 heads in two tosses}) = P(HH) = 1/4$

		Y		
		0	1	2
X	0	1/4	1/4	0
	1	0	1/4	1/4

$$\sum_{x=0}^{x=1} \sum_{y=0}^{y=2} f(x, y) = 1$$

Marginal pmfs

The individual probability distribution of a random variable is referred to as its marginal probability distribution.

- In general, the marginal probability distribution of X can be determined from the joint probability distribution of X and other random variables.
  —— For example, to determine $P(X = x)$, we sum $P(X = x, Y = y)$ over all points in the range of (X, Y) for which $X = x$.

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

- to determine $P(Y = y)$, we sum $P(X = x, Y = y)$ over all points in the range of (X, Y) for which $Y = y$.

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

Example 2: Based on the joint mass function of X and Y in example 1, find the marginal pmfs of X and Y.

		Y			
		0	1	2	$f_X(x)$
X	0	1/4	1/4	0	1/2
	1	0	1/4	1/4	1/2
	$f_Y(y)$	1/4	1/2	1/4	

$$f_X(0) = P(X = 0) = f(0,0) + f(0,1) + f(0,2) = 1/2$$

$$f_X(1) = P(X = 1) = f(1,0) + f(1,1) + f(1,2) = 1/2$$

$$f_X(0) + f_X(1) = 1/2 + 1/2 = 1$$

$$f_Y(0) = P(Y = 0) = f(0,0) + f(1,0) = 1/4$$

$$f_Y(1) = P(Y = 1) = f(0,1) + f(1,1) = 1/2$$

$$f_Y(2) = P(Y = 2) = f(0,2) + f(1,2) = 1/4$$

$$f_Y(0) + f_Y(1) + f_Y(2) = 1$$

Practice: a joint distribution function is defined as follows:

x	y	$f_{X,Y}(x,y)$
1	1	1/4
1.5	2	1/8
1.5	3	1/4
2.5	4	2k
3.0	5	k

Determine the following:
(a) Find k so that the joint function $f_{X,Y}(x,y)$ satisfies the properties of a joint probability mass function.
(b) Marginal probability mass function of X and Y.
(c) Find $P(X > 1.8, Y > 4.7)$.

Covariance and Correlation

Expected value of a function of two random variables:

$$\mu_X = E[X] = \sum_x \sum_y x f_{X,Y}(x,y) = \sum_x x f_X(x)$$

$$\mu_Y = E[Y] = \sum_y \sum_x y f_{X,Y}(x,y) = \sum_y y f_Y(y)$$

$$\sigma_X = V(X) = E(X^2) - E^2(X), \sigma_Y = V(Y) = E(Y^2) - E^2(Y)$$

Covariance: the covariance between two random variables X and Y, denoted as $cov(X, Y)$ or $\sigma_{X,Y}$, is

$$\sigma_{X,Y} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y$$

———Note: covariance is a measure of linear relationship between random variables.

———

$$E(XY) = \sum_x \sum_y xy f_{X,Y}(x,y)$$

Correlation: the correlation between two random variables X and Y, denoted as ρ_{XY}, is

$$\rho_{XY} = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

For any two random variable X and Y, $-1 \leq \rho_{X,Y} \leq 1$.

Example 3:

Find $E(X), E(Y), V(X), V(Y), E(XY), cov(X,Y)$ and $\rho_{X,Y}$ of the following joint distribution function.

		Y			
		0	1	2	$f_X(x)$
X	0	1/4	1/4	0	1/2
	1	0	1/4	1/4	1/2
	$f_Y(y)$	1/4	1/2	1/4	

$$
\begin{aligned}
E(XY) &= \sum_x \sum_y xyf_{X,Y}(x,y) \\
&= 0*0*f(0,0) + 0*1*f(0,1) + 0*2*f(0,2) \\
&+ 1*0*f(1,0) + 1*1*f(1,1) + 1*2*f(1,2) \\
&= \frac{1}{4} + 2*\frac{1}{4} = \frac{3}{4}
\end{aligned}
$$

Recall Example 1: Suppose you toss a coin 2 times. Let random variable X

$$X = \begin{cases} 1 & \text{if first toss is a head} \\ 0 & \text{if first toss is a tail} \end{cases}$$

and let Y be the random variable with number of heads in the 2 tosses.

- X is a bernoulli distribution with $p = 1/2$

$$E(X) = 1/2, V(X) = 1/2 * (1 - 1/2) = 1/4$$

  or using joint distribution function table,

$$E(X) = 0 * 1/2 + 1 * 1/2 = 1/2$$

- Y is a binomial distribution with $n = 2, p = 1/2$

$$E(Y) = 2 * 1/2 = 1, V(Y) = 2 * 1/2 * (1 - 1/2) = 1/2$$

- $cov(X, Y) = E(XY) - E(X)E(Y) = \frac{3}{4} - \frac{1}{2} * 1 = \frac{1}{4}$
- $\rho_{XY} = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}} = \frac{\frac{1}{4}}{\sqrt{1/4} * \sqrt{1/2}} = \sqrt{2}/2 = 0.707$

Practice: suppose r.v X and Y has the following joint distribution

		Y			
		7	9	11	$f_X(x)$
X	1	0.2	0	0	0.2
	2	0	0.6	0	0.6
	3	0	0	0.2	0.2
	$f_Y(y)$	0.2	0.6	0.2	

Find correlation of X and Y.

5-1 Two Discrete Random Variables

5-1.3 Conditional Probability Distributions

Given discrete random variables X and Y with joint probability mass function $f_{XY}(x, y)$ the **conditional probability mass function** of Y given $X = x$ is

$$f_{Y|x}(y) = f_{XY}(x, y)/f_X(x) \qquad \text{for } f_X(x) > 0 \tag{5-3}$$

5-1 Two Discrete Random Variables

5-1.3 Conditional Probability Distributions

Because a conditional probability mass function $f_{Y|x}(y)$ is a probability mass function for all y in R_x, the following properties are satisfied:

(1) $f_{Y|x}(y) \geq 0$

(2) $\sum_y f_{Y|x}(y) = 1$

(3) $P(Y = y | X = x) = f_{Y|x}(y)$ (5-4)

5-1 Two Discrete Random Variables

Definition: Conditional Mean and Variance

The **conditional mean** of Y given $X = x$, denoted as $E(Y|x)$ or $\mu_{Y|x}$, is

$$E(Y|x) = \sum_y y f_{Y|x}(y) \tag{5-5}$$

and the **conditional variance** of Y given $X = x$, denoted as $V(Y|x)$ or $\sigma^2_{Y|x}$, is

$$V(Y|x) = \sum_y (y - \mu_{Y|x})^2 f_{Y|x}(y) = \sum_y y^2 f_{Y|x}(y) - \mu^2_{Y|x}$$

5-1 Two Discrete Random Variables

5-1.4 Independence

For discrete random variables X and Y, if any one of the following properties is true, the others are also true, and X and Y are **independent.**

(1) $f_{XY}(x, y) = f_X(x) f_Y(y)$ for all x and y

(2) $f_{Y|x}(y) = f_Y(y)$ for all x and y with $f_X(x) > 0$

(3) $f_{X|y}(x) = f_X(x)$ for all x and y with $f_Y(y) > 0$

(4) $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for any sets A and B in the range of X and Y, respectively. (5-6)

Example, practice problem continued

r.vs X and Y has the following joint distribution

x	y	$f_{X,Y}(x,y)$
1	1	$1/4$
1.5	2	$1/8$
1.5	3	$1/4$
2.5	4	$1/4$
3.0	5	$1/8$

- find conditional probability distribution of Y given $X = 1.5$
- find conditional probability distribution of X given $Y = 2$
- find $E[Y|X = 1.5]$
- are X and Y independent?
- find $cov(X, Y)$

Recall binomial distribution

- 1st trial, Fail or Success, 2nd trial Fail or Success, ...nth trial Fail or Success
- Trials are independent
- Let X be number of successes, $X \sim Bin(n, p)$
  ——p is the probability of success

$$P(X = x) = \left(\begin{array}{c} n \\ x \end{array} \right) p^x (1 - p)^{n-x}$$

Now with more possible outcomes

Example: Digital channel of 20 bits received.

- Assume that each individual bit is classified as excellent (E), good (G), fair (F) or poor (P) with probabilities of 0.6, 0.3, 0.08 and 0.02 respectively.
- Assume each individual bits are independent.
- What is the probability that among 20 bits received, 14 are excellent, 3 are good, 2 are fair and 1 is poor?

$$P(14E, 3G, 2F, 1P) = \frac{20!}{14!3!2!1!}0.6^{14}0.3^{3}0.08^{2}0.02^{1} = 0.0063$$

5-1 Two Discrete Random Variables

5-1.6 Multinomial Probability Distribution

Suppose a random experiment consists of a series of n trials. Assume that

(1) The result of each trial is classified into one of k classes.

(2) The probability of a trial generating a result in class 1, class 2, … , class k is constant over the trials and equal to $p_1, p_2, \ldots, p_k$, respectively.

(3) The trials are independent.

The random variables $X_1, X_2, \ldots, X_k$ that denote the number of trials that result in class 1, class 2, … , class k, respectively, have a **multinomial distribution** and the joint probability mass function is

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k} \qquad (5\text{-}12)$$

for $x_1 + x_2 + \cdots + x_k = n$ and $p_1 + p_2 + \cdots + p_k = 1$.

5-1 Two Discrete Random Variables

5-1.6 Multinomial Probability Distribution

Each trial in a multinomial random experiment can be regarded as either generating or not generating a result in class i, for each $i = 1, 2, \ldots, k$. Because the random variable X_i is the number of trials that result in class i, X_i has a binomial distribution.

If $X_1, X_2, \ldots, X_k$ have a multinomial distribution, the marginal probability distribution of X_i is binomial with

$$E(X_i) = np_i \quad \text{and} \quad V(X_i) = np_i(1 - p_i) \tag{5-13}$$

Example continued:

(1) marginal distribution of X_2 (Good bits) is binomial with $n = 20$ and $p = 0.3$

$$E(X_2) = 20 * 0.3 = 6, V(X) = 20 * 0.3 * (1 - 0.3) = 4.2$$

(2)

$$P(X_2 = 6, X_3 = 2) = \frac{20!}{6!2!12!} 0.3^6 0.08^2 0.62^{12}$$

(3)

$$\begin{aligned} P(X_1 = 12 | X_2 = 6) &= \frac{P(X_1 = 12, X_2 = 6)}{P(X_2 = 6)} \\ &= \frac{\frac{20!}{12!6!2!} 0.6^{12} 0.3^6 0.1^2}{\frac{20!}{6!14!} 0.3^6 0.7^{14}} \\ &= 1.44 * 10^{-6} \end{aligned}$$

5-5 Linear Combinations of Random Variables

Definition

Given random variables $X_1, X_2, \ldots, X_p$ and constants $c_1, c_2, \ldots, c_p$,

$$Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p \tag{5-34}$$

is a **linear combination** of $X_1, X_2, \ldots, X_p$.

Mean of a Linear Combination

If $Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$,

$$E(Y) = c_1E(X_1) + c_2E(X_2) + \cdots + c_pE(X_p) \tag{5-35}$$

5-5 Linear Combinations of Random Variables

Variance of a Linear Combination

If $X_1, X_2, \ldots, X_p$ are random variables, and $Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$, then in general

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p) + 2\sum_{i<j}\sum c_ic_j \operatorname{cov}(X_i, X_j) \quad (5\text{-}36)$$

If $X_1, X_2, \ldots, X_p$ are **independent**,

$$V(Y) = c_1^2V(X_1) + c_2^2V(X_2) + \cdots + c_p^2V(X_p) \quad (5\text{-}37)$$

5-5 Linear Combinations of Random Variables

Example 5-33

An important use of equation 5-37 is in **error propagation** that is presented in the following example.

A semiconductor product consists of three layers. If the variances in thickness of the first, second, and third layers are 25, 40, and 30 nanometers squared, what is the variance of the thickness of the final product.

Let X_1, X_2, X_3, and X be random variables that denote the thickness of the respective layers, and the final product. Then

$$X = X_1 + X_2 + X_3$$

The variance of X is obtained from equaion 5-39

$$V(X) = V(X_1) + V(X_2) + V(X_3) = 25 + 40 + 30 = 95 \text{ nm}^2$$

Consequently, the standard deviation of thickness of the final product is $95^{1/2} = 9.75$ nm and this shows how the variation in each layer is propagated to the final product.

- Independent $\rightarrow cov(X, Y) = 0$
- but $cov(X, Y) = 0$ doesn't imply independence of X and Y

Example:

x	y	$f_{X,Y}(x, y)$
-2	4	1/5
-1	1	1/5
0	0	1/5
1	1	1/5
2	4	1/5

$$cov(X, Y) = 0,$$

but

$$f_{XY}(x, y) \neq f_X(x) f_Y(y)$$

5-5 Linear Combinations of Random Variables

Mean and Variance of an Average

If $\overline{X} = (X_1 + X_2 + \cdots + X_p)/p$ with $E(X_i) = \mu$ for $i = 1, 2, \ldots, p$

$$E(\overline{X}) = \mu \tag{5-38a}$$

if $X_1, X_2, \ldots, X_p$ are also independent with $V(X_i) = \sigma^2$ for $i = 1, 2, \ldots, p$,

$$V(\overline{X}) = \frac{\sigma^2}{p} \tag{5-38b}$$

5-5 Linear Combinations of Random Variables

Reproductive Property of the Normal Distribution

If $X_1, X_2, \ldots, X_p$ are independent, normal random variables with $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$, for $i = 1, 2, \ldots, p$,

$$Y = c_1X_1 + c_2X_2 + \cdots + c_pX_p$$

is a normal random variable with

$$E(Y) = c_1\mu_1 + c_2\mu_2 + \cdots + c_p\mu_p$$

and

$$V(Y) = c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + \cdots + c_p^2\sigma_p^2 \qquad (5\text{-}39)$$

5-5 Linear Combinations of Random Variables

Example 5-34

Let the random variables X_1 and X_2 denote the length and width, respectively, of a manufactured part. Assume that X_1 is normal with $E(X_1) = 2$ centimeters and standard deviation 0.1 centimeter and that X_2 is normal with $E(X_2) = 5$ centimeters and standard deviation 0.2 centimeter. Also, assume that X_1 and X_2 are independent. Determine the probability that the perimeter exceeds 14.5 centimeters.

Then, $Y = 2X_1 + 2X_2$ is a normal random variable that represents the perimeter of the part. We obtain, $E(Y) = 14$ centimeters and the variance of Y is

$$V(Y) = 4 \times 0.1^2 + 4 \times 0.2^2 = 0.2$$

Now,

$$P(Y > 14.5) = P[(Y - \mu_Y)/\sigma_Y > (14.5 - 14)/\sqrt{0.2}]$$
$$= P(Z > 1.12) = 0.13$$

Practice: soft-drink cans are filled by an automated filling machine and the standard deviation is 0.5 fluid ounce. Assume that the fill volumes of the cans are independent normal r.vs.

(a) what is the standard deviation of the average fill volume of 100 cans?

(b) if the mean fill volume is 12.1 ounces, what is the probability that the average fill volume of the 100 cans is below 12 ounces?

(c) what should the mean fill volume equal so that the probability that the average of 100 cans is below 12 fluid ounces is 0.005?