UNIVERSITY OF LINCOLN
SCHOOL OF MATHEMATICS AND PHYSICS

MTH1005M PROBABILITY AND STATISTICS PRACTICAL 6

Review

The expectation of a joint discrete random variable is defined as

$$E[X,Y] = \sum_{y}\sum_{x} xyp(x,y)$$

The covariance of a discrete random variable is defined as

$$\begin{aligned} Cov(X_1,X_2) &= E\left[(X_1 - E[X_1])(X_2 - E[X_2])\right] \\ &= E[X_1X_2] - E[X_1]E[X_2] \\ &= \sum_{all x_1}\sum_{all x_2}(x_1-\mu_{X_1})(x_2-\mu_{X_2})p(x_1,x_2) \end{aligned}$$

In general it can be shown that a positive Cov(X,Y) is an indication that Y increases when X does. A negative Cov(X,Y) is an indication that Y decreases when X increases. A better indicator than the covariance is **the correlation** that is defined as

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

It can be shown that the correlation will always have a value betwen -1 and $+1$. The significance of the correlation is similar to just discussed for the covariance:

- Positive correlation between two variables means that X increases as Y increases.

- Negative correlation means that X decreases as Y increases.

The joint probability density function $f_{XY}(x,y)$ of a continuous random variables (X,Y) is a function whose integral in the set of possible values for (X, Y) is the rectangle $D = (x,y): a \leq x \leq b, u \leq y \leq w.$ gives the likelihood of the subset in the sample space of (X,Y) :

$$P\{a \leq X \leq b; u \leq Y \leq w\} = \int_a^b \int_u^w f_{XY}(x,y)dxdy$$

To be a valid probability density function of a random variable, we must have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1 \quad (0.1)$$

MEASURES OF CENTRAL TENDENCY

- **Mean.** For a discrete data set $\{x_i \in X, i = 1, \ldots N\}$, the arithmetic mean, also called the mathematical expectation or average, is the central value of the numbers (x_i) in the the set: specifically, the sum of the values divided by the number of values (N)

  $$\mu = \frac{1}{N} \sum_{i=0}^{N} x_i.$$

  The **expectation value or mean** of a continuous random variable Y by

  $$E[Y] = \mu_Y = \int_{-\infty}^{\infty} y f_Y(y)dy$$

  The **expectation value or mean** of a continuous joint random variables (X,Y)

  $$E[XY] = \mu_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x,y)dxdy$$

  with $f_{XY}(x,y)$ the joint probability density function.

- The **variance of the random variable X** is defined by

  $$\sigma_X^2 = Var(X) = E[(X-\mu)^2]$$

  and it can be also written as

  $$\sigma_X^2 = Var(X) = E[X^2] - (E[X])^2.$$

  with the second central moment of a continuous random variable given by

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx$$

It is the expected squared distance of a value from the centre of the distribution.

Question 1

A variable Y has a probability density function

$$f_Y(y) = \begin{cases} \frac{1}{4}(2y+3) & 0 \leq y \leq 1\,, \\ 0 & \text{otherwise} \end{cases}$$

- compute the expectation value of Y and aY
- compute the expectation value of Y^2 and $(aY)^2$
- compute the variance of Y and aY
- compute the variance of $(aY + b)$
- would this result apply to other random variables?

Question 2

A service station has both self-service and full-service islands. On each island, there is a single regular unleaded pump with two hoses. Let X denote the number of hoses being used on the self-service island at a particular time, and let Y denote the number of hoses on the full-service island in use at that time. The joint pmf of X and Y appears in the accompanying tabulation.

		y		
$p(x, y)$		0	1	2
	0	0.10	0.04	0.02
x	1	0.08	0.20	0.06
	2	0.06	0.14	0.30

a. What is $P(X = 1 \text{ and } Y = 1)$?

b. Compute $P(X \leq 1 \text{ and } Y \leq 1)$.

c. Give a word description of the event $\{X \neq 0 \text{ and } Y \neq 0\}$, and compute the probability of this event.

d. Compute the marginal pmf of X and of Y. Using $p_X(x)$, what is $P(X \leq 1)$?

e. Are X and Y independent random variables? Explain.

Question 3

Ada is a room usage surveyor. She took data on the number of people using classrooms in the MTH building.

	small	medium	large
morning	17	7	3
afternoon	8	19	15

Let X be a random variable taking values 0,1,2 for small, medium and large room sizes.

Also, let Y be a random variable taking values 0 and 1 for when the observation was in the morning or afternoon, respectively.

- Determine the joint probability mass function for X and Y, and the marginal probabilities of X and Y.
- Are X and Y independent? - (for independence $P(X = x, Y = y) = P(X = x)P(Y = y)$ for all x, y.
- Find the covariance and correlation of X and Y.

Question 4

Let X be the discrete random variable 'number of tails shown when two coins are thrown'.

Define two more random variables X_1 = the number of tails shown on the first coin and X_2 the number of tails shown on the second coin.

- Calculate the covariances and correlations of the random variables.

Question 5

Suppose the joint pdf of (X, Y) is given by

$$f(x, y) = \begin{cases} \frac{6}{5}(x + y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

- Verify that this is a legitimate pdf.
- The probability $P(0 \le x \le \frac{1}{4}, 0 \le y \le \frac{1}{4})$.