Computational Foundations of Cognitive Science 1 (2009–2010)

School of Informatics, University of Edinburgh
Lecturers: Frank Keller, Miles Osborne

Solution for Tutorial 8: Joint Distributions; Expectation and Variance

Week 9 (08–12 March, 2010)

1. **Joint and Marginal Distributions**

   Two psychometric tests are administered to a group of 10 experimental subjects. The first test is a standardized aptitude test used for university admissions. It returns a score between 0 and 100. Let X be the random variable denoting the aptitude test score. The second psychometric test is a memory test. It measures the digit span, i.e., the number of digits in a sequence that a subject is able to remember before they make a mistake. Let Y be the random variable that denotes the digit span, ranging from 1 to 5. The results of the experiment are given in the following table:

Subject	1	2	3	4	5	6	7	8	9	10
x	60	50	40	50	60	30	30	20	30	50
y	5	3	3	3	3	3	3	4	4	4

   (a) Compute the distributions of X and Y.

   (b) Compute the joint distribution of X and Y.

   (c) Compute the marginal distributions of X and Y.

   (d) Are X and Y independent?

   (e) Compute the conditional distributions of X given $Y = 3$.

   Solution:

   (a)

x	20	30	40	50	60
$f(x)$	0.1	0.3	0.1	0.3	0.2
y	3	4	5		
$f(y)$	0.6	0.3	0.1		

   (b)

(x,y)	20	30	40	50	60
3	0	0.2	0.1	0.2	0.1
4	0.1	0.1	0	0.1	0
5	0	0	0	0	0.1

   (c)

(x,y)	20	30	40	50	60	$\sum_x f(x,y)$
3	0	0.2	0.1	0.2	0.1	0.6
4	0.1	0.1	0	0.1	0	0.3
5	0	0	0	0	0.1	0.1
$\sum_y f(x,y)$	0.1	0.3	0.1	0.3	0.2	

   (d) No, for example $f(40,3) = 0.1$, but $\sum_y f(40,y) \cdot \sum_x f(x,3) = 0.1 \cdot 0.6 = 0.06$

   (e)

x	20	30	40	50	60
$y = 3$	0	0.67	1	0.67	0.5

2. **Expectation and Variance**

(a) For the discrete random variable X with the following probability distribution:

$$f(x) = \frac{|x-2|}{7} \text{ for } x = -1, 0, 1, 2, 3$$

determine $E(X)$ and $\text{var}(X)$. Now assume the functions $g(X) = 3X+2$ and $h(X) = X^2$ and determine $E(g(X))$ and $E(h(X))$.

Solution:

$$E(X) = \sum_x x \cdot f(x) = -1\frac{3}{7} + 0\frac{2}{7} + 1\frac{1}{7} + 2\frac{0}{7} + 3\frac{1}{7} = \frac{1}{7}$$

$$\text{var}(X) = \sum_x (x-\mu)^2 \cdot f(x) = (-1-\frac{1}{7})^2\frac{3}{7} + (0-\frac{1}{7})^2\frac{2}{7} + (1-\frac{1}{7})^2\frac{1}{7} + (2-\frac{1}{7})^2\frac{0}{7} + (3-\frac{1}{7})^2\frac{1}{7}$$

$$E(g(X)) = \sum_x 3x + 2 \cdot f(x) = -1\frac{3}{7} + 2\frac{2}{7} + 5\frac{1}{7} + 8\frac{0}{7} + 11\frac{1}{7} = \frac{16}{7}$$

$$E(h(X)) = \sum_x x^2 \cdot f(x) = (-1)^2\frac{3}{7} + 0^2\frac{2}{7} + 1^2\frac{1}{7} + 2^2\frac{0}{7} + 3^2\frac{1}{7} = \frac{13}{7}$$

(b) In Chebyshev's theorem, which form does the inequality take for $k = 1, 2, 3, 4$?

Solution: The general form of Chebyshev's theorem is:

$$P(|x-\mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

So we get for $k = 1, 2, 3, 4$:

$$P(|x-\mu| < \sigma) \geq 0$$

$$P(|x-\mu| < 2\sigma) \geq \frac{3}{4}$$

$$P(|x-\mu| < 3\sigma) \geq \frac{8}{9}$$

$$P(|x-\mu| < 4\sigma) \geq \frac{15}{16}$$

3. **Covariance**

The *covariance* of two random variables X and Y with the joint distribution $f(x,y)$ is defined as:

$$\text{cov}(X,Y) = E((X-\mu_X)(Y-\mu_Y)) = \sum_x \sum_y (x-\mu_X)(y-\mu_Y) \cdot f(x,y)$$

where μ_X and μ_Y are the means of X and Y.

Assume that X and Y have the following joint distribution:

(x,y)	0	1	2
0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$
1	$\frac{2}{9}$	$\frac{1}{6}$	0
2	$\frac{1}{36}$	0	0

(a) Compute the marginal distributions of X and Y.

(b) Use the marginal distributions to compute μ_X and μ_Y.

(c) Now compute the covariance of X and Y.

Solution:

(a) The marginal distributions are:

(x,y)	0	1	2	
0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{7}{12}$
1	$\frac{2}{9}$	$\frac{1}{6}$	0	$\frac{7}{18}$
2	$\frac{1}{36}$	0	0	$\frac{1}{36}$
	$\frac{5}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	

(b)

$$\mu_X = 0\frac{5}{12} + 1\frac{1}{2} + 2\frac{1}{12} = \frac{2}{3}$$

$$\mu_Y = 0\frac{7}{12} + 1\frac{7}{18} + 2\frac{1}{36} = \frac{4}{9}$$

(c)

$$\text{cov}(X,Y) = \sum_x \sum_y (x - \frac{2}{3})(y - \frac{4}{9}) \cdot f(x,y) = -\frac{7}{54}$$