Formal Modeling in Cognitive Science 1 (2005–2006)

School of Informatics, University of Edinburgh
Lecturers: Mark van Rossum, Frank Keller

Solutions for Tutorial 8: Expectation and Variance; Special Distributions

Week 9 (6–10 March, 2006)

1. **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 $k = 1, 2, 3, 4$:

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

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

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

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

2. **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)

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

3. **Special Distributions**

(a) A scientist claims that 1 in 10 car accidents are due to driver fatigue. Using the formula for the binomial distribution, compute the probability that at most 3 of 5 accidents that happen on a given day are due to driver fatigue.

Solution:

$$b(0;5,0.1) + b(1;5,0.1) + b(2;5,0.1) + b(3;5,0.1) =$$

$$\binom{5}{0}0.1^0 0.9^5 + \binom{5}{1}0.1^1 0.9^4 + \binom{5}{2}0.1^2 0.9^3 + \binom{5}{3}0.1^3 0.9^2 = 0.0086$$

(b) In a reaction time experiment, the response latency in seconds is distributed according to the standard normal distribution. What is the probability that the reaction time is between 0 and 1 seconds?

Solution:

$$P(0 < X < 1) = \int_0^1 n(x;0,1)dx = \int_0^1 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = 0.3413$$

We can work this out by integrating, or because we know that for the normal distribution, $P(|x - \mu| < \sigma) = 0.6826$, hence for the standard normal distribution $P(x < 1) = 0.6826/2 = 0.3413$.