COMS10003 Work Sheet 22

Random Variables I

Andrew Calway January 27, 2015

1. Continuous random variable X and discrete random variable Y have probability density function (pdf) and probability mass function (pmf) given below. In each case determine and sketch the corresponding cumulative distribution function (cdf).

$$p(x) = \begin{cases} 1/a & 0 \leq x \leq a \\ 0 & else \end{cases} \qquad P(y) = \frac{1}{4}\sum_{i=1}^{4} \delta(y-i)$$

   where $\delta(y) = 1$ if $y = 0$ and zero otherwise.

   Answer:

   $c(x) = 0$ for $x < 0$, $c(x) = x/a$ for $0 \leq x \leq a$, $c(x) = 1$ for $x > a$

   $c(x)$ is staircase function between 0 and 4, $c(x) = 0$ for $x < 0$ and $c(x) = 1$ for $x > 4$.

2. A discrete random variable X can be one of three values with probabilities given by its pmf $P(-1) = 0.2$, $P(2) = 0.5$, $P(6) = 0.3$. Compute the expected value, the variance, $Var(X)$ and the standard deviation of X.

   Answer:

   $E(X) = 2.6$, $Var(X) = 6.24$, standard deviation $= \sqrt{(6.24)}$

3. Determine the sample mean and the sample variance of the following set $\{2, 5, 3, 7, 8, 12, 3, 2\}$

   Answer:

   $\bar{x} = 5.25$, $\bar{\sigma}^2 = 12.5$

4. The joint pmf for two discrete random variables X and Y taking values $0, 1, 3$ and $-1, 1, 2$, respectively, is shown below.

$X : Y$	-1	1	2
0	0.12	0.08	0.1
1	0.2	0.04	0.25
3	0.08	0.1	0.03

   (a) Compute the marginal mass functions

   (b) Are X and Y independent? Justify your answer.

   (c) Compute $E(X)$ and $E(Y)$.

(d) Compute the pmf for the RV $Z = X + Y$

(e) Compute the conditional pmf $P(X|Y = 2)$.

Answer:

(a) *P(X=0)=0.3,P(X=1)=0.49,P(X=3)=0.21*
$P(Y = -1) = 0.4,\ P(Y = 1) = 0.22,\ P(Y = 2) = 0.38$

(b) *No, P(X=x) conditionally depends on Y*

(c) $E(X) = 1.12$ *and* $E(Y) = 0.58$.

(d)

z	*-1*	*0*	*1*	*2*	*3*	*4*	*5*
P(Z=z)	*0.12*	*0.2*	*0.08*	*0.22*	*0.25*	*0.1*	*0.03*

(e) $P(X = 0|Y = 2) = P(X = 0, Y = 2)/P(Y = 2) = 0.1/0.38,\ P(X = 1|Y = 2) = 0.25/0.38,\ P(X = 3|Y = 2) = 0.03/0.38$

5. Prove the following for scalar a and random variables X and Y:

$$E(aX) = aE(X) \qquad E(X+Y) = E(X) + E(Y) \qquad Var(aX) = a^2Var(X)$$

Answer:

$E(aX) = \int axp(x)dx = a\int xp(x)dx = aE(X)$

$E(X+Y) = \int\int(x+y)p(x,y)dxdy = \int\int xp(x,y)dxdy + \int\int yp(x,y)dxdy = \int x\int p(x,y)dydx + \int y\int p(x,y)dxdy = \int xp(x)dx + \int yp(y)dy = E(X) + E(Y)$

$Var(aX) = \int(ax - aE(X))^2p(x)dx = a^2\int(x - E(X))^2p(x)dx = a^2Var(X)$

6. If X and Y are independent random variables, prove the following:

$$E(XY) = E(X)E(Y) \qquad Var(X+Y) = Var(X) + Var(Y)$$
$$Var(XY) = E(X^2)E(Y^2) - E(X)^2E(Y)^2$$

Answer:

If indepedent then $p(x,y) = p(x)p(y)$, *so* $E(XY) = \int\int xyp(x)p(y)dxdy = \int xp(x)dx\int yp(y)dy = E(X)E(Y)$

$Var(X+Y) = \int\int(x+y-E(X)-E(Y))^2p(x)p(y)dxdy = \int(x-E(X))^2p(x)dx + \int(y-E(Y))^2p(y)dy + 2\int\int(x-E(X))(y-E(Y))p(x)p(y)dxdy = \int(x-E(X))^2p(x)dx + \int(y-E(Y))^2p(y)dy = Var(X) + Var(Y)$

$Var(XY) = \int\int(xy - E(X)E(Y))^2p(x)p(y)dxdy$
$= \int\int x^2y^2p(x)p(y)dxdy - 2E(X)E(Y)\int\int xyp(x)p(y)dxdy + E(X)^2E(Y)^2 = E(X^2)E(Y^2) - E(X)^2E(Y)^2$

7. A fair die is thrown 10 times. Compute the probability that the die will show a number less than 5 on exactly 3 of the 10 throws.

   Answer:

   Binomial distribution with $p = 4/6 = 2/3$ *and* $n = 10$. $P(k = 3) = 10!/(3!7!)(2/3)^3(1/3)^7 = 0.0163$

8. Over one week, Alice received 50 spam emails and 150 non-spam emails. She then receives 10 emails during one day in the following week. Estimate the probability that exactly 6 of those will be spam and the probability that she will receive less than 3 spam emails.

   Answer:

   Binomial distribution with $p = 50/200 = 1/4$ *and* $n = 10$. $P(k = 6) = 10!/(6!4!)(1/4)^6(3/4)^4 = 0.0162$

   $P(k < 3) = (3/4)^10 + 10!/9!(1/4)(3/4)^9 + 10!/(2!8!)(1/4^2)(3/4)^8 = 0.5256$

9. Prove that the expected value and variance of a random variable which follows the binomial distribution is $E(X) = np$ and $Var(X) = np(1-p)$.

   Answer:

   Assume successive experiments, eg flips of coin, are independent Bernoulli trials. Thus expected value of sequence of experiments is sum of individual expected values, ie given n *experiments* $E(X) = np$. *Similarly for variance - individual variance is* $p(1-p)$ *for Bernoulli, hence for Binomial* $var(X) = np(1-p)$

10. (Buffon's Needle - version taken from the Wikipedia entry - try not to peep!) A needle of length L is dropped onto a plane which is ruled with parallel lines T units apart. Show that the probability that the needle will lie across a line is given by $2L/T\pi$. *Hint: consider the problem in terms of two random variables - the distance of the needle centre from one of the parallel lines and the angle the needle makes with the parallel lines. Assume that they are independent (are they really?) and then determine the joint pdf. That should get you started.*

    Answer: *See wikipedia page*