ST3009 Mid-Term Test 2018

Attempt **all** questions. Time: 1 hour 30 mins.

1.

(i)

(a) Solve the equation $-5x+20=25$ (find the value of x) [1 mark]
(b) Solve the equation $5(2x+1)+3=10$ (find the value of x) [1 mark]
(c) Suppose $x-y=0$ and $x+y=1$. What are the values of x and y? [1 mark]
(d)) Is $x/2+y/2+z/2 = (x+y+z)/2$? Briefly explain. [1 mark]
(e) Simplify $(xy+xz)/x$ [1 mark]

(ii) Define the term "random variable" and give an example. [5 marks]

(iii)What is the probability mass function of a discrete random variable? Give an example. [5 marks]

Let X and Y be independent random variables that take values in set {-1,0,+1}. Assume that X and Y are uniformly distributed on {-1, 0, 1} i.e. the probability of each value occurring is the same. Let V =2X+2Y.

(iv) Calculate E[X] and E[V] [5 marks]

(v) Define what it means for two random variables to be independent. [5 marks]

(vi) Are V and X independent? Explain with respect to the definition of independence. [5 marks]

Model Solution

(i)(a) $x=(25-20)/(-5)= -1$, (b) $x=((10-3)/5-1)/2 = 0.2$, (c) from first equation x=y, substituting this into the second equation gives x=0.5=y, (d) yes, we can take the division by 2 on the RHS inside the brackets, (e) $x(y+z)/x=y+z$.

(ii) A random variable maps from the sample space of a random experiment to a real number. For example, we might define random variable X to take value 1 when a coin comes up heads and value 0 when the coin comes up tails.

(iii) For a random variable X taking values in $\{x_1, x_2, .., x_n\}$ then the set of probabilities $P(X=x_1), P(X=x_2), \ldots, P(X=x_n)$ is the PMF. For example, when X takes value 1 when a coin comes up heads and 0 otherwise then $P(X=1)=0.5=P(X=0)$ is the PMF.

(iv) $E[X] = -1.P(X=-1)+0.P(X=0)+1.P(X=1) = -1/3+0/3+1/3=0$. The long way to get E[V] is using $E[V]=-4\times P(X=-1, Y=-1)-2\times (P(X=-1,Y=0)+ P(X=0,Y=-1))+0\times (P(X=-1,Y=1)+P(X=0,Y=0)+ P(X=1,Y=-1)) + 2\times (P(X=0,Y=1)+P(X=1,Y=0))+4\times P(X=1,Y=1) = -4/9-2.2/9+2.2/9+4/9=0$. A shorter way is to use the linearity of the expectation i.e. $E[V]=E[2X+2Y]=2E[X]+2E[Y]=0$ since $E[X]=0=E[Y]$.

(v) Two RVs X and V are independent if $P(X=x \text{ and } V=v)=P(X=x)P(V=v)$ for all pairs of values (x,v) that the RVs can take.

(vi) No. For example, P(X=-1 and V=4)=P(V=4|X=-1)P(X=-1)=0 and P(X=-1)=1/3 and P(V=4)=P(X=1 and Y=1)=1/9. So P(X=-1 and V=4)=0 is not equal to P(X=-1) P(V=4)=1/3.1/9

2.

(i) Define the conditional probability of a random event and state Bayes Theorem. [5 marks]

(ii) Suppose two websites A and B take hotel bookings. Site A takes 60% of all bookings and site B takes 40%. However, only 75% of the bookings made on site A result in positive reviews after the hotel stay, while on site B it is 90%. Given that a booking received a positive review, what is the probability that booking was made on site B ? Hint: use Bayes Theorem. [10 marks]

Model Solution

(i) For random events E and F the conditional probability P(E|F)=P(E∩F)/P(F).
Bayes Theorem is: P(E|F)= P(F|E)P(E)/P(F)

(ii) Let event F be that the booking received a positive review and event E that the booking was made on site B. We want to calculate P(E|F). P(E)=0.4, P(F|E)=0.9 and $P(F)=P(F|E)P(E)+P(F|E^c)P(E^c)=0.9\times0.4+0.75\times0.6=0.81$. Using Bayes it follows that $P(E|F)=0.4\times0.9/0.81=0.44$

3. Data is stored in encoded form across 10 disks to provide some protection against disk failures. To read a file data needs to be successfully read from any 3 of the 10 disks.

(i) Suppose a server selects 3 disks uniformly at random to read from. What is the probability that disk 1 is read ? Hint: think of drawing balls from a bag without replacement. [10 points]

(ii) Suppose now that disks 1 and 2 cannot be read together (the set of disks that can be read includes disk 1 or disk 2 or neither, but not both). What is the probability that disk 1 is read now ? [10 points]

(iii) Each disk fails independently with probability 0.01. Remember 3 disks need to be read successfully to reconstruct a file. When the server reads 3 disks what is the probability that the file fails to be reconstructed ? [5 points]

(iv) With the same setup as in (iii) what is the probability when the server now reads 4 disks ? [5 points]

Model Solution

(i) Number of ways we can select 3 disks out of 10 without replacement is $10\times9\times8$. Suppose disk 1 is selected in the first position, then there are $1\times9\times8$ ways for this to happen, similarly for the disk 1 in the second and third positions so in total there are $3\times1\times9\times8$ ways to select disk 1 and the probability is $3\times1\times9\times8/(10\times9\times8)=3/10=0.3$

(ii) Suppose disk 1 is selected in the first position, then there are $1\times8\times7$ ways for this to happen since disk 2 cannot be selected. Similarly, when disk 1 is in positions 2 and 3. So there are $3\times1\times8\times7$ ways to select disk 1. Using the same argument there are also $3\times1\times8\times7$ ways to select disk 2, and also $8\times7\times6$ ways to select neither disk 1 or disk 2. So the probability of selecting disk 1 is $3\times1\times8\times7/(3\times1\times8\times7+3\times1\times8\times7+8\times7\times6)= 1/4=0.25$

(iii) The probability that **none** of the three disks fails is $(1-0.01)^3$. So the probability than one or more fails is $1-(1-0.01)^3\approx0.03$.

(iv) The probability that no disks fail is $(1-0.01)^4$. The probability that one disk fails is $4\times0.01\times(1-0.01)^3$. So the probability that the server is successful is $(1-0.01)^4+4\times0.01\times(1-0.01)^3$ and that it fails is $1-(1-0.01)^4-4\times0.01\times(1-0.01)^3\approx0.0006$