1. (i) Define the terms "sample space", "event" and "random variable" and give an example of each.

Sample space: Set of all possible outcomes for an experiment. Eg. for dice roll, sample space = {1,2,3,4,5,6}

Event: Event E is a subset of sample space S, $E \subset S$, a set of possible outcomes when an experiment is performed. If E and F are events then so are: $E \cup F$, $E \cap F$, E^c, F^c. Eg. rolling a 5 when rolling a dice.

Random variable: Function that maps from sample space S to real line R, maps every event to a real number, eg. takes value 1 if Event E occurs and 0 if does not occur.

(ii) What is an indicator random variable and what is the probability mass function of a discrete random variable ?

Indicator random variable: A random variable that has the value 1 or 0, according to whether a specified event occurs or not

Probability Mass Function of Discrete Random Variable: A probability is associated with each value that a discrete random variable can take. P(X = x) for the probability that random variable X takes value x. P(X = x1), P(X = x2), . . . , P(X = xn)

(iii) Define the conditional probability of an event and state Bayes Theorem.

Conditional probability: The probability that event E occurs given that event F has already occurred

Bayes Theorem: P(E|F) = (P(F|E)P(E))/P(F)

(iv) Explain what is meant by "marginalization".

Sum of values of joint probability distribution.

2. Suppose we have two bags, labeled A and B. Bag A contains 3 white balls and 1 black ball, bag B contains 1 white ball and 3 black balls. We toss a fair coin and select bag A if it comes up heads and otherwise bag B. From the selected bag we now draw 5 balls, one after another, replacing each ball in the bag after it has been selected (the bag always contains 4 balls each time a ball is drawn). We observe 4 white balls and 1 black ball. What is the probability that we selected bag A? Hint: use Bayes Rule.

A = $(0.75^4 * 0.25)$ = 0.0791015625
B = $(0.25^4 * 0.75)$ = 0.0029296875

E = choose bag A
F = observe 4 white and 1 black

$P(E) = ½$
$P(F) = (0.0791015625 * 0.5) + (0.0029296875 * 0.5) = 0.041015625$
$P(E|F) = (0.0791015625 * 0.5)/ 0.041015625 = 0.96428571428$

3. (i) Define the expected value of a random variable. Give a proof that the expected value is linear i.e. E[X+Y]=E[X]+E[Y] for random variables X and Y.

Expected value: For random variable X taking values x1, …, xn the expected value is E[X]=x1*P(X=x1)+… xn*P(X=xn)

E[aX + b] = aE[X] + b
E[aX + bY] = (sum x)(sum y)(ax + by)(P X = x and Y = y)
= a (sum x) (sum y) xP(X = x and Y = y) + b (sum y)(sum x) yP(X = x and Y = y)
= a (sum x) xP(X = x) + b (sum y) yP(Y = y) - using marginalisation
= aE[X] + bE[Y]

(ii) Define what it means for two random variables to be independent.Give a proof that when two random variables X and Y are independent then E[XY]=E[X]E[Y].

Observing one does not affect the outcome of the other.
E[XY] = (sum x)(sum y) xyP(X = x and Y = y)
= (sum x)(sum y) xyP(X = x)P(Y = y)
= (sum x) xP(X = x) sum y yP(Y = y)
= E[X]E[Y]

(iii) Define the covariance and correlation of two random variables X and Y.

Covariance: Cov(X, Y) = E[(X − μX)(Y − μY)] = E[XY] − E[X]E[Y]
Cov(X, Y) is positive if X and Y tend to increase together, and negative if and increase in one tends to correspond to a decrease in the other

Correlation: Corr(X, Y) = Cov(X, Y)/(root(Var(X)Var(Y)))

4. (i) A bag contains 30 balls, of which 10 are red and the other 20 blue. Suppose you take out 8 balls from this bag, with replacement. What is the probability that among the 8 balls in this sample exactly 3 are red and 5 are blue?

$(⅓)^3 (⅔)^5 * 8!/(5!*3!)$

(ii) Now suppose that the balls are taken out of the bag without replacement. What is the probability that out of 8 balls exactly 3 are red and 5 are blue?

(20C5 * 10C3)/30C10 = 0.31787183331