ST003: Statistical Methods For Computer Science

Assignment 5

Question 1. A box contains 5 red and 5 blue marbles. Two marbles are withdrawn randomly. If they are the same colour, then you win $1.10; if they are different colours, then you lose $1.00. Calculate:

(a) The expected value of the amount you win

E[X] = Xn i=1 xiP(X = xi) // from notes

P(winning) * win + P(losing) * loss = expected value

E[x] = 1.1 * 4/9 + (-1) * 5/9 = ***-0.06667***

(b) The variance of the amount you win.

Var(X) = E[x^2] - (E[x])^2 // from notes

E[X^2] = (1.1)^2 * 4/9 + (-1)^2 * 5/9 = 1.0933

(E[X])^2 = (-0.06667)^2 = 0.003086419

= ***1.08889***

Question 2. Suppose you carry out a poll following an election. You do this by selecting n people uniformly at random and asking whether they voted or not, letting Xi = 1 if person i voted and Xi = 0 otherwise. Suppose the probability that a person voted is 0.6.

(a) Calculate E[Xi] and Var(Xi).

letting Xi = 1 if person i voted and Xi = 0 otherwise. // from questions

E[Xi] = P(vote) * 1 + (1-p(vote)) * 0 // values of Xi multiplied by their probability

E[Xi] = 0.6 * 1 + 0.4 * 0 = ***0.6***

Var(Xi) = E[Xi^2] – E[Xi]^2

= 0.6 * 1^2 + 0.4 * 0^2

Var(Xi) = 0.6 – (0.6)^2

Var(Xi) = ***0.24***

(b) What is E[Y] ? Is it the same as E[X] or different, and why ?

E[X] = E[Xi] // since the people are sampled independently

Thus, E[X] = 0.6

E[Y] = the expected number of people who did vote (i=1) from the sample (n) // from formula

Thus, E[Y] = 0.6n // likelihood of vote multiplied by sample size

0.6 != 0.6n therefore, ***E[Y] is not the same as E[X]***

(c) What is E[1/ n Y] ?

Since E[Y] = 0.6n // found in previous question

E[(1/n)Y] = (1/n) * 0.6n

0.6n / n is simply 0.6

Thus, E[(1/n)Y] = ***0.6***

(d) What is the variance of 1 /n Y (express in terms of V ar(X)) ? Hints: use linearity of the expectation and the fact that people are sampled independently.

Var((1/n)*Y) = (1/n)^2Var(Y) // taken from notes

Var(Y) = Var(∑(n, i=1) Xi) // E[Y] = E[Xn] , calculated in previous answer thus n will be multiplied

= ∑(n, i=1)Var(Xi) // linearity notes

= nVar(X) // sum =n and var(Xi) = Var(X) since people are sampled independently

Thus, Var((1/n)*Y) = (1/n)^2 * n (Var(X))

= ***(1/n)Var(X)***

Question 3. Suppose that 2 balls are chosen without replacement from an urn consisting

of 5 white and 8 red balls. Let Xi equal 1 if the i'th ball selected is white, and let it equal

0 otherwise.

(a) Give the joint probability mass function of X1 and X2

	x = 0	x = 1	P(X2 = y)
y = 0	***14/39***	***10/39***	***24/39***
y = 1	***10/39***	***5/39***	***15/39***
P(X1 = x)	***8/13***	***5/13***	***1***

Probability of X1 is out of 13 potential balls, X2 is 12. Since X2 must come after X1.

Probabilities were calculated using this assumption.

(b) Are X1 and X2 independent ? (Use the formal definition of independence to determine this)

The formal definition of independence states that two events E and F are independent if:

P(E ∩ F) = P(E) * P(F)

P(X1 ∩ X2) = 5/13 * 15/39 = 25/169 // values calculated for previous question

P(X1) * P(X2) = 5/39 // value calculated in previous question

5/39 ! = 25/169 therefore ***X1 and X2 are not independent.***

(c) Calculate E[X2]

E[X2] = P(X2=white) * 1 + (1-p(X2=white)) * 0 // values of X2 multiplied by their probability

E[Xi] = 15/39 * 1 + 24/39 * 0 = ***15/39***

(d) Calculate E[X2|X1 = 1]

E[X|Y = y] = ∑ x xP(X = x|Y = y) // taken from notes

P(X = x|Y = y) = P((X=x and Y =y) / P(Y =y) // notes

P((X=x and Y =y) / P(Y =y) = (5/39) / (5/13) =1/3 // values taken from q3a

E[X2|X1 = 1] = (p(true) * 1) + (1-p(true) * 0)

= 1/3 + (1-1/3 * 0) = ***1/3 = 0.3333***