Machine Learning

Homework 2 solution - Shyam Arumugaswamy

1. Let A be a Bernoulli random variable that shows that passenger is terrorist or not
   P (A) = 0.01 for A = 1 …. Terrorist
   0.99 for A = 0 …. Not terrorist

   Let B be another Bernoulli random variable that depicts the scanner result
   P(B | A = 1) = 0.95 for B = 1 i.e. P(B = 1 | A = 1) = 0.95
   0.05 for B = 0 i.e. P(B = 0 | A = 1) = 0.05

   Similarly
   P(B = 1 | A = 0) = 0.05
   P(B = 0 | A = 0) = 0.95

   So we want probability that passenger is terrorist given that scanner result is terrorist
   so by Baye's rule

   $$P(A = 1 \mid B = 1) = \frac{\mathrm{P(B = 1 \mid A = 1)*P(A=1)}}{(\mathrm{P(B = 1 \mid A = 1)*P(A=1)})+(\mathrm{P(B = 1 \mid A = 0)}*P(A=0))}$$

   $$= \frac{0.95*0.01}{(0.95*0.01)+(0.05*0.99)}$$

   $$= \frac{19}{118}$$

   = 0.161

2. A coin is tossed twice. So sample space will be S = {HH, HT, TH, TT}
   For each H, Red ball(R) is placed or else white ball(W)
   so various possibility of balls in the box will be {RR, RW,WR,WW}
   each with a probability of ¼
   P(RR) = P(RW) = P(WR) = P(WW) = ¼

The event of picking 3 red balls from the above possibilities are:

P(RRR | RR) = 1
P(RRR | RW) = P(RRR | WR) = ½ * ½ * ½ = 1/8
P(RRR | WW) = 0

By Baye's rule, probability of picking from box of red balls
P(RR | RRR) =

$$\frac{\mathrm{P(RRR \mid RR) * P(RR)}}{\mathrm{(P(RRR \mid RR) * P(RR)) + (P(RRR \mid RW) * P(RW)) + (P(RRR \mid WR) * P(WR)) + (P(RRR \mid WW) * P(WW))}}$$

$$= \frac{1/4}{\frac{1}{4} + \left(\frac{1}{4} * \frac{1}{8}\right) + \left(\frac{1}{4} * \frac{1}{8}\right) + 0}$$

$$= \frac{4}{5} \quad = 0.8$$

3. We stop the coin toss as soon as we get Heads,
   So expected number of heads E(H) = 1

   let X be number of coin flips required until head is received. E(X) be first flip, E(X|T) is number of remaining coin flips given tail is received in first flip and E(X|H) is number of remaining coin flips given head is received in first flip
   E(X) = ½ * (1 + E(X|T)) + ½ * (1 + E(X|H))
   E(X|H) = 0, as no need to flip after head is received in first flip and
   E(X|T) = E(X) as no progress was made towards getting head

   E(X) = ½ * (1 + E(X)) + ½ * (1 + 0)
   E(X) = 2

   Expected total flips = 2
   and expected number of heads is 1
   so the expected number of tails is 1.

4. P(x) $= \frac{1}{b-a}$ a <= x <= b

0 elsewhere

Mean : $E[X] = \int_{-\infty}^{+\infty} x\, p(x) dx$

$$= \int_a^b \frac{x}{b-a} dx$$

$$= \frac{1}{b-a} \left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{1}{b-a} \left[\frac{b^2}{2} - \frac{a^2}{2}\right]$$

$$= \frac{a+b}{2}$$

$E[X^2] = \int_{-\infty}^{+\infty} x^2\, p(x) dx$

$$= \int_a^b \frac{x^2}{b-a} dx$$

$$= \frac{1}{b-a} \left[\frac{x^3}{3}\right]_a^b$$

$$= \frac{1}{b-a} \left[\frac{b^3}{3} - \frac{a^3}{3}\right]$$

$$= \frac{a^2 + ab + b^2}{3}$$

Variance : $Var[X] = E[X^2] - E[X]^2$

$$= \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{a^2 - 2\,ab + b^2}{12}$$

$$= \frac{(a-b)^2}{12}$$

5. To prove :

$E[X] = E_y [E_{x|y} [X]]$

$Var[X] = E_y [Var_{x|y} [X]] + Var_y [E_{x|y} [X]]$

1st Proof:

<u>RHS</u>

$E_y [E_{x|y} [X]] = \int [\int x\, p(x|y)\, dx]\, p(y)\, dy$

$= \int\int x\, p(x|y)\, p(y)\, dx\, dy$

$= \int\int x\, p(x,y)\, dx\, dy$

$= \int x \int p(x,y)\, dy\, dx$

$= \int x\, p(x)\, dx$

$= E[X]$

2nd Proof

<u>RHS</u>

$E_y [Var_{x|y} [X]] + Var_y [E_{x|y} [X]]$

$= E_y [E_{x|y} [X^2]] - E_y [(E_{x|y} [X])^2] + E_y [(E_{x|y} [X])^2] - E_y ([E_{x|y} [X])^2$

$= E_y [E_{x|y} [X^2]] - E_y ([E_{x|y} [X])^2$

$= E_y [E_{x|y} [X^2]] - Ex^2[X]$ --------using first tower property

$= E_X [X^2] - Ex^2[X]$ --------using first tower property

$= Var[X]$

6. $P(|\frac{1}{n}\sum_{i=1}^{n} Xi - \sum[Xi]| > \varepsilon)$

$= P(|\sum_{i=1}^{n} Xi - n\sum[Xi]| > n\varepsilon)$

we take $Y = \sum_{i=1}^{n} Xi$ and $E[Y] = \sum_{i=1}^{n} Xi = nE[X]$

$= p(|Y – E[Y]| > n\varepsilon)$

From Chebyshev inequality, we can write above eqn as

$= p(|Y – E[Y]| > n\varepsilon) < \frac{Var[Y]}{n^2\varepsilon^2} = \frac{Var[X]}{n\varepsilon^2}$ which tends to 0