Name : Akshay Tumkur Renukaprasad
Matriculation number : 03694060

Machine Learning – Second Assignment : Probability refresher

Solution 1:

Let T be the random variable that illustrates if a person is terrorist or not and let S indicate the scanner test, whether terrorist or not.

Probability of being a terrorist $P(T=1) = 0.01$

Hence, probability of not being terrorist $P(T=0) = 0.99$

Below mentioned scenarios can be considered for different values of S and T

$P(S=1 \mid T=1) = 0.95$
$P(S=1 \mid T=0) = 0.05$
$P(S=0 \mid T=0) = 0.95$
$P(S=0 \mid T=1) = 0.05$

To calculate, probability that the person is terrorist

$P(T=1 \mid S=1) \Rightarrow P(T=1 \cap S=1) / P(S=1)$

By laws of probability

$P(T=1 \cap S=1) \Rightarrow P(S=1 \mid T=1)*P(T=1)$
$P(S=1) \Rightarrow P(S=1 \mid T=1)*P(T=1) + P(S=1 \mid T=0)*P(T=0)$

Then,

$P(T=1 \mid S=1) \Rightarrow P(S=1 \mid T=1)*P(T=1) / P(S=1 \mid T=1)*P(T=1) + P(S=1 \mid T=0)*P(T=0) \Rightarrow$

$0.95*0.01 / (0.95*0.01 + 0.05*0.99) \approx \mathbf{0.16}$

Solution 2:

Consider the balls drawn event 3 times in succession as RRR. Similarly considering the ball placing event as ‘rr’ (red and red), ‘rw’ (red and white), ‘wr’ (white and red) and ‘ww’ (white and white).

To calculate both balls in the box are red, $P(rr \mid RRR)$

$P(rr \mid RRR) \Rightarrow P(rr \cap RRR) / P(RRR)$
By laws of probability
$P(rr \cap RRR) \Rightarrow P(RRR \mid rr)*P(rr)$

P(RRR) => P(RRR | rr)*P(rr) + P(RRR | rw)*P(rw) + P(RRR | wr)*P(wr) + P(RRR | ww)*P(ww) Then

final representation will be:

=>P(RRR | rr)*P(rr) / (P(RRR | rr)*P(rr) + P(RRR | rw)*P(rw) + P(RRR | wr)*P(wr) + P(RRR | ww)*P(ww))

P(RRR | rr) => 1	P(rr) = 1/4
P(RRR | rw) => 1/2*1/2*1/2 = 1/8	P(rw) = 1/4
P(RRR | wr) => 1/2*1/2*1/2 = 1/8	P(wr) = 1/4
P(RRR | ww) => 0	P(ww) = 1/4

=> 1*1/4 / (1*1/4 + 1/8*1/4 + 1/8*1/4 + 0*1/4)
=> 1/4 / (5/16) => **4/5**

<u>Solution 3:</u>

As given in the hint, according to the geometric distribution. Let X denote the number of tails until first head shows up.

The geometric distribution is P(X = ∞) => $q^{x-1}p$ where

p is the probability of occurrence of tail = 1/2

Then expectation of this distribution $E(X)=\sum_{x=1}^{\infty}(x)q^{(x-1)}(p)$

=>p + 2qp + $3q^2p$ + $4q^3p$ + …..

=> p (1 + 2q + $3q^2$ + $4q^3$ + …..)

we know that : 1 + 2q + $3q^2$ + $4q^3$ + ….. = $1/(1 – q)^2$
then,

=> $p/((1 – q)^2$

=> p/p^2 => 1/p => 1/0.5 => 2

Expected number tails in any run of experiment = 2

Hence, the expected number of heads H in one run of this experiment is 2.

Since we stop the experiment after encountering the first head, the expected number of tails T in one run of this experiment is

E(H) – 1 = 2 -1 = 1.

Solution 4:

Given uniform distribution $f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \le x \le b, \\ 0 & \text{for } x < a \text{ or } x > b \end{cases}$ f(x) = 1/(b – a)

f(x) = 1/(b – a)

the mean of uniform distribution E[X] = $\int_a^b xf(x)dx$ => $\int_a^b x/(b-a)dx$

=> $1/(b-a)\int_a^b xdx$ => $1/(b-a)\left|x^2/2\right|_a^b$ => $(b^2-a^2)/(2(b-a))$

=> $(b-a)(b+a)/2(b-a)$ => $b+a/2$

$E[X]=b+a/2$

Var(X) = E[X^2] – (E[X])^2

E[X^2] = $\int_a^b x^{2f}(x)dx$ => $\int_a^b x^2/(b-a)dx$ => $1/(b-a)\int_a^b x^2dx$

=> $1/(b-a)\left|x^3/3\right|_a^b$ => $(b^3-a^3)/3(b-a)=(b^2+ab+a^2)/3$

Var(X) = E[X^2] – (E[X])^2 => $(b^2+ab+a^2)/3-((b+a)/2)^2$

=> $(4b^2+4ab+4a^2-3b^2-3a^2-6ab)/12$

=> $(b^2+a^2-2ab)/12=(b-a)^2/12$

$Var(X)=(b-a)^2/12$

Solution 5 :

To prove

$$E[X] = E_Y[E_{X|Y}[X]],$$
$$\mathrm{Var}[X] = E_Y[\mathrm{Var}_{X|Y}[X]] + \mathrm{Var}_Y[E_{X|Y}[X]].$$

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

To prove

$$p\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E[X_i]\right| > \epsilon\right) \to 0 \qquad \text{as } n \to \infty.$$

$= \quad var[X] = E[X^2] - (E[X])^2$

$= \quad \int (\int x^2 p(x,y)dx)p(y)dy - (E[X])^2$

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

$= \quad E_y([E_{(x|y)}[X^2]] - (\int xp(x/y)dx)^2 + (\int xp(x/y)dx)^2) - E[X]^2$

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

$= \quad E_y[Var_{(X|Y)}[X]] + Var_{(X|Y)}[E[X]]$

Solution 6:

As given in the question, we can assume that X_i have a finite variance $Var[X_i]$. We know that

$\frac{1}{n}\sum_{i=1}^{i=n} E[X_i] = \mu$ and $\frac{1}{n}\sum_{i=1}^{i=n} Var[X_i] = \frac{\sigma^2}{n}$. Therefore a direct application of Chebyshev's

inequality shows that $P\left[\left|\frac{1}{n}\sum_{i=1}^{i=n} Xi - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{n}$

And we can see that $P\left[\left|\frac{1}{n}\sum_{i=1}^{i=n} Xi - \mu\right| > \epsilon\right] \leq \frac{\sigma^2}{n}$ $\rightarrow 0$ as $n \rightarrow \infty$ Proved.