SMML: Exercise 02

Abhijeet Parida

Section 1-Quiz

If you are training your machine learning task for every input with corresponding target, it is called supervised learning, which will be able to provide target for any new input after sufficient training.

eg- Teaching alphabets to a child.

If you are training your machine learning task only with a set of inputs, it is called unsupervised learning, which will be able to find the structure or relationships between different inputs.

eg- vegetable shop arrangement potatoes with potatoes, tomatoes with tomatoes.

Joint probability is used to esimate the possibility of two events occuring simultaneously.
eg- probability of me cooking and going to market today.

Conditional probablity is the estimation of probabilty of an event happening provided another event has taken place.
eg- probability of me cooking mushroom today if buy mushroom today

Marginal probabilty is the probability of an event occuring which is not conditioned on any other event.(Unconditonal prbability).

The model developed can be tested using the validation test created out of the initial test cases. The overfitting of a model for curve fitting can be overcome by

1) limiting the coefficients of the polynomial

2) using a large no.of points

Bayes theorem describes the probability of an event, based on prior knowledge of conditions after an observation has been made pertaining to the event.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)}$$

Section 2-Bayesian Reasoning

$D = \{0, 1\}$ — 0: no disease, 1: disease

$T = \{0, 1\}$ — 0: test false, 1: test true

$P(D=1) = \frac{1}{10^6}$ $\quad P(T=1 \mid D=1) = 1$ $\quad P(T=1 \mid D=0) = \frac{01}{10^4}$

1) $P(T=1) = 10^{-6} + 10^{-4}\frac{(10^6 - 1)}{10^6}$

$= 10^{-6} + 10^{-4} - 10^{-10}$

≈ 0.000101

2) $P(D=1 \mid T=1) = \frac{P(T=1 \mid D=1) \cdot P(D=1)}{P(T=1)} = \frac{P(D=1 \cap T=1)}{P(T=1)}$

$= \frac{10^{-6}}{10^{-6} + 10^{-4} - 10^{-10}}$

3) $P(k_1 = c) = \frac{1}{3}$

$P(k_1 \neq c) = 1 - P = \frac{2}{3}$

$P(k_2 = c) = \sum_{k_1} P(k_2 = c, k_1) = \sum_{k_1} P(k_2 = c \mid k_1) \cdot P(k_1)$

no switch $\Rightarrow p(k_2 = c \mid k_1 = c) \cdot p(k_1 = c)$

$= \frac{1}{3}$

switch $\Rightarrow p(k_2 = c \mid k_1 \neq c) \cdot p(k_1 \neq c)$

$=$

Section 3-Coin Tossing

4) $X = \{0, 1\}$

$0 \rightarrow$ Tail, $1 \rightarrow$ Head

$P(x=1) = p$

$P(x=0) = 1-p$

$P(x|p) = p^{x}(1-p)^{1-x}$

likelyhood.

$$P(x_1 \dots x_n | p) = \prod_{i=1}^{N} P(x_i|p) = p^{x_m}(1-p)^{N-x_m}$$

log likelyhood =

$$x_m \log p + (N - x_m)\log(1-p)$$

MLE

$$\frac{\partial}{\partial p}\left(x_m \log p + (N-x_m)\log(1-p)\right) = 0$$

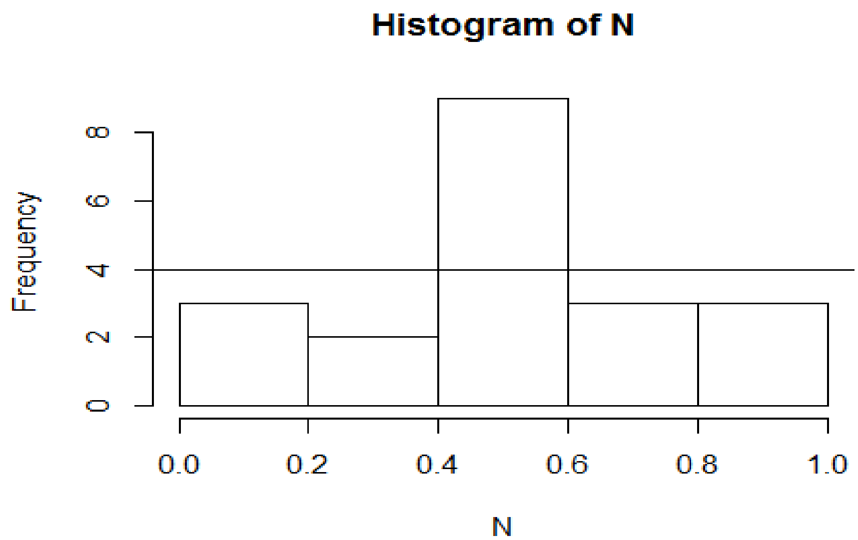
5) $$P(G=1 | B \geq 1) = \frac{P(G=1, B=1)}{P(B \geq 1)} = \frac{1/2}{1/2 + 1/4} = 4/6 = \frac{2}{3}$$

6) $$P(G=1 | B=1) = \frac{P(B=1 | G=1) \cdot P(G=1)}{P(B=1)} = \frac{1/2 \cdot 3/4}{3/4} = 1/2$$

Question 7

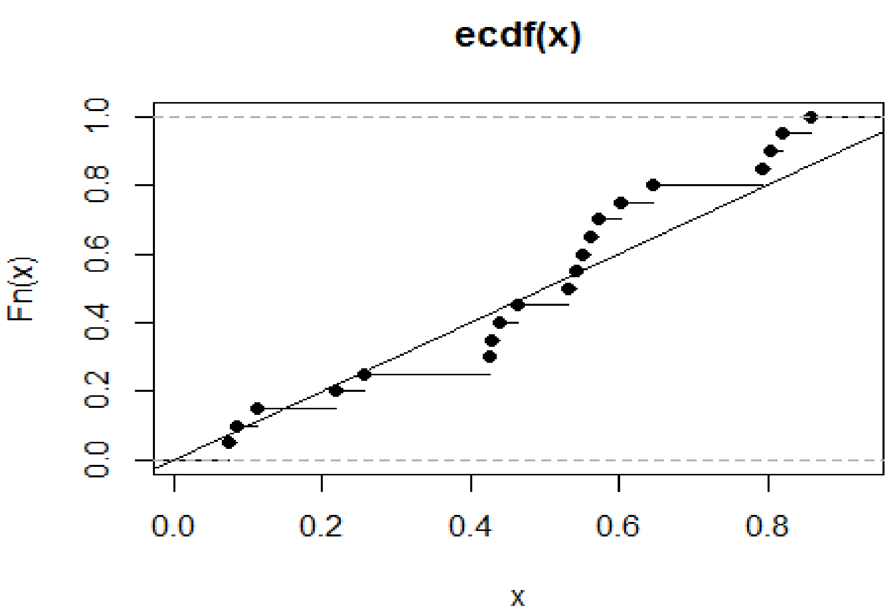
```
N=runif(20,0,1)

hist(N,breaks = 5)
abline(h=4)
```



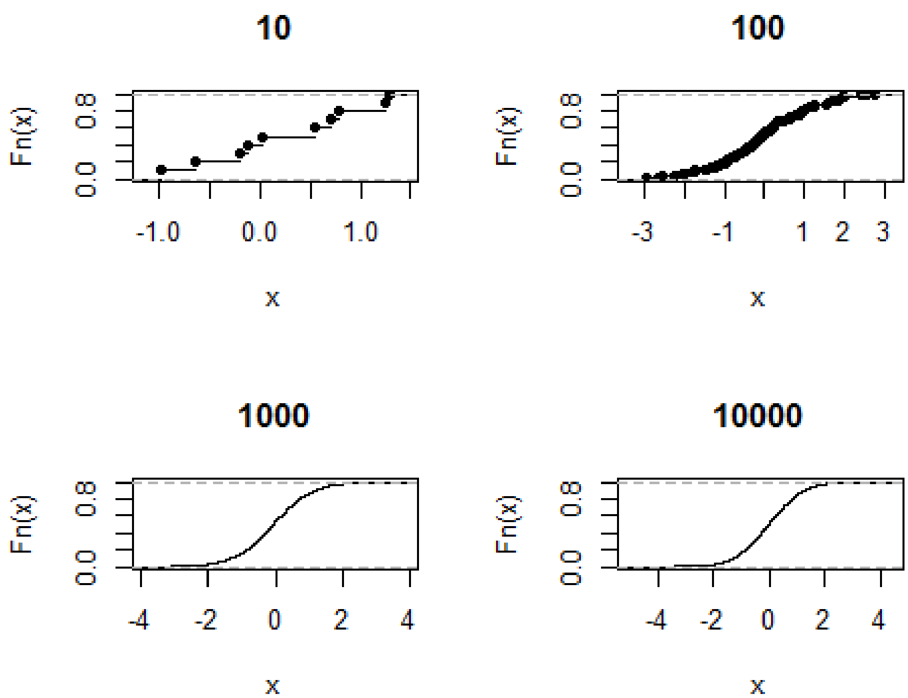

```
## It is expected when we get a uniform distrubution we expect
## they will be equally divided into the no breaks

plot.ecdf(N)
abline(0,1)
```

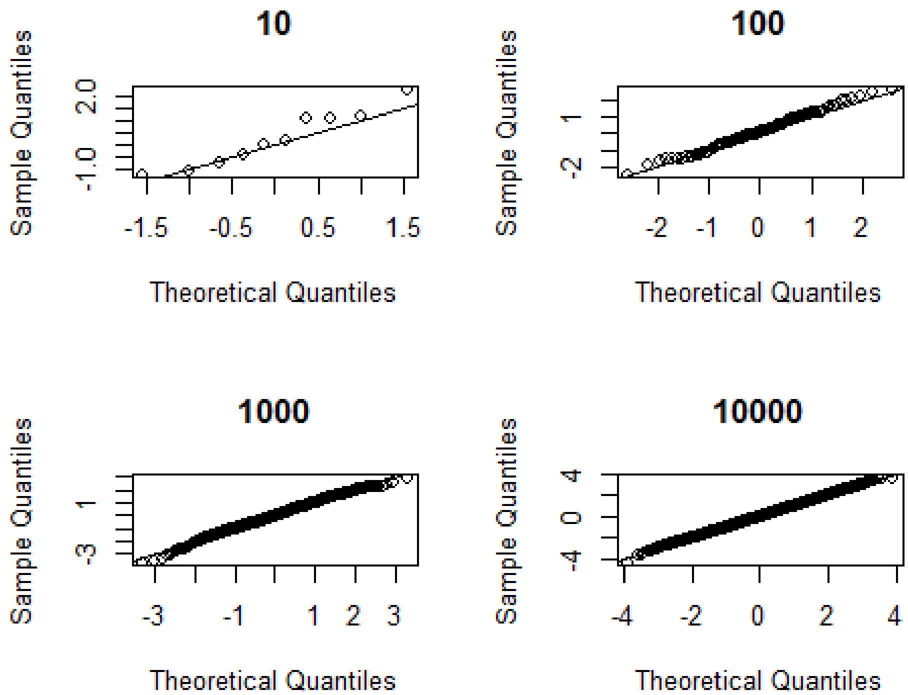


```
## It is expected when we get a uniform distrubution we expect
## when the horizontal line is integrated it will be a str. line with slope 1

par(mfrow=c(2,2))
for (N in c(10,100,1000, 10000)){
  x=rnorm(N)
  plot(ecdf(x), main=N)
}
```



```r
par(mfrow=c(2,2))
for (N in c(10,100,1000, 10000)){
  x=rnorm(N)
  qqnorm(x,main=N)
  abline(0,1)
}
```



Section 4-Variance and Covariance

$$
\begin{aligned}
(8)\quad \mathrm{var}[f] &= \mathbb{E}[(f(x) - \mathbb{E}(f(x)))^2] \\
&= \mathbb{E}[(f(x))^2 + (\mathbb{E}(f(x)))^2 - 2f(x)\,\mathbb{E}(f(x))] \\
&= \mathbb{E}(f(x))^2 + \mathbb{E}(f(x))^2 - 2\,\mathbb{E}(f(x))^2) \\
&= \underline{\underline{\mathbb{E}(f(x)^2) - \mathbb{E}(f(x))^2}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{cov}[x,y] &= \mathbb{E}[\{x - \mathbb{E}(x)\}\{y - \mathbb{E}(y)\}] \\
&= \mathbb{E}[xy - y\,\mathbb{E}(x) - x\,\mathbb{E}(y) + \mathbb{E}(x)\cdot\mathbb{E}(y)] \\
&= \mathbb{E}(xy) - \cancel{\mathbb{E}(y)}\cdot\mathbb{E}(x) - \mathbb{E}(x)\,\mathbb{E}(y) \\
&\qquad + \cancel{\mathbb{E}(x)}\cdot\mathbb{E}(y) \\
&= \underline{\underline{\mathbb{E}(xy) - \mathbb{E}(x)\cdot\mathbb{E}(y)}}
\end{aligned}
$$

$$\mathrm{cov}[x_i, x_j] = \mathbb{E}_{x,y}[x_i x_j] - \mathbb{E}[x_i]\,\mathbb{E}[x_j]$$

$$\mathrm{cov}[\vec{x}, \vec{x}] = \mathbb{E}_{xy}[\vec{x}\vec{x}^T] - \mathbb{E}[\vec{x}]\cdot\mathbb{E}[\vec{x}^T]$$

$$(\mathrm{cov}[\vec{x}])_{ij} = \mathbb{E}_{xy}[x_i x_j] - \mathbb{E}[x_i]\cdot\mathbb{E}[x_j]$$

Section 7- Homework

(14) $I\{0, 1\} \to$ identical

not identical

$S = \{0, 1\} \to$ same sex

not same sex

$P(I=1) = 1/3$ $\quad$ $P(S=1, I=0) = 1/2$

$P(S=1, I=1) = 1$ $\quad$ $P(S=0, I=0) = 1/2$

$$P(I=1|S=1) = \frac{P(S=1|I=1) \cdot P(I=1)}{P(S=1)}$$

$$= \frac{1 \cdot 1/3}{1*1/3 + 1/2*2/3} = 1/2$$