CSE 250 A

ASSIGNMENT - 1

PRERNA BHAVSAR

PID A59002324

1.1 Conditioning on background evidence

(a) Prove $P(X,Y|E) = P(X|Y,E) \cdot P(Y|E)$

$\Rightarrow$ Product Rule $\Rightarrow P(X,Y,E) = P(E) \cdot P(X,Y|E)$

$\therefore P(X,Y|E) = \dfrac{P(X,Y,E)}{P(E)}$ —— ①

LHS $= P(X,Y|E)$

$= \dfrac{P(X,Y,E)}{P(E)}$ —— From ①

$= \dfrac{\cancel{P(E)} \cdot P(Y|E) \cdot P(X|Y,E)}{\cancel{P(E)}}$ [Product Rule]

$= P(Y|E) \cdot P(X|Y,E)$

$=$ RHS

Hence Proved, $P(X,Y|E) = P(X|Y,E) \cdot P(Y|E)$

(b) Prove $P(X|Y,E) = \dfrac{P(Y|X,E) \cdot P(X|E)}{P(Y|E)}$

Product Rule $\Rightarrow P(X,Y,E) = P(E) \cdot P(Y|E) \cdot P(X|Y,E)$

$\therefore P(X|Y,E) = \dfrac{P(X,Y,E)}{P(E) \cdot P(Y|E)}$ —— ①

LHS $= \dfrac{P(X,Y,E)}{P(E) \cdot P(Y|E)}$ [From ①]

$= \dfrac{P(E) \cdot P(X|E) \cdot P(Y|X,E)}{P(E) \cdot P(Y|E)}$ [Product Rule]

$= \dfrac{P(X|E) \cdot P(Y|X,E)}{P(Y|E)} =$ RHS

Hence Proved $P(X|Y,E) = \dfrac{P(Y|X,E) \cdot P(X|E)}{P(Y|E)}$

(1b) Prove $P(X|E) = \sum_y P(X, Y=y | E)$

$\Rightarrow$ Product Rule $\Rightarrow$ $P(X,E) = P(E) \cdot P(X|E)$

$\therefore \; P(X|E) = \dfrac{P(X,E)}{P(E)}$

$$
\begin{aligned}
LHS &= P(X|E) \\
&= \frac{P(X,E)}{P(E)} \\
&= \sum_y \frac{P(X, Y=y, E)}{P(E)} \quad \text{[Marginalization]} \\
&= \sum_y \frac{P(E) \cdot P(X, Y=y | E)}{P(E)} \quad \text{[Product Rule]} \\
&= \cancel{P(E)} \sum_y \frac{P(X, Y=y | E)}{\cancel{P(E)}} \quad \text{[E is not dependent on Y]} \\
&= \sum_y P(X, Y=y | E) \\
&= RHS
\end{aligned}
$$

Hence Proved. $P(X|E) = \sum_y P(X, Y=y | E)$

1.2 Conditional Independence

1. To show that (3) implies (1)

Consider (3) $\Rightarrow$ $P(Y|X,E) = P(Y|E)$ —— (A)

To Prove : (1) $\Rightarrow$ $P(X,Y|E) = P(X|E) \cdot P(Y|E)$

$P(X,Y,E) = P(E) \cdot P(X,Y|E)$ [Product Rule]

$\therefore$ $P(X,Y|E) = \dfrac{P(X,Y,E)}{P(E)}$ —— (1)

$$\begin{aligned}
LHS &= P(X,Y|E) \\
&= \frac{P(X,Y,E)}{P(E)} && \text{[From (1)]} \\
&= \frac{\cancel{P(E)} \cdot P(X|E) \cdot P(Y|X,E)}{\cancel{P(E)}} && \text{[Product Rule]} \\
&= P(X|E)\,P(Y|E) && \text{[From (A)]} \\
&= RHS
\end{aligned}$$

$\therefore$ Given (3) implies (1)
i.e. $P(Y|X,E) = P(Y|E)$ implies
$P(X,Y|E) = P(X|E) \cdot P(Y|E)$

2. To show that (1) implies (2)

Given (1) $\Rightarrow$ $P(X,Y|E) = P(X|E) \cdot P(Y|E)$ — (A)

To Prove (2) $\Rightarrow$ $P(X|Y,E) = P(X|E)$

$P(X,Y,E) = P(E) \cdot P(Y|E) \cdot P(X|Y,E)$ [Product Rule]

$\therefore \quad P(X|Y,E) = \dfrac{P(X,Y,E)}{P(E) \cdot P(Y|E)}$ —— (B)

$$LHS = P(X|Y,E)$$

$$= \frac{P(X,Y,E)}{P(E) \cdot P(Y|E)} \quad \text{[From (B)]}$$

$$= \frac{\cancel{P(E)} \cdot P(X,Y|E)}{\cancel{P(E)} \cdot P(Y|E)} \quad \text{[Product Rule]}$$

$$= \frac{P(X|E) \cdot P(Y|E)}{P(Y|E)} \quad \text{[From (A)]}$$

$$= P(X|E)$$

$$= RHS$$

$\therefore$ We can prove that given Equation (1) implies Equation (2).

i.e. given $P(X,Y|E) = P(X|E) \cdot P(Y|E)$
implies $P(X|Y,E) = P(X|E)$

3. To show that (2) implies (3)

Given (2) $\Rightarrow$ $P(X|Y,E) = P(X|E)$ —— (A)

To Prove (3) $\Rightarrow$ $P(Y|X,E) = P(Y|E)$

$P(X,Y,E) = P(E) \cdot P(X|E) \cdot P(Y|X,E)$ [Product Rule]

$\therefore$ $P(Y|X,E) = \dfrac{P(X,Y,E)}{P(E) \cdot P(X|E) \cdot \cancel{P(Y|X,E)}}$

$$
\begin{aligned}
LHS &= P(Y|X,E) \\
&= \frac{P(X,Y,E)}{P(E) \cdot P(X|E)} \\
&= \frac{\cancel{P(E)} \cdot P(Y|E) \cdot P(X|Y,E)}{\cancel{P(E)} \cdot P(X|E)} \quad \text{[Product Rule]} \\
&= \frac{P(Y|E) \cdot \cancel{P(X|E)}}{\cancel{P(X|E)}} \quad \text{[From (A)]} \\
&= P(Y|E) \\
&= RHS
\end{aligned}
$$

$\therefore$ We have proved that given (2) implies (3).

i.e. $P(X|Y,E) = P(X|E)$ implies

$P(Y|X,E) = P(Y|E)$

$\therefore$ It is Proved that from 1, 2, 3. Solutions that

(1) implies (2) & (3)

(2) implies (1) & (3) and

(3) implies (2) & (1)

1.3. (a) Cumulative Evidence:

X = Your skin is sunburned.
Y = You did not apply sunscreen
Z = You went to the beach.

$\therefore$ ~~Proba~~ $P(X) < P(X=1|Y=1) < P(X=1|Y=1, Z=1)$.

(b) Explaining away:

$$P(X=1|Y=1) > P(X=1)$$

X = Swimming across the pool.
Y = Having floaters on.

$\therefore$ The probability of swimming across a pool increases when we know that the person is wearing floaters

Now Z = Person does not know how to swim.

$\therefore P(X|Y=1, Z=1) < P(X=1|Y=1)$
Because probability of swimming across a pool knowing the person has floaters but does not know how to swim ~~decreases~~ is less than just knowing that the person has floaters on.

(c) Conditional Independence:

$$P(X=1, Y=1) \neq P(X=1) \cdot P(Y=1)$$
$$P(X=1, Y=1 \mid Z=1) = P(X=1 \mid Z=1) \cdot P(Y=1 \mid Z=1)$$

X = Person can see the sun.
Y = It is Raining
Z = It is a cloudy day.

∴ Probability of seeing the sun depends on the knowledge of a cloudy day. Likewise, probability of it raining is dependent on whether it is a cloudy day.

X and Y are not independent events as the probability of seeing the sun is dependent on whether it is raining or not.

1.4 Let D denote performance enhancing drugs by cyclists.

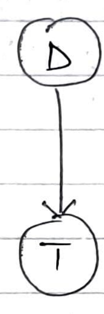
D = 1 denote use of performance enhancing drugs
D = 0 denote no use of performance enhancing drugs.
T denote outcome of drug test.

(a) $\therefore$ $P(D=1) = 0.01$
$P(D=0) = 0.99$.

given D, T is conditionally dependent on D.

Belief Network

$\boxed{P(D=1) = 0.01}$



D	$P(T=1/D)$

?

Given: False Positive Rate (FPR) = 5%.
$\therefore$ True Negative Rate (TNR) = 95%.

Given: False negative Rate (FNR) = 10%
True Positive Rate (TPR) = 90%.

$\therefore$ ~~P(T=0~~ $P(T=1 \mid D=0) = 0.05 = FPR$.
$P(T=1 \mid D=1) = 0.9 = TPR$.

$\therefore$ $P(T=0 \mid D=0) = TNR = 0.95$
$P(T=0 \mid D=1) = FNR = 0.1$

$\therefore$ Belief Network :

$P(D=1) = 0.01$

D	$P(T=1 \mid D)$
0	0.05
1	0.9

(b) To Find : $P(D=0 \mid T=0)$

From Bayes Rule,

$$P(D=0 \mid T=0) = \frac{P(T=0 \mid D=0) \cdot P(D=0)}{P(T=0)}$$

We know $P(T=0 \mid D=0) = 0.95$ and $P(D=0) = 0.99$ from results of (a)

$$\therefore \quad P(T=0) = \sum_d P(T=0, D=d) \quad \text{[Marginalization]}$$

$$= \sum_d P(D=d) \cdot P(T=0 \mid D=d)$$

[Product rule]

$$= P(D=0) \cdot P(T=0 \mid D=0) + P(D=1) \cdot P(T=0 \mid D=1)$$

$$= 0.99 \times 0.95 + 0.01 \times \cancel{0.9}\ 0.1$$

$$= 0.9415 \quad \text{— (A)}$$

$$\therefore \quad P(D=0 \mid T=0) = \frac{0.95 \times 0.99}{0.9415} \quad \text{[From A]}$$

$$= \boxed{0.9989}$$

(c) To find : $P(D=1 \mid T=1)$

From Bayes Rule,

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

We know, $P(T=1 \mid D=1) = 0.9$ and $P(D=1) = 0.01$ from (a)

$$\therefore \; P(T=1) = \sum_d P(T=1, D=d) \quad \text{[Marginalization]}$$

$$= \sum_d P(D=d) \cdot P(T=1 \mid D=d)$$

$$= P(D=0) \cdot P(T=1 \mid D=0) + P(D=1) \cdot P(T=1 \mid D=1)$$

$$= 0.99 \times 0.05 + 0.01 \times 0.9$$

$$= 0.0585 \quad \text{--- (A)}$$

$$\therefore \; P(D=1 \mid T=1) = \frac{0.9 \times 0.01}{0.0585} \quad \text{[From A]}$$

$$= \boxed{0.1538}$$

1.5

(a) $P(X = x_i) = p_i$

Entropy $H[X] = -\sum_{i=1}^{n} p_i \log p_i$

To show that $H[X]$ is maximized when $p_i = \frac{1}{n}$ for all i

Constraint $\Rightarrow \quad \sum p_i = 1$

$\sum p_i - 1 = 0 \quad —— \text{(A)}$

Computing gradient with respect to p_i and using Lagrange multipliers to enforce constraint (A),

$$L = -\sum_{i=1}^{n} p_i \log p_i + \left(\sum_i p_i - 1\right)\lambda$$

Taking derivative w.r.t p_i.

$$\frac{\delta L}{\delta p_i} = -(\log p_i + 1) + 1 \cdot \lambda$$

To maximize. $\frac{\delta L}{\delta p_i} = 0$

$\therefore \quad -\log p_i - 1 = \lambda$

$\log p_i = \lambda - 1$

$\therefore \quad p_i = e^{\lambda - 1} \quad —— \text{(B)}$

Constraint $\sum p_i = 1$

$\sum_{i=1}^{n} e^{\lambda - 1} = 1 \qquad$ [From B]

$n \cdot e^{\lambda - 1} = 1$

$\therefore \quad e^{\lambda - 1} = \frac{1}{n} \quad —— \text{(C)}$

Substituting (C) in (B) we get,

$$p_i = e^{\lambda - 1}$$

$$p_i = \frac{1}{n} \quad \text{for all } i.$$

$\therefore$ $H[X]$ is maximized at $p_i = \frac{1}{n}$ for all i.

(b) $$H(X_1, X_2, \cdots X_n) = -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} P(x_1, x_2, \cdots x_n) \cdot \log P(x_1, x_2, \cdots x_n)$$

If X_i variables are independent, show that

$$P(x_1, x_2, \cdots x_n) = \prod_{i=1}^{n} P(x_i) \quad \text{(A)}$$ implies

$$H(X_1, X_2, \cdots X_n) = \sum_{i=1}^{n} H(X_i).$$

$$H(X_1, X_2 \cdots X_n) = -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} P(x_1, x_2, \cdots x_n) \cdot \log P(x_1, x_2 \cdots x_n)$$

$$= -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} P(x_1, x_2, \cdots x_n) \left[\log (P(x_1) \cdot P(x_2) \cdots P(x_n)) \right]$$

[From (A)]

$$= -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} P(x_1, x_2 \cdots x_n) \left[\log P(x_1) + \log P(x_2) + \cdots + \log P(x_n) \right]$$

Considering summation over x_1 term.

$$-\sum_{x_1} P(x_1, x_2, \cdots x_n) \cdot \log P(x_1) = -\sum_{x_1} P(x_1) \cdot P(x_2) \cdots P(x_n) \cdot \log P(x_1)$$

$$= -\sum_{x_1} P(x_1) \log P(x_1)$$

because x_i are independent variables.

$$\therefore -\sum_{x_1} P(x_1, x_2 \cdots x_n) \log P(x_1) = -\sum_{x_1} P(x_1) \log(P(x_1))$$
$$= H[x_1]$$

Similarly the summation of $x_2 \cdots x_n$ will be equal to $H[x_i]$ as.

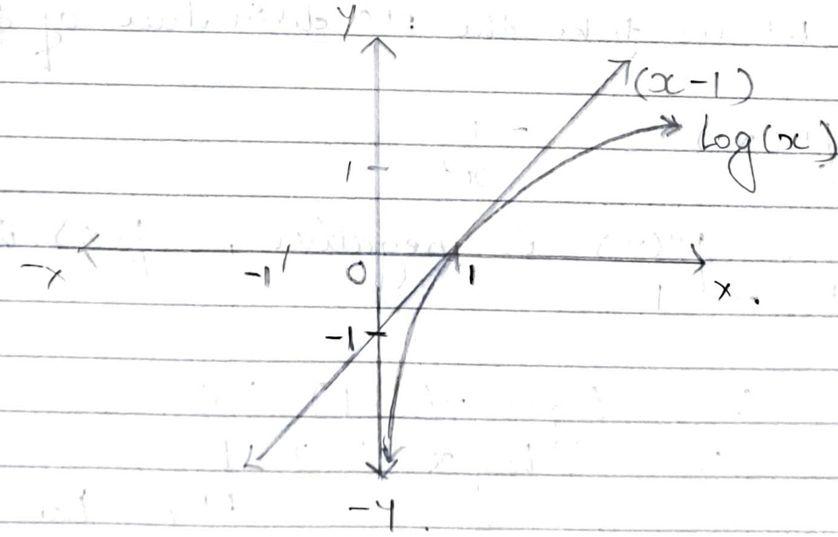
$$H[x_1] = -\sum_{x_1} P(x_1) \log(P(x_1))$$

$$\therefore H(x_1, x_2 \cdots x_n) = \sum_{i=1}^{n} H[x_i]$$

1.6. Kullback-Leibler Distance.

$$KL(p,q) = \sum_i p_i \log(p_i/q_i)$$

(a)



(i) ⇒ Graph for $(x-1)$ and $\log(x)$

From graph (i) we can verify that,

$$\log x \leq x - 1$$

with the equality if and only if $x = 1$.

To prove : $\log x \le x-1$ by differentiation.

$f(x) = \log(x) - (x-1)$

differentiating w.r.t x,

$f'(x) = \frac{1}{x} - 1$

$\therefore f'(x) = 0 = \frac{1}{x} - 1$

$\frac{1}{x} = 1$

$\therefore \boxed{x = 1}$

Hence $\log x = x-1$ when $x = 1$.

Now let us take the 2nd derivative of $f(x)$ w.r.t x,

$f''(x) = -\frac{1}{x^2}$

Since $f''(x)$ is negative, $f(x)$ is maximum at $x = 1$

$\therefore \log x - (x-1) \le 0$

$\therefore \log x \le x - 1$

Hence Proved

$\doteq$

(b.) given $\log x \leq x-1$ — (A)

To prove $KL(p,q) \geq 0$ if and only if p_i and q_i are equal.

$$KL(p,q) = \sum_{i} p_i \log(p_i/q_i)$$

$$= -\sum_i p_i \log\left(\frac{q_i}{p_i}\right) \quad \text{— (B)}$$

From (A) : $\log x \leq x-1$

$$-\log x \geq -(x-1)$$

Substituting $x = \frac{q_i}{p_i}$.

$$-\log \frac{q_i}{p_i} \geq -\left(\frac{q_i}{p_i} - 1\right) \quad \text{— (C)}$$

From (C), we can rewrite (B) as,

$$KL(p,q) \geq -\sum_i p_i \left(\frac{q_i}{p_i} - 1\right)$$

$$\geq \sum_i p_i - q_i$$

$$\geq \sum_i p_i - \sum_i q_i$$

We know $\sum_i p_i = \sum_i q_i = 1$

$\therefore KL(p,q) \geq 1-1$

$$\boxed{KL(p,q) \geq \underline{0}}$$

(c) Given : $\log x \leq x - 1$ and $\log x = 2 \log \sqrt{x}$

Derive : $KL(p,q) \geq \sum_i (\sqrt{p_i} - \sqrt{q_i})^2$

$\Rightarrow$ ~~$KL(q,p)$~~

$$KL(p,q) = -\sum_i p_i \log\left(\frac{q_i}{p_i}\right)$$

By using $\log x = 2 \log \sqrt{x}$, we can rewrite $KL(p,q)$ as,

$$KL(p,q) = -\sum_i p_i \, 2 \log \sqrt{\frac{q_i}{p_i}}$$

$$= -2 \sum_i p_i \log \sqrt{\frac{q_i}{p_i}} \quad \text{— (A)}$$

Now from $\log x \leq x - 1$

$\Rightarrow -\log x \geq -(x-1)$

Substituting $x = \sqrt{\frac{q_i}{p_i}}$ in the equation

$$\Rightarrow -\log \sqrt{\frac{q_i}{p_i}} \geq -\left(\sqrt{\frac{q_i}{p_i}} - 1\right) \quad \text{— (B)}$$

From (B) we can rewrite (A) as,

$$KL(p,q) \geq -\sum_i 2p_i \left(\sqrt{\frac{q_i}{p_i}} - 1\right)$$

$$\geq -\sum_i 2\sqrt{q_i p_i} - 2p_i$$

$$\geq \sum_i 2p_i - 2\sqrt{q_i p_i}$$

But, $\sum_i 2p_i = \sum p_i + p_i = \sum_i p_i + \sum_i p_i = \sum_i p_i + \sum_i q_i$

$\therefore \quad KL(p,q) \geq \sum_i p_i + q_i - 2\sqrt{p_i q_i}$

$$\boxed{KL(p,q) \geq \sum_i (\sqrt{p_i} - \sqrt{q_i})^2}$$

(d) To show that $KL(p,q) \neq KL(q,p)$

Consider $p_1 = 0.4$ and $q_1 = 0.1$
$p_2 = 0.6$ $q_2 = 0.9$

$$KL(p,q) = \sum p_i \log(p_i / q_i)$$

$$= 0.4 \log\left(\frac{0.4}{0.1}\right) + 0.6 \log\left(\frac{0.6}{0.9}\right)$$

$$= 0.135$$

$$KL(q,p) = \sum q_i \log(q_i / p_i)$$

$$= 0.1 \log\left(\frac{0.1}{0.4}\right) + 0.9 \log\left(\frac{0.9}{0.6}\right)$$

$$= 0.098$$

$$\therefore \quad \boxed{KL(p,q) \neq KL(q,p)}$$

1.7 Mutual Information.

(a) To show that if, $I(X,Y) = \sum_x \sum_y P(x,y) \log\left[\frac{P(x,y)}{P(x)\cdot P(y)}\right]$

$I(X,Y) \geq 0$.

Given : $\log x \leq x-1$

$\therefore -\log x \geq -(x-1)$

Substituting $x = \cancel{P(x,y)}\ \frac{P(x)\cdot P(y)}{P(x,y)}$

$$-\log\left[\frac{P(x)\cdot P(y)}{P(x,y)}\right] \geq -\left[\frac{P(x)\cdot P(y)}{P(x,y)} - 1\right] \quad \text{— (A)}$$

Now $I(X,Y) = \sum_x \sum_y P(x,y) \log\left[\frac{P(x,y)}{P(x)\cdot P(y)}\right]$

$$= -\sum_x \sum_y P(x,y) \log\left[\frac{P(x)\, P(y)}{P(x,y)}\right]$$

From (A)

$$I(X,Y) \geq -\sum_x \sum_y P(x,y)\left[\frac{P(x)\cdot P(y)}{P(x,y)} - 1\right]$$

$$\geq -\sum_x \sum_y \cancel{P(x,y)}\ P(x)\cdot P(y) - P(x,y)]$$

$$\geq \sum_x \sum_y P(x,y) - P(x)\cdot P(y)$$

$$\geq \sum_x \sum_y P(x,y) - \sum_y P(x)\cdot P(y)$$

$$\geq \sum_x P(x) - P(x)\cdot \sum_y P(y)$$

$\therefore \quad I(X,Y) \geq \sum_{x} P(x) - P(x) \cdot 1$

$[\because \sum_{y} P(y) = 1]$

$\boxed{I(X,Y) \geq 0}$ Hence Proved.

(b) To Show that $I(X,Y) = 0$ if X and Y are independent random variables.

$$I(X,Y) = \sum_{x} \sum_{y} \log \left[\frac{P(x,y)}{P(x) \cdot P(y)} \right] \cdot P(x,y)$$

$$= - \sum_{x} \sum_{y} P(x,y) \log \left[\frac{P(x) \cdot P(y)}{P(x,y)} \right]$$

$$= - \sum_{x} \sum_{y} P(x,y) \left[\log \left(P(x) \cdot P(y) \right) - \log \left[P(x,y) \right] \right]$$

$\because$ Because X and Y are independent

$P(X,Y) = P(X) \cdot P(Y)$

$$\therefore \quad I(X,Y) = - \sum_{x} \sum_{y} P(x,y) \left[\log [P(x) \cdot P(y)] - \log [P(x) \cdot P(y)] \right]$$

$\boxed{I(X,Y) = 0}$

1.8 Compare and Contrast

(a) Yes there is a conditional independence implied in belief network 1 which is not implied in second.

In first belief network Y and Z are conditionally independent.

i.e. $P(Y | X, Z) = P(Y | X)$

$P(Z | X, Y) = P(Z | X)$

In the second, Z is conditionally dependent on Y.

$P(Z | X, Y) = P(Z | Y)$.

(b) No, the second belief network does not imply a statement of marginal or conditional independence that is not implied by third

(c) Yes, there is a conditional independence implied in third that is not implied by first.

In third, X and Z are conditionally independent.

i.e. $P(X | Y, Z) = P(X | Y)$

Whereas in first X and Z are conditionally dependent

i.e. Z is dependent on X.

$P(Z | Y, X) = P(Z | X)$

HW1_Hangman

October 5, 2021

[24]:
```
!wget -nc https://raw.githubusercontent.com/brpy/colab-pdf/master/colab_pdf.py
from colab_pdf import colab_pdf
colab_pdf('HW1_Hangman.ipynb')
```

```
File colab_pdf.py already there; not retrieving.


WARNING: apt does not have a stable CLI interface. Use with caution in scripts.


WARNING: apt does not have a stable CLI interface. Use with caution in scripts.

[NbConvertApp] Converting notebook /content/drive/MyDrive/Colab
Notebooks/HW1_Hangman.ipynb to pdf
[NbConvertApp] Writing 38807 bytes to ./notebook.tex
[NbConvertApp] Building PDF
[NbConvertApp] Running xelatex 3 times: [u'xelatex', u'./notebook.tex',
'-quiet']
[NbConvertApp] Running bibtex 1 time: [u'bibtex', u'./notebook']
[NbConvertApp] WARNING | bibtex had problems, most likely because there were no
citations
[NbConvertApp] PDF successfully created
[NbConvertApp] Writing 30778 bytes to /content/drive/My Drive/HW1_Hangman.pdf

<IPython.core.display.Javascript object>


<IPython.core.display.Javascript object>
```

[24]: 'File ready to be Downloaded and Saved to Drive'

[4]:
```
%cd '/content/drive/MyDrive/CSE250A/HW1'
```

```
/content/drive/MyDrive/CSE250A/HW1
```

[5]:
```python
def processing(file):

    word_count = {}
    f = open(file,'r')

    for line in f:
        line = line.split(" ")
        word_count[line[0]] = int(line[1])
    return word_count
```

[6]:
```python
def priorProbability(word_dict):
    word_prob = {}
    total  = sum(list(word_dict.values()))
    for key, value in word_dict.items():
        word_prob[key] = value / total

    return word_prob
```

[7]:
```python
#initial word count
word_count = processing("hw1_word_counts_05.txt")
#initial word probabilities
word_prob = priorProbability(word_count)
```

[8]:
```python
sorted_word_probabilities = sorted(word_prob.items(), key = lambda x: x[1],␣
 ↪reverse = True)


print("Top 15 most frequent words:")
for i in range(15):
    print(sorted_word_probabilities[i][0])


print("14 least frequent words:")
for i in range(-1, -15, -1):
    print(sorted_word_probabilities[i][0])
```

```
Top 15 most frequent words:
THREE
SEVEN
EIGHT
WOULD
ABOUT
THEIR
WHICH
AFTER
FIRST
FIFTY
OTHER
FORTY
```

```
YEARS
THERE
SIXTY
14 least frequent words:
TROUP
OTTIS
MAPCO
CAIXA
BOSAK
YALOM
TOCOR
SERNA
PAXON
NIAID
FOAMY
FABRI
CLEFT
CCAIR
```

[9]:
```
def calcPosteriorProbs(possible_words, total):
  posterior_prob = {}
  for word, value in possible_words.items():
    posterior_prob[word] =  value/total
  return posterior_prob
```

[10]:
```
def getPossibleLetters(possible_words, correct_guess, wrong_guess):
  possible_letters = set()
  for word in possible_words.keys():
    l = list(map(str, word))
    possible_letters.update(l)
  correct_letters = [l for l, v in correct_guess.items()]
  possible_letters = possible_letters - set(wrong_guess) - set(correct_letters)
  return possible_letters
```

[11]:
```
def calcLetterProb(letters, posterior_probs):
  letter_predictive_prob ={}
  next_guess, val = None, 0

  for letter in letters:
    letter_predictive_prob[letter] = 0
    for word, value in posterior_probs.items():
      if letter in word:
        letter_predictive_prob[letter] += value
    if letter_predictive_prob[letter] > val:
      next_guess = letter
      val = letter_predictive_prob[letter]

  return letter_predictive_prob, next_guess, val
```

[12]:
```python
def getPossibleWords(word_probs, correct_guess, wrong_guess):
  possible_words = {}
  total = 0
  for word,prob in word_probs.items():
    sub = {}
    flag = 0
    for letter in wrong_guess:
      if letter in word:
        flag = 1
        break
    if flag:
      continue

    for i,ch in enumerate(word):
      sub[ch] = sub.get(ch,[]) + [i]
    for letter, value in correct_guess.items():
      if letter not in word:
        flag = 1
        break
      else:
        if sub[letter] != value:
          flag = 1
          break
    if flag:
      continue
    possible_words[word] = prob
    total += prob
  return possible_words, total
```

[13]:
```python
def predictNextLetter(word_count, word_prob, correct_guess, wrong_guess):

  possible_words, total_poss_prob  = getPossibleWords(word_prob, correct_guess,␣
↪wrong_guess)

  word_posterior_prob = calcPosteriorProbs(possible_words, total_poss_prob)
  possible_letters = getPossibleLetters(possible_words, correct_guess,␣
↪wrong_guess)
  letter_probs, next_guess, val = calcLetterProb(possible_letters,␣
↪word_posterior_prob)

  return (next_guess, val)
```

[14]:
```python
samples = []
correct = []
wrong = []

for i in range(9):
  word = input("Enter correctly guessed:")
```

```
  samples.append(word)
  w_guess = input("Enter incorrectly guessed:")
  word_w = w_guess.split(',')
  if word_w[0] =='':
    wrong.append([])
  else:
    wrong.append(word_w)
```

```
Enter correctly guessed:-----
Enter incorrectly guessed:
Enter correctly guessed:-----
Enter incorrectly guessed:E,A
Enter correctly guessed:A---S
Enter incorrectly guessed:
Enter correctly guessed:A---S
Enter incorrectly guessed:I
Enter correctly guessed:--O--
Enter incorrectly guessed:A,E,M,N,T
Enter correctly guessed:-----
Enter incorrectly guessed:E,O
Enter correctly guessed:D--I-
Enter incorrectly guessed:
Enter correctly guessed:D--I-
Enter incorrectly guessed:A
Enter correctly guessed:-U---
Enter incorrectly guessed:A,E,I,O,S
```

[15]:
```
print(samples)
print(wrong)
```

```
['-----', '-----', 'A---S', 'A---S', '--O--', '-----', 'D--I-', 'D--I-',
'-U---']
[[], ['E', 'A'], [], ['I'], ['A', 'E', 'M', 'N', 'T'], ['E', 'O'], [], ['A'],
['A', 'E', 'I', 'O', 'S']]
```

[16]:
```
for word in samples:
  s = {}
  for i,ch in enumerate(word):
    if ch != '-':
      s[ch] = s.get(ch,[])+[i]
  correct.append(s)
print(correct)
```

```
[{}, {}, {'A': [0], 'S': [4]}, {'A': [0], 'S': [4]}, {'O': [2]}, {}, {'D': [0],
'I': [3]}, {'D': [0], 'I': [3]}, {'U': [1]}]
```

[22]:
```
print('correctly guessed incorrectly guessed\tbest next guess Prob')
for i in range(9):
  next_guess, next_prob = predictNextLetter(word_count, word_prob, correct[i],␣
 ↪wrong[i])
  print('{:20s}{:20s}{:20s}{:4.4f}'.format(samples[i],' '.join(wrong[i]),␣
 ↪next_guess, next_prob))
```

```
correctly guessed incorrectly guessed   best next guess Prob
-----                                   E                   0.5394
-----               E A                 O                   0.5340
A---S                                   E                   0.7715
A---S               I                   E                   0.7127
--O--               A E M N T           R                   0.7454
-----               E O                 I                   0.6366
D--I-                                   A                   0.8207
D--I-               A                   E                   0.7521
-U---               A E I O S           Y                   0.6270
```