Foundations of Probabilistic of Information Retrieval
Assignment 2
Rui Tang - 10010250
13 Nov 2017

Problem 1
(a)

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}_{n\times 1} \qquad A = \begin{pmatrix} a_{11} & \cdots & a_{1d} \\ a_{21} & \cdots & a_{2d} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nd} \end{pmatrix}_{n\times d} \qquad \cancel{y = z^TA = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}}$$

$$y = z^TA = (z_1\ z_2 \cdots z_n) \begin{pmatrix} a_{11} & \cdots & a_{1d} \\ a_{21} & \cdots & a_{2d} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nd} \end{pmatrix} = (\underbrace{z_1a_{11} + z_2a_{21} + \cdots + z_na_{n1}}_{y_1}, \cdots, \underbrace{z_1a_{1d} + z_2a_{2d} + \cdots + z_na_{nd}}_{y_d})$$

$$= (y_1, \cdots, y_d)_{1\times d}$$

$P(X=i) = s_i \ , i \in [1, d]$

$s = (s_1, \cdots, s_d) = \text{softmax}(y)$

The softmax of v: $\text{softmax}(v)_i = \dfrac{\exp(v_i)}{\sum_{j=1}^{k} \exp(v_j)}$

so the softmax of d-dimensional vector y is defined as:

$$\text{softmax}(y)_i = \frac{\exp(y_i)}{\sum_{j=1}^{k} \exp(y_j)}$$

Because we have

$P(X=i) = s_i$

$s_i = \text{softmax}(y)_i$

so $P(X=i) = \text{softmax}(y)_i = \dfrac{\exp(y_i)}{\sum_{j=1}^{k} \exp(y_j)}$

(b)

$$P(X=i)=\frac{\exp(y_i)}{\sum_{j=1}^{K}\exp(y_j)}\quad , Q(X=i)=r_i \quad r=(r_1,\dots,r_d)$$

The cross-entropy is

$$H(Q;P)=-E_{X\sim Q}\log(P(X))=-\sum_{i=1}^{d}Q(x_i)\log P(x_i) \cancel{= -\sum_{i=1}^{d} r_i \cdot}$$

$$=-\sum_{i=1}^{d} r_i\left(y_i-\left(\log\sum_{j=1}^{K}\exp(y_j)\right)\right)$$

We should find y_K that is ~~minimal~~ minimum in y.

If y_K is minimum, r_K should be 1. Then the cross-entropy $H(Q;P)$ is minimal.

Problem 2

(a)

We can show the log-likelihood of this:

$$l(P_1,P_2,\cancel{P_3}\cdots P_K)=\log n!-\sum_{i=1}^{K}\log x_i!+\sum_{i=1}^{K}x_i\log P_i$$

We should take the constraint into account by using the Lagrange multipliers.

$$L(P_1,P_2,\dots,P_K,\lambda)=l(P_1,P_2,\dots,P_K)+\lambda\left(1-\sum_{i=1}^{K}P_i\right)$$

By posting all the derivatives to be 0, we get the most natural estimate

$$\hat{P}_i=\frac{x_i}{n}$$

(b)

We first produce a sample:

In [2]:
```
TRUE_LAMBDA = 6
X = np.random.poisson(TRUE_LAMBDA, 10000)
```

For our sample, we estimate a value for λ using MLE:

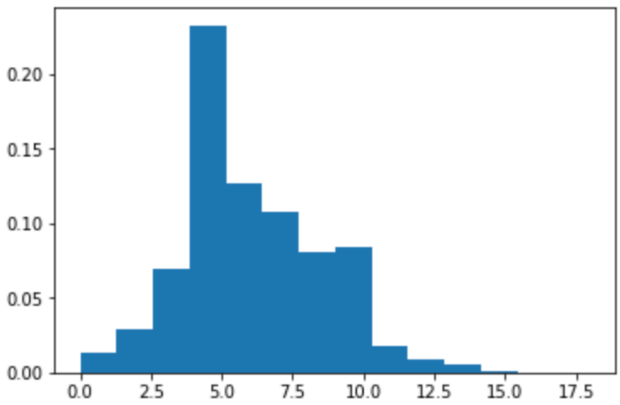
In [3]:
```
def poisson_lambda_MLE(X):
    # TODO
    # total number of observations
    n = X.size
    x = np.ones(n)
    lamda = x.dot(X)/n
    return lamda

lambda_mle = poisson_lambda_MLE(X)
lambda_mle
```

Out[3]: 6.0266000000000002

We finally plot the sample and the resulting distribution:

In [4]:
```
# TODO
count, bins, ignored = plt.hist(X, 14, normed=True)
plt.show()
```



Problem 3

The log likelihood of the Poisson distribution is:

$$l(\lambda) = \log \prod_{i=1}^{K} \left(\frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \right) = \sum_{i=1}^{K} \log \left(\frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \right)$$

$$= -K\lambda + \left(\sum_{i=1}^{K} x_i \right) \log \lambda - \sum_{i=1}^{K} \log (x_i!)$$

We find the maximum by finding the derivative:

$$\frac{\partial}{\partial \lambda} l(\lambda) = -K + \left(\sum_{i=1}^{K} x_i \right) \frac{1}{\lambda} = 0$$

which implies that the estimate should be:

$$\hat{\lambda} = \frac{1}{K} \sum_{i=1}^{K} x_i$$

Problem 4

materiell — -bash — 114×30

```
tangruideMacBook-Pro:materiell loisdawn$
tangruideMacBook-Pro:materiell loisdawn$ python multinomial_solution_rui_tang.py "shakespeare.txt" "stopwords.txt"
estimated probabilities:
p_1     ,       0.44781
p_2     .       0.19983
p_3     ;       0.09453
p_4     :       0.05768
p_5     !       0.05760
p_6     ?       0.05580
p_7     thou    0.02925
p_8     thy     0.02086
p_9     shall   0.01966
p_10    thee    0.01698

enter observation:
X_1=1000
X_2=990
X_3=800
X_4=700
X_5=450
X_6=440
X_7=400
X_8=300
X_9=200
X_10=100

result: {',': 0.18587360594795538, '.': 0.18401486988847585, ';': 0.14869888475836432, ':': 0.13011152416356878, '
!': 0.08364312267657993, '?': 0.08178438661710037, 'thou': 0.07434944237918216, 'thy': 0.055762081784386616, 'shal
l': 0.03717472118959108, 'thee': 0.01858736059479554}
tangruideMacBook-Pro:materiell loisdawn$
```