Quiz 1

1. We wish to classify animals based on how far they jump. Presume we use an exponential likelihood for distance $d$. It’s probability function follows the Exponential distribution:

$$P(x|\lambda) = \lambda e^{-\lambda x}$$

We wish to classify solely based on this feature. What is the maximum likelihood estimate of parameter $\sigma$ for class animal=rabbit, to maximize:

$$L = \prod_{i \in \text{rabbit}} P(d_i|\lambda)$$

(Show your work for partial credit!)

$$\ln L = \sum_{i \in \text{rabbit}} \ln P(d_i|\lambda) = \sum_{i \in \text{rabbit}} \ln (\lambda e^{-\lambda x_i}) = \sum_{i \in \text{rabbit}} [\ln(\lambda) - \lambda x_i]$$

Derivative: $\sum_{i \in \text{rabbit}} \left[ \frac{1}{\lambda} - x_i \right] = 0 \rightarrow \lambda = \frac{N}{\sum_{i \in \text{rabbit}} x_i}$

2. We wish to classify a hospital patient as heart-attack-risk or NOT-heart-attack-risk based on ten numeric features (e.g., including blood pressure, glucose level, body temperature). We will use a Naïve Bayes classification with Gaussian likelihood $P(x|y)$. How many parameters do we need to learn for this Gaussian Naïve Bayes problem?

$# \text{ classes x } # \text{ params/feat x } # \text{ feats } + \text{ prior } = 2 \times 2 \times 10 = 40 + 1 = 41$

Each class has mean and standard deviation; no joint variances (co-variances) computed since Naive
3. Consider three multi-valued random variables $A$ (age), $H$ (height), and $S$ (salary). We know that $S$ is independent of $A$ and $H$; $A$ and $H$ are NOT independent of one another. We are provided the probability tables/functions for the following four joint, marginal, and conditional probabilities.

The four probabilities:

\[
\begin{align*}
    P(A) & \quad P(S,H) \\
    P(A\mid H) & \quad P(H)
\end{align*}
\]

For example, we are told:

$P(A=\text{young}) = 0.2$, $P(A=\text{middleAge})=0.5$, $P(A=\text{old})=0.3$

We are not provided any other probability tables.

Explain how to combine the four probabilities from above (and the knowledge that $S$ is independent) to compute each probability below, or write “not possible” if it is not possible.

a) $P(H=\text{tall}\mid A=\text{young}) = P(A=\text{young}, H=\text{tall})/P(A=\text{young}) = (P(A=\text{young}\mid H=\text{tall})xP(H=\text{tall})) / P(A=\text{young})$

b) $P(S=\text{rich}, A=\text{middleAge}) = P(S=\text{rich}) x P(A=\text{middleAge}) = P(A=\text{middleAge})x\sum_h P(S = \text{rich}, H = h)$

Several other solutions possible here

c) $P(H=\text{short} \mid S=\text{middleIncome}) = (P(H=\text{short}) * P(S=\text{middleIncome})) / P(S=\text{middleIncome}) = P(H=\text{short}) / P(S=\text{middleIncome})$

Several other solutions possible here

4. Two vectors are orthogonal if they have a zero dot product: $a \cdot b = 0$. Find a unit-magnitude orthogonal vector for each of the two following vectors.

\[
v^1 = \begin{bmatrix} 2 \\ -4 \\ 0 \\ -3 \end{bmatrix} \quad v^2 = \begin{bmatrix} -0.5 \\ 0.5 \\ -0.5 \\ 0.5 \end{bmatrix}
\]

Remember each vector $a$ must have unit magnitude here, $|a|=1$

\[
a^1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad a^2 = \begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}
\]