CS5691: Pattern recognition and machine learning
Quiz - 1
Course Instructor : Prashanth L. A.
Date : Feb-1, 2019 **Duration** : 30 minutes

I. Multiple Choice Questions

1. Suppose X is uniformly distributed over $[0,5]$ and Y is uniformly distributed over $[0,4]$. If X and Y are independent, then $\mathbb{P}(\max(X,Y) > 3)$ is
   (a) $\frac{9}{20}$
   (b) $\frac{1}{20}$
   (c) $\frac{11}{20}$
   (d) 1

> **Solution:** (c). Use $\mathbb{P}(\max(X,Y) > 3) = 1 - \mathbb{P}(\max(X,Y) \leq 3) = 1 - \mathbb{P}(X \leq 3)\,\mathbb{P}(Y \leq 3)$, and plug in the values for $\mathbb{P}(X \leq 3), \mathbb{P}(Y \leq 3)$ use their distributional information to arrive at the answer.

2. Let $X_i, i = 1,\ldots,4$ be independent Bernoulli r.v.s each with mean $p = 0.1$ and let $S = \sum_{i=1}^{4} X_i$. Then,
   (a) $\mathbb{E}(X_1 \mid S = 2) = 0.1$.
   (b) $\mathbb{E}(X_1 \mid S = 2) = 0.5$.
   (c) $\mathbb{E}(X_1 \mid S = 2) = 0.25$.
   (d) $\mathbb{E}(X_1 \mid S = 2) = 0.75$.

> **Solution:** (b) since
> $$\mathbb{E}(X_1 \mid S=2) = \mathbb{P}(X_1 = 1 \mid S = 2) = \frac{\mathbb{P}(X_1 = 1, S = 2)}{\mathbb{P}(S = 2)} = \frac{p \times 3p(1-p)^2}{\binom{4}{2}p^2(1-p)^2} = 0.5.$$

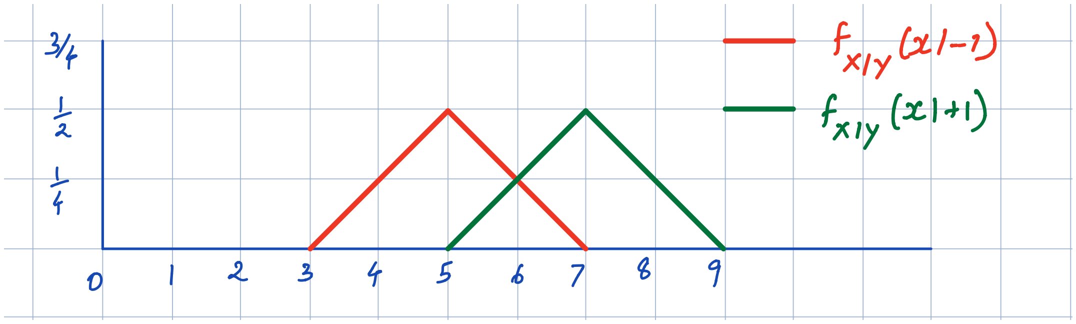
3. Let $v_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, v_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, and $v_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. Let $C(A)$ and $N(A)$ denote the column and null space, respectively of any matrix A. Then, which of the following statements is **false**?
   (a) $v_1, v_2 \in C(A)$, and $v_3 \in N(A)$ for some matrix A.
   (b) $v_1, v_2 \in C(A)$, and $v_3, v_4 \in N(A)$ for some matrix A.
   (c) $v_1 \in C(A)$, and $v_3, v_4 \in N(A)$ for some matrix A.
   (d) $v_1 \in C(A)$, and $v_3 \in N(A)$ for some matrix A.

> **Solution:** (b) since the dimensions of the column space and null space add up to 4, while using the rank-nullity theorem, these dimensions can only sum up to 3.

4. Let $Z = (X, Y)$ be a bivariate normal random variable. Then, which of the following statements is **false**?
   (a) X and Y are independent if and only if they are uncorrelated.
   (b) $X + Y$ is univariate normal.
   (c) $Y \mid X = x$ is distributed as a normal random variable.
   (d) $X + Y$ and $X - Y$ are independent.

> **Solution:** (d)

5. Let $P(Y=-1)=\frac{1}{3}$, and $P(Y=+1)=\frac{2}{3}$. The class-conditionals $P(X|Y)$ are given by the graph below. The Bayes classifier is then given by which option below? (Triangle on left is the class conditional for $Y=-1$).



(a) $h^*(x)=\begin{cases}-1 & \text{if } x\le 6\\ +1 & \text{if } x>6\end{cases}$

(b) $h^*(x)=\begin{cases}-1 & \text{if } x\le \frac{17}{3}\\ +1 & \text{if } x>\frac{17}{3}\end{cases}$

(c) $h^*(x)=\begin{cases}-1 & \text{if } x\le \frac{19}{3}\\ +1 & \text{if } x>\frac{19}{3}\end{cases}$

(d) $h^*(x)=\begin{cases}-1 & \text{if } x> 6\\ +1 & \text{if } x\le 6\end{cases}$

Solution: Options (d) is clearly wrong, while option (a) would be right if the prior probabilities were $\frac{1}{2}$. Now, between options (b) and (c), observe that the former has a threshold to the left of 6, and the latter to the right of 6. Given that prior probabilities make class with label $+1$ more likely, the threshold for Bayes classifier would be to the left of 6, making option (b) the right choice. A more rigorous proof would involve explicit calculatation of the threshold, by using the density for a triangular distribution, together with the prior probabilities.

II. A problem that requires a detailed solution

1. Let X and Y be r.v.s with the joint density given by

$$f(x,y)=\frac{1}{8\sqrt{3}\pi}\exp\left(-\frac{x^2}{6}-\frac{y^2}{24}+\frac{xy}{12}+\frac{x}{12}+\frac{y}{6}-\frac{7}{24}\right). \tag{1}$$

Answer the following: (3+2 marks)

(a) Find the means and variances of X and Y. Also, find the covariance between X and Y.

(b) Find the conditional density of Y given $X=x$. Also, calculate $\mathbb{E}[Y\mid x]$.

Solution: Recall that the bivariate normal density $f(x,y)$ is given by

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left(-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_1}{\sigma_1}\right)^2+\left(\frac{y-\mu_2}{\sigma_2}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right)\right). \tag{2}$$

From the coefficients of x^2, y^2 and xy terms in (1) and (2), we obtain

$$\sigma_1^2(1-\rho^2)=3,\quad \sigma_2^2(1-\rho^2)=12,\quad \sigma_1\sigma_2(1-\rho^2)=6\rho.$$

Further, from the factor outside the exponential, we have

$$2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}=8\sqrt{3}\pi$$
$$\Rightarrow \sigma_1^2\sigma_2^2(1-\rho^2)=48.$$
$$\text{Using } \sigma_1^2(1-\rho^2)=3,\ \text{we obtain } 3\sigma_2^2=48,\ \text{or } \sigma_2^2=16.$$
$$\text{Similarly, using } \sigma_2^2(1-\rho^2)=12,\ \text{we obtain } \sigma_1^2=4.$$
$$\text{Substituting } \sigma_1,\sigma_2 \text{ in } \sigma_1^2\sigma_2^2(1-\rho^2)=48,\ \text{leads to } \rho=1/2, (\text{ why not } -1/2?)$$

To infer the values of μ_1, and μ_2, subsitute the values of σ_1,σ_2,ρ in (2), and equate the coefficients of x and y terms to obtain the following set of equations:

$$-2\mu_1+\mu_2=-1,\quad \mu_1-\mu_2=-2.$$

Solving the above equations, we obtain $\mu_1=1$, and $\mu_2=3$.

Thus, (X, Y) is bivariate normal with means $1, 3$, variances $4, 16$, and correlation coefficient $\frac{1}{2}$. Further, $Cov(X, Y) = \rho\sigma_1\sigma_2 = 4$. Now, answer to part (a) is immediate.

For part (b), let $U = \frac{X-1}{2}, V = Y - 34$. Then (U, V) is standard bivariate normal with correlation $\rho = \frac{1}{2}$, and from the class notes, it can be seen that

$$V \mid U = u \text{ is normally distributed with mean } \rho u = \frac{u}{2}, \text{ and variance } (1 - \rho^2) = \frac{3}{4}.$$

Now, $V \mid X = x$ is the same as $V \mid u = (x - 1)/2$, implying $V \mid X = x$ is normal with mean $\frac{x-1}{4}$, and variance $\frac{3}{4}$.

Finally, $Y \mid X = x$ is the same as $4V + 3 \mid X = x$, implying $Y \mid X = x$ is normal with mean $x + 2$ and variance 12.