Machine Learning Worksheet 4

Probability Theory - II

1 Basic Probability

Problem 1. We say that two random variables are pairwise independent, if $P(X_2|X_1) = P(X_2)$ and hence

$$P(X_2, X_1) = P(X_1)P(X_2|X_1) = P(X_1)p(X_2)$$

We say that n random variables are mutually independent if

$$P(X_i|X_S) = P(X_i) \qquad \forall S \subseteq \{1, \ldots, n\} \setminus \{i\}$$

and hence

$$P(X_{1:n}) = \prod_{i=1}^{n} P(X_i)$$

Show that *pairwise* independence between all pairs of variables however does not necessarily imply mutual independence. It suffices to give a counter example.

Problem 2. Let X and Y be two random variables. Express $Var[X + Y]$ in terms of $Var[X]$, $Var[Y]$ and $Cov[X, Y]$.

Problem 3. Let X and Y be two random variables. Prove that $-1 \leq \rho(X, Y) \leq 1$. You may want to use the result from the previous problem.

Problem 4. Let X be a random variable. Show that, if $Y = aX + b$ for some parameters $a > 0$ and b, then $\rho(X, Y) = 1$. Similarly show that if $a < 0$, then $\rho(X, Y) = -1$.

Problem 5. Let $X \sim U(-1, 1)$ and $Y = X^2$. Obviously, Y is dependent on X (in fact, Y is uniquely determined by X). However, show that $\rho(X, Y) = 0$ (i.e. uncorrelated does not imply independent, in general; but see next problem for a special case).

Problem 6. Let $Z = (X, Y)$ be a bivariate normal distributed random variable. Furthermore, let $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$. Assume that $\rho(X, Y) = 0$. Show that in this case X and Y are independent.

Problem 7. Using Jensen's Inequality, show that for a finite random variable X (with n different values), its entropy is always bounded above by $\ln n$. Additionally, prove that the Kullback-Leibler divergence between any two discrete probability distributions is always non-negative.