Machine Learning Worksheet 3

Probability Theory Refresher I

1 Basic Probability

Problem 1: A secret government agency has developed a scanner which determines whether a person is a terrorist. The scanner is fairly reliable; 95% of all scanned terrorists are identified as terrorists, and 95% of all upstanding citizens are identified as such. An informant tells the agency that exactly one passenger of 100 aboard an aeroplane in which you are seated is a terrorist. The agency decide to scan each passenger and the shifty looking man sitting next to you is tested as ”TERRORIST”. What are the chances that this man *is* a terrorist? Show your work!

Problem 2: Two balls are placed in a box as follows: A fair coin is tossed and a white ball is placed in the box if a head occurs, otherwise a red ball is placed in the box. The coin is tossed again and a red ball is placed in the box if a tail occurs, otherwise a white ball is placed in the box. Balls are drawn from the box three times in succession (always with replacing the drawn ball back in the box). It is found that on all three occasions a red ball is drawn. What is the probability that both balls in the box are red? Show your work!

Problem 3: Calculate the mean and the variance of the uniform random variable X with PDF $p(x) = 1, \forall\ x \in [0,1]$, and 0 elsewhere.

Problem 4: Consider two random variables X and Y with joint density $p(x,y)$. Prove the following two results:

$$E[X] = E_Y[E_{X|Y}[X]] \tag{1}$$

$$Var[X] = E_Y[Var_{X|Y}[X]] + Var_Y[E_{X|Y}[X]] \tag{2}$$

Here $E_{X|Y}[X]$ denotes the expectation of X under the conditional density $p(x|y)$, with a similar notation for the conditional variance.

2 Probability Inequalities

Inequalities are useful for bounding quantities that might otherwise be hard to compute. We'll begin with a simple inequality, called the Markov inequality after Andrei A. Markov, a student of Pafnuty Chebyshev.

2.1 Markov Inequality

Let X be a non-negative, discrete random variable, and let $c > 0$ be a positive constant.

Problem 5: Show that

$$P(X > c) \leq \frac{E[X]}{c}.$$

Now, consider flipping a fair coin n times. Using the Markov Inequality, what is the probability of getting more than $(3/4)n$ heads?

2.2 Chebyshev Inequality

Apply the Markov Inequality to the deviation of a random variable from its mean, i.e. for a general random variable X we wish to bound the probability of the event $\{|X - E[X]| > a\}$, which is the same as the event $\{(X - E[X])^2 > a^2\}$.

Problem 6: Prove that

$$P(|X - E[X]| > a) \leq \frac{Var(X)}{a^2}$$

holds. Again, consider flipping a fair coin n times. Now use the Chebyshev Inequality to bound the probability of getting more than $(3/4)n$ heads.

2.3 Jensen's Inequality

Let f be a convex function. If $\lambda_1, \ldots, \lambda_n$ are positive numbers with $\lambda_1 + \cdots + \lambda_n = 1$, then for any $x_1, \ldots, x_n \in I$:

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)$$

Problem 7: Prove this inequality by using induction on n.