Machine Learning Worksheet 1

Probability Theory

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: A fair coin is tossed twice. Whenever it turns up heads, a red ball is placed into a box, otherwise a white ball. Afterwards, balls are drawn from the box three times in succession (replacing the drawn ball ever time). 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: A fair coin is flipped until heads shows up for the first time. What is the expected number of tails T and the expected number of heads H in any one run of this experiment? Show your work.

Hint: While there is a very short solution to this problem for people with a good intuition, the rest of us might need to look at the geometric series *and its properties. You may use them without proof.*

Problem 4: Calculate mean and variance of a uniform random variable X on the interval $[a, b], a < b$ with probability density function

$$p(x) = \begin{cases} \frac{1}{b-a} & 0 \leq x \leq 1, \\ 0 & \text{elsewhere.} \end{cases}$$

Problem 5: Let X and Y be random variables with joint density $p(x, y)$. Prove the *tower properties*,

$$\begin{aligned} E[X] &= E_Y[E_{X|Y}[X]], \\ \mathrm{Var}[X] &= E_Y[\mathrm{Var}_{X|Y}[X]] + \mathrm{Var}_Y[E_{X|Y}[X]]. \end{aligned}$$

$E_{X|Y}[X]$ and $\mathrm{Var}_{X|Y}[X]$ denote the expectation and variance of X under the conditional density $p(x \mid y)$.

2 Probability Inequalities

Inequalities are useful for bounding quantities that might otherwise be hard to compute. A famous example is the Markov inequality

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

for a *non-negative* random variable X and a constant $c > 0$. From it, it is relatively easy to prove the Chebyshev inequality

$$p(|X - E[X]| > c) \leq \frac{\mathrm{Var}(X)}{c^2}$$

for arbitrary X with finite variance.

With the help of Chebyshev's inequality, one can prove (a weak version of) the *law of large numbers*, which roughly states that the empirical mean of n i.i.d. random variables X_i converges to the true mean for $n \to \infty$. More formally, for any $\epsilon > 0$

$$p\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - E[X_i]\right| > \epsilon\right) \to 0 \qquad \text{as } n \to \infty. \tag{1}$$

Problem 6: Prove eq. (1). You may assume that the X_i have finite variance $\mathrm{Var}[X_i]$. You may further use Markov's and Chebyshev's inequalities without proof.

(We highly recommend to practise your "proof skills" on them, though. The proofs are technical, but very short.)