Machine Learning: Homework #2

Due on November 06, 2017 at 10:00am

Professor Dr. Stephan Guennemann

Marc Meissner

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 airplane in which you are seated is a terrorist. The agency decides 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!

Solution
Let S symbolize the correct result of the scan and T the fact that the subject is a terrorist. According to the description, the following holds:

$$P(S|T) = P\left(\bar{S}|\bar{T}\right) = 0.95 \ , \ P\left(S|\bar{T}\right) = 0.05 \ , \ P(T) = 0.01 \text{ and } P(\bar{T}) = 0.99$$

We can then acquire the solution by Bayes Law:

$$P(T|S) = \frac{P(S|T) \cdot P(T)}{P(S|T) \cdot P(T) + P\left(S|\bar{T}\right) \cdot P(\bar{T})} = \frac{19}{118} \approx 0.161$$

The probability is only 16.1%. But since the man is looking shifty, he probably is the terrorist, anyway.

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 (placing the drawn ball back into the box every 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!

Solution
Let RR, WW and RW be the three possible combinations of colored balls in the box. Let $D3R$ denote the event that three red balls have been drawn. Then:

$$P(RR) = P(WW) = \frac{1}{4} \ , \ P(RW) = \frac{1}{2}$$

$$P(D3R|RR) = 1 \ , \ P(D3R|WW) = 0 \ , \ P(D3R|RW) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

Bayes Law leads to the solution:

$$P(RR|D3R) = \frac{P(D3R|RR) \cdot P(RR)}{P(D3R|RR) \cdot P(RR) + P(D3R|RW) \cdot P(RW)} = \frac{4}{5} = 0.8$$

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.

Problem 3 continued on next page...

Solution

F	1	2	3	4	...	k
n_T	0	1	2	3	...	$k-1$
n_H	1	1	1	1	...	1
$p(F)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	...	$\frac{1}{2^k}$

Table 1: The probabilities of the different outcomes.

In the table , F denotes the number of flips , n_t and n_H are the number of tails and heads and $p(E)$ is the probability of F.
The derivation of $E[n_H] = 1$ is trivial. For $E[n_T]$, one has to build a sum:

$$E[n_T] = \lim_{K\to\infty} \sum_{k=1}^{K} p_k \cdot n_{T,k} = \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^k \cdot (k-1)$$

This equation can be solved by using the derivation property of the geometric series. However, since n_T changes incrementally with F, one can simplify this to the geometric series:

$E[n_T] = \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k - 1 = \frac{1}{1-\frac{1}{2}} - 1 = 1$

Problem 4

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

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

Solution
Just apply the formulas/definitions of mean and variance:

$E[X] = \int_{-\infty}^{\infty} x \cdot p(x)\,\mathrm{d}x = \frac{1}{b-a} \cdot \int_a^b x\,\mathrm{d}x = \frac{1}{2} \cdot \frac{b^2-a^2}{b-a} = \frac{a+b}{2}$

$E[X^2] = \frac{1}{b-a} \cdot \int_a^b x^2\,\mathrm{d}x = \frac{1}{3\cdot(b-a)} \cdot [b^3 - a^3] = \frac{b^2+ab+a^2}{3}$
$\mathrm{Var}[X] = E[X^2] - (E[X])^2 = \frac{b^2+ab+a^2}{3} - \frac{b^2+ab+a^2}{4} = \frac{(b+a)^2}{12}$

Problem 5

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

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

Problem 5 continued on next page...

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

Solution
First equation:

$E[X] = E_Y[E_{X|Y}[X]]$
$\int_{-\infty}^{\infty} x\ p(x)\ dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p(x|y)\ dx\ p(y)\ dy$

$p(y)$ is a constant, so we can pull it into the integral

$\int_{-\infty}^{\infty} x\ p(x)\ dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p(x|y)\ p(y)\ dx\ dy$
$\int_{-\infty}^{\infty} x\ \int_{-\infty}^{\infty} p(x,y)\ dy\ dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p(x,y)\ dx\ dy$

According to Fubini's theorem, the integration order can be flipped. Thus, both statements are identical, which yields the proof.

Second equation using $\mathrm{Var}[X] = E[X^2] - E[X]^2$:

$\mathrm{Var}[X] = E_Y[\mathrm{Var}_{X|Y}[X]] + \mathrm{Var}_Y[E_{X|Y}[X]] = E_Y[E_{X|Y}[X^2] - E_{X|Y}[X]^2] + E_Y[E_{X|Y}[X]^2] - E_Y[E_{X|Y}[X]]^2$

$\mathrm{Var}[X] = E_Y[E_{X|Y}[X^2]] - E_Y[E_{X|Y}[X]^2] + E_Y[E_{X|Y}[X]^2] - E_Y[E_{X|Y}[X]]^2$

$\mathrm{Var}[X] = E_Y[E_{X|Y}[X^2]] - E_Y[E_{X|Y}[X]]^2$

$\mathrm{Var}[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2\ p(x|y)dx\ p(y)\ dy - \left(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ p(x|y)\ dx\ p(y)\ dy\right)^2$

$\mathrm{Var}[X] = \int_{-\infty}^{\infty} x^2\ p(x,y)dx - \left(\int_{-\infty}^{\infty} x\ p(x,y)\ dx\right)^2$

$\mathrm{Var}[X] = E[X^2] - E[X]^2$

The last property is taken as a proven equation, which proofs the second equation.

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.

Solution
$E[X] = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i = \mu$
$\mathrm{Var}[X] = \mathrm{Var}[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathrm{Var}[\frac{X_i}{n}] = n \cdot \frac{\sigma^2}{n^2} = \frac{\sigma^2}{n}$
$P(|X - E[X]| > \epsilon) = P(|\frac{1}{n} \cdot \sum_{i=1}^{n} X_i - E[X_i]| > \epsilon) \leq \frac{\sigma^2}{\epsilon^2 \cdot n} \to 0$ for $n \to \infty$.