TDA231 Homework 0: Theory

NA

May 11, 2018

Problem 1

What we want to calculate is the probability that we are ill given the positive test result: $P(\text{ill}|\text{positive})$. By Bayes' Theorem, this is given by

$$P(\text{ill}|\text{positive}) = \frac{P(\text{positive}|\text{ill})P(\text{ill})}{P(\text{positive})}$$

One by one, the probabilities on the right-hand side can be found. $P(\text{positive}|\text{ill})$ is given in the text as 0.99. Using historical data for our prior, we set $P(\text{ill})$ to be 10^{-4}. Finally, the probability of any random person receiving a positive result is found by the law of total probability:

$$P(\text{positive}) = P(\text{positive}|\text{ill})P(\text{ill}) + P(\text{positive}|\text{healthy})P(\text{healthy})$$

This is assuming, of course, that every test result is either positive or negative, and that every person taking it is either ill or healthy (with regards to this particular condition). This also gives us $P(\text{healthy}) = 1 - P(\text{ill})$. Finally, $P(\text{positive}|\text{healthy}) = 10^{-2}$ as well, by the claim that the test is equally accurate on positive and negative results.

The final calculation becomes

$$P(\text{ill}|\text{positive}) = \frac{(1 - 10^{-2})10^{-4}}{(1 - 10^{-2})10^{-4} + (1 - 10^{-4})10^{-2}} = 9.8 \cdot 10^-3$$

The probability I should assign to being afflicted is about $\frac{1}{100}$.

Problem 2

The covariance of two random variables is given as

$$COV(x, y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

where the last equality follows from the linearity of expectation and the fact that $E[X]$ and $E[Y]$ are constants.

However, in this case $Y = X^2$, meaning the first term is $E[X^3]$. But the random variable $Z = X^3$ must have the expected value 0, since its probably density function is an odd function on an interval centred on 0. The same goes for $E[X]$ itself, and so both the terms of the covariance become 0.

Problem 3

A variable $\theta \sim \text{Beta}(a, b)$ has the following properties:

$$\begin{aligned} E[\theta] &= \frac{a}{a+b} \\ \text{var}(\theta) &= \frac{ab}{(a+b)^2(a+b+1)} \end{aligned}$$

We want to assign a and b so that $E[\theta] = m$ and $\text{var}(\theta) = v$. The rest is simply equation solving. The expectation:

$$\frac{a}{a+b} = m \Leftrightarrow a = \frac{m}{1-m}b$$

The variance:

$$\frac{ab}{(a+b)^2(a+b+1)} = \frac{mb^2}{1-m} \cdot \frac{1}{(\frac{b}{1-m})^2(\frac{b}{1-m} - 1)} = \frac{m(1-m)^2}{b+1-m}$$

which we set to be equal to v. Solving for b, we get

$$\frac{((1-m)m - v)(1-m)}{v} = b$$

which in turn gives

$$\frac{((1-m)m - v)m}{v} = a$$

When checking if this really does yield the wanted expectation and variance, it helps to note the useful equalities

$$\begin{aligned} a+b &= \frac{((1-m)m-v)}{v} \\ a+b+1 &= \frac{(1-m)m}{v} \end{aligned}$$

Now we get

$$\begin{aligned} E[\theta] &= \frac{a}{a+b} = \frac{\frac{((1-m)m-v)m}{v}}{\frac{((1-m)m-v)}{v}} = m \\ \text{var}(\theta) &= \frac{ab}{(a+b)^2(a+b+1)} = \frac{(\frac{((1-m)m-v)}{v})^2(1-m)m}{(\frac{((1-m)m-v)}{v})^2 \cdot \frac{(1-m)m}{v}} = v \end{aligned}$$

Finally, the numbers given in the lecture were $m = 0.7, \sigma = 0.2$, the latter giving $\sigma^2 = \text{var}(\theta) = 0.04$. Inserting these values into the above we get $a = 2.975, b = 1.275$, which is correct to three digits compared to the values given in the lecture.