Probabilistic Machine Learning - Final Quiz

1. Calculate using Python and a Google Colab notebook, the following probabilities when three 20-sided dice are rolled together:

   (a) $P(D_1 + D_2 + D_3 > 50)$

   (b) $P(D_1 + D_2 + D_3 = 30 \mid \text{All dice are even})$

   (c) $P(D_1 + D_2 > D_2 + D_3)$

   For (b) you can make use of the function below:

   ```
   def isEven(values):
       return (values & 1 == 0)
   ```

2. An autonomous drone that is capable of detecting smoke is surveying an area of housing looking for dangerous fires. Dangerous fires are rare and only occur 1% of the time. However, smoke is fairly common, occurring about 10% of the time due to barbecues. If 90% of dangerous fires make smoke, and the drone detects smoke, what is the probability that there is a dangerous fire.

3. A PhD student is measuring the time it takes to send a TCP packet across an experimental network. Due to network traffic, the time is different each time, so she makes ten measurements:

   1.4ms, 4.1ms, 3.7ms, 3.2ms, 0.4ms, 3.7ms, 1.8ms, 0.6ms, 6.4ms, 2.0ms

   Since these measurement represent elapsed time, they cannot be assumed to be generated by a Gaussian distribution.

   (a) Use PyMC3 to model to case that these measurements of elapsed time are drawn from a Gamma distribution with unknown mean and standard deviation. Using appropriate priors, calculate posterior estimates for the mean and standard deviation of the Gamma distribution. (Hint: See the code in `measurement_noise.ipynb` for inspiration).

   (b) Using the estimates for the mean and standard deviation found above, plot the resulting Gamma distribution. (Hint: See the code in `continuous_distributions.ipynb` for plotting distributions.)

   (c) Generate a PDF of the plot above. Describe your results in an Overleaf document. Your description should follow the example included in the note. Describe the priors and the likelihood used by your model. You can generate additional plots of the posterior distributions to include within the description.