Exercises 4 for MA 413 – Statistics for Data Science

This sheet will cover lecture material from the lecture 7/10/2019 and later material.

1. The mutual information is defined as

$$I(X,Y) = \int\int f_{X,Y}(x,y)\log\left(\frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)}\right)dxdy. \tag{1}$$

   Show that $I(X,Y) = 0 \Leftrightarrow X$ and Y are independent.

2. Rewrite the Poisson distribution in the natural parameterisation recognizing the parameterization in terms of the Poisson mean.

3. Rewrite the Gamma distribution in the natural parameterisation recognizing the parameterization in terms of the Gamma parameters.

4. Moment generating functions and PDFs have a 1-1 correspondence. Using the result for the sum of independent random variables determine the distribution of a sum of m independent geometric random variables with parameter p.

5. Determine the distribution of k independent Gamma random variables with β parameter β and α parameters $\alpha_1, \dots, \alpha_k$ using MGFs.

6. Show from first principles that the MGF of a Poisson with mean μ is $M(t) = \exp(\lambda(e^t - 1))$.

7. Using MGFs show that the sum of k independent Gaussian random variables with parameters $(\mu_1, \sigma_1^2), \dots, (\mu_k, \sigma_k^2)$ is Gaussian and identify its mean and variance.

8. For two continuous random variables X and Y prove the law of total variance, e.g.

$$\mathrm{Var}_Y(Y) = \mathrm{E}_X\left\{\mathrm{Var}_{Y|X}(Y|X)\right\} + \mathrm{Var}_X\left(\mathrm{E}_{Y|X}\{Y|X\}\right). \tag{2}$$

   Show how this can be extended to the law of total covariance:

$$\mathrm{Cov}_{X,Y}(X,Y) = \mathrm{E}_Z\left\{\mathrm{Cov}_{X,Y|Z}(X,Y|Z)\right\} + \mathrm{Cov}_Z\left(\mathrm{E}_{X|Z}\{X|Z\}, \mathrm{E}_{Y|Z}\{Y|Z\}\right). \tag{3}$$