STAT743 FOUNDATIONS OF STATISTICS (Part I)

Assignment 1 Due in class at 2:30pm on Thursday October 3, 2019

Instructions:

1. Ensure that your full name and student number are clearly marked on each page of your solutions.

2. Solutions need not be typed (although they can be) but must be readable.

3. Start each question on a new page and clearly indicate where each part of the question begins.

4. Submit questions in the same order as given below.

5. You are expected to show all details of your solution and any results taken from my notes or the textbook must be clearly and properly referenced.

6. No extensions to the due date and time will be given except in extreme circumstances and late assignments will not be accepted. Assignments are due at the **start** of class.

7. Students are reminded that submitted assignments must be their own work. Submission of someone else's solution under your name is academic misconduct and will be dealt with as such. Penalties for academic misconduct can include a 0 for the assignment, an F for the course with an annotation on your transcript and/or dismissal from your program of study.

Q. 1 **a)** If A and B are two events in a sigma-algebra, prove, using only the Axioms of probability as a starting point, that

$$P(A \bigcap B) \leqslant P(A) \leqslant P(A \bigcup B) \leqslant P(A) + P(B)$$

b) If A, B, C are three events in a sigma-algebra, prove that

$$P(A \bigcup B \bigcup C) = P(A)+P(B)+P(C)-P(A \bigcap B)-P(A \bigcap C)-P(B \bigcap C)+P(A \bigcap B \bigcap C)$$

c) Casella and Berger exercise 1.24

Q. 2 Suppose that X is a random variable whose moment generating function $M_X(t)$ exists for $|t| < t_0$. The *Cumulant Generating Function* of X is defined as

$$K_X(t) = \log\left(M_X(t)\right) \qquad \text{for } |t| < t_0.$$

It can be shown that $K_X(t)$ has a convergent Taylor expansion

$$K_X(t) = \sum_{r=1}^{\infty} \frac{t^r}{r!} \kappa_r(X)$$

The quantities $\kappa_1(X), \kappa_2(X), \ldots$ are called the cumulants of X and are often very useful. In this question we examine the relationship between cumulants and moments.

a) Show that

$$\kappa_r(X) = \frac{d^r}{dt^r} K_X(t)\bigg|_{t=0} \qquad r = 1, 2, \ldots$$

b) Give expressions for the first three cumulants of X in terms of its mean and central moments.

c) If $X \sim \mathrm{N}(\mu, \sigma^2)$ give the values of all of the cumulants of X.

d) Let X and Y be independent random variables with finite cumulant generating functions in some neighbourhood of 0. Show that

$$\kappa_r(X + Y) = \kappa_r(X) + \kappa_r(Y).$$

Q. 3 Suppose that Y is a negative binomial random variable with probability mass function

$$f_Y(y \mid r, p) = \binom{y+r-1}{r-1} p^r (1-p)^y \qquad y = 0, 1, \ldots$$

a) Show that the moment generating function of Y is

$$M_Y(t) = \left(\frac{p}{1-e^t(1-p)}\right)^r \qquad \text{for } t < -\log(1-p)$$

b) Use the above result to verify the mean and variance of the negative binomial random variable given in my notes.

c) Prove that as $p \to 0$, the distribution of the random variable $Z = pY$ converges to a gamma distribution and give the parameters of that distribution.

Q. 4 **a)** The *log-normal* distribution is used a lot to model heavily skewed data. We say that Y has log-normal distribution with parameters μ and σ^2 if, and only if, $X = \log(Y)$ has a normal(μ, σ^2) distribution. Find the mean and variance of the log-normal distribution.

b) Suppose that a random variable X is in exponential family form with natural parameterization

$$f(x \mid \boldsymbol{\eta}) = h(x) c^*(\boldsymbol{\eta}) \exp\left\{\sum_{i=1}^{k} \eta_i t_i(x)\right\}$$

Prove that

$$\begin{aligned}
\mathrm{E}\left(t_j(X)\right) &= -\frac{\partial}{\partial \eta_j} \log c^*(\boldsymbol{\eta}) \\
\mathrm{Var}(t_j(X)) &= \frac{\partial}{\partial \eta_j} \mathrm{E}\left(t_j(X)\right) = -\frac{\partial^2}{\partial \eta_j^2} \log c^*(\boldsymbol{\eta})
\end{aligned}$$

c) Show that the normal and Poisson distributions are members of the exponential family. Make sure you give the natural parameters η_i and the corresponding functions of the variable $t_i(x)$.