**THE UNIVERSITY OF NEW SOUTH WALES**

**DEPARTMENT OF STATISTICS**

**Exercises for MATH3811/MATH3911, Statistical Inference/Higher Statistical Inference**

**Sufficiency**

1. Use the factorization criterion to find a sufficient statistic for the parameter when $X_1, X_2, \ldots, X_n$ are independent random variables each with distribution

   a) $N(\mu, 1)$,
   b) $N(0, \sigma^2)$,
   c) Uniform $(\theta, \theta + 1)$,
   d) Poisson $(\lambda)$.
   e) Distribution with a density $f(x; \theta) = \theta(1 + x)^{-(1+\theta)}, x > 0; \theta > 0$.
   f) Distribution with a density $f(x; \theta) = e^{-(x-\theta)} I_{[\theta,\infty)}(x), -\infty < \theta < \infty$.

   Check your answer in (d) using the definition of sufficiency.

   Answers: a) $\sum_{i=1}^n X_i$; b) $\sum_{i=1}^n X_i^2$; c) $(X_{(1)}, X_{(n)})$; d) $\sum_{i=1}^n X_i$; e)$\sum_{i=1}^n \log(1 + X_i)$; f) $X_{(1)}$.

2. Let $X_1, X_2, X_3$ be a sample of size 3 from the Bernoulli $(p)$ distribution. Consider the 2 statistics $S = X_1 + X_2 + X_3$ and $T = X_1X_2 + X_3$. Show that $S$ is sufficient for $p$ and $T$ is not.

3. A random variable $X = (X_1, X_2)$ has the following distribution (with $1 < \theta < 3$):

| $(x_1, x_2)$ | (0,0) | (0,1) | (1,0) | (1,1) |
|---|---|---|---|---|
| $P(X_1 = x_1,\ X_2 = x_2)$ | $\frac{1}{12}(12 - 7\theta + \theta^2)$ | $\frac{\theta}{12}(4 - \theta)$ | $\frac{\theta}{12}(4 - \theta)$ | $\frac{\theta}{12}(\theta - 1)$ |

   Check whether $X_1 + X_2$ or $X_1X_2$ is sufficient for $\theta$.

4. If $X_1, X_2, \ldots, X_n$ are independent Bernoulli $(p)$ random variables, prove that $X_1$ is not sufficient for $p$.

5. Given that $\theta$ is an integer and that $X_1$ and $X_2$ are independent random variables which are Uniformly distributed on the integers $1, 2, \ldots, \theta$, prove that $X_1 + X_2$ is not sufficient for $\theta$.

6. Suppose $X_1, X_2, \ldots, X_n$ are independent discrete random variables each with probability function $f(x; \theta), \theta$ unknown. Prove that $(X_1, X_2, \ldots, X_{n-1})$ is not sufficient for $\theta$.

7. Find a minimal sufficient statistic for the parameter when $X_1, X_2, \ldots, X_n$ are independent random variables each with distribution

   a) Poisson $(\lambda)$,
   b) $N(0, \sigma^2)$,
   c) Gamma $(\alpha)$, (With a density $f(x, \alpha) = \frac{1}{\Gamma(\alpha)} \exp(-x)x^{\alpha-1}, x > 0$. (Here the Gamma function is defined as $\Gamma(\alpha) = \int_0^\infty e^{-x}x^{\alpha-1}dx$ and has the property $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$. In particular, for a natural number $n : \Gamma(n + 1) = n!$ holds).
   d) Uniform $(0, \theta)$.

e) Uniform $(\theta, \theta+1)$.

f) Uniform $(\theta_1, \theta_2)$.

8. If $X_1, X_2, \ldots, X_n$ are i.i.d. random variables with densities $f(x;\theta)$ given below, find a sufficient statistics for $\theta$.

   a) $f(x;\theta) = \theta x^{\theta-1} I_{(0,1)}(x), \theta \in (0,\infty)$.

   b) $f(x;\theta) = \frac{1}{6\theta^4} x^3 e^{-x/\theta} I_{(0,\infty)}(x), \theta \in (0,\infty)$.

   Answers: a) $\prod_{i=1}^n X_i$; b) $\sum_{i=1}^n X_i$.

9. (*Ex 2.12 from the textbook) Consider the linear model $E[\mathbf{Y}] = \mathbf{X}\beta$ where $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ is a vector of $n$ random variables with $V(Y_i) = \sigma^2, i = 1, 2, \ldots n, Cov[Y_i, Y_j] = 0, i \neq j$. Here $\mathbf{X}$ is a $n \times p$ full-rank matrix of fixed constants and $\beta$ is a vector of unkown parameters. Assume the elements of $\mathbf{Y}$ are normally distributed. Show, by using the definition, that

   a) When $\sigma^2$ is assumed to be known, $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ is minimal sufficient for $\beta$.

   b) When $\sigma^2$ is also assumed to be unknown then the pair of statistics $[\hat{\beta}, (\mathbf{Y}\text{-}\mathbf{X}\hat{\beta})'(\mathbf{Y}\text{-}\mathbf{X}\hat{\beta})]$ is minimal sufficient for $(\beta, \sigma^2)$.

10. a) Find a vector statistic that is sufficient for the parameter vector $\theta = (\theta_1, \theta_2)'$ of the Beta distribution with a density

$$f(x;\theta_1,\theta_2) = \frac{1}{B(\theta_1,\theta_2)} x^{\theta_1-1}(1-x)^{\theta_2-1} I_{(0,1)}(x)$$

   (where $B(\theta_1,\theta_2) = \int_0^1 x^{\theta_1-1}(1-x)^{\theta_2-1}dx$ and $\theta_1 > 0, \theta_2 > 0$ are unknown parameters.

   b) Assume that $\theta_1 = \theta_2 = \psi$ holds. Find a minimal sufficient statistic for $\psi$.

**Homework revision questions on probability (general 2801/2901 stuff)**

11. How many children should a family be planning for if they want to have with probability 0.95 at least one boy and at least one girl? (Assuming that the probability of having a boy is the same as the probability of having a girl).

12. If $Z \sim N(0,1)$ find the density of $Z^2$.

13. The two parameter Gamma distribution Gamma $(\alpha, \beta)$ has a density

$$f_X(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} \exp(-x/\beta) x^{\alpha-1}, x > 0$$

   where $\alpha > 0$ and $\beta > 0$ are shape parameters. Show that if $X \sim$ Gamma $(\alpha, \beta)$ then $EX = \alpha\beta, EX^2 = \beta^2\alpha(\alpha+1), Var(X) = \alpha\beta^2$.