STAT 613 Final (2 hours) **May 9th, 2016**

(1) $\{X_i; i = 1, \ldots, n\}$ are iid Gaussian random variables and $\{Y_j; j = 1, \ldots, m\}$ are iid exponential random variables with densities

$$f_X(x; \mu, \sigma^2) = (2\sigma^2\pi)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right), \quad x \in \mathbb{R}$$
$$f_Y(y; \theta) = \theta^{-1} \exp(-y/\theta) \quad y > 0$$

respectively ($\mu \in \mathbb{R}, \sigma^2, \theta > 0$). $\{X_i; i = 1, \ldots, n\}$ and $\{Y_j; j = 1, \ldots, m\}$ are independent of each other.

(i) State the joint log likelihood of $\{X_i; i = 1, \ldots, n\}$ and $\{Y_j; j = 1, \ldots, m\}$.

(ii) Obtain the sufficient statistics of μ, σ^2 and θ.

(iii) Suppose $\mu = \gamma$, $\sigma^2 = \gamma^2$ and $\theta = \gamma$ obtain the maximum likelihood estimator of γ.

Hint: Use good notation and think carefully about the solution to use.

(2) $\{Y_i; i = 1, \ldots, n\}$ are iid random variables with mean $\mu(\theta)$ and variance $V(\theta)$. We test the hypothesis $H_0 : \theta = \theta_0$ against $H_A : \theta \neq \theta_0$ using the test statistic

$$T = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{Y_i - \mu(\theta_0)}{\sqrt{V(\theta_0)}}.$$

(i) What is the asymptotic distribution of T under the null hypothesis?

Hint: Use the classical CLT.

(ii) Obtain the power function of T under the alternative hypothesis.

Hint: Use the alternative $\theta_1 = \theta_0 + \phi/\sqrt{n}$.

(3) $\{X_i\}$ are iid random variables where

$$X_i = \begin{cases} U_i & \text{if } U_i \leq V_i \\ V_i & \text{if } U_i > V_i. \end{cases}$$

(i) Let F_U denote the distribution of U and f_V denote the density of V. Show that the distribution of $Y = U - V$ is

$$F_{U-V}(y) = \int_{-\infty}^{\infty} F_U(y+v) f_V(v) dv.$$

(i) Suppose that U and V are exponential distributed with density $\theta_1 \exp(-\theta_1 u)$ and $\theta_2 \exp(-\theta_2 v)$ $(u, v, \theta_1, \theta_2 \geq 0)$ obtain the density of X.

You may use that the distribution function of the density $\theta^{-1} \exp(-x/\theta)$ is $\exp(-x/\theta)$.

(iii) Show how the EM-algorithm can be utilized to obtain the maximum likelihood estimators of θ_1 and θ_2.

(4) Y_i are independent random variables with mean $\mathbb{E}[Y_i] = g(\beta x_i)$ (where x_i is an observed regressor, g is a known function and β is an unknown parameter). Suppose $\text{var}[Y_i] = V(\beta x_i)$. We define the estimator

$$\widehat{\beta} = \arg\min_{\beta \in \Theta} \sum_{i=1}^{n} [Y_i - g(\beta x_i)]^2 .$$

(i) Assuming that $\widehat{\beta}$ consistently estimates β, derive the asymptotic distribution of $\widehat{\beta}$.

Hint: There are some technical issues in solving this question; this can be ignored. You may use the result that if $\{X_i\}$ are independent mean zero random variables that are not necessarily identically distributed and $[\sum_{i=1}^{n} \text{var}[X_i]]^{-1} \stackrel{P}{\to} 0$ as $n \to \infty$ (you can assume this condition holds true), then

$$\frac{1}{\sqrt{\sum_{i=1}^{n} \text{var}[X_i]}} \sum_{i=1}^{n} X_i \stackrel{D}{\to} N(0,1).$$

(ii) Briefly outline how you would estimate the "asymptotic" variance of $\widehat{\beta}$.