18.655 Midterm Exam 1, Spring 2016
Mathematical Statistics
Due Date: 4/4/2016

1. **Censored Survival Times.** Let $Y_1, \ldots, Y_n$ be iid $Exponential(\theta)$,

   $f(y \mid \theta) = \frac{1}{\theta} e^{-y/\theta}, y > 0, \theta > 0.$

   Suppose that the n survival times Y_i are censored at time $c > 0$, we observe:

   $$X_i = \begin{cases} Y_i, & if \quad Y_i < c \\ c, & if \quad Y_i \geq c \end{cases}$$

   (a). Find $\hat{\theta}(Y_1, \ldots, Y_n)$ the maximum-likelihood estimate of θ based on the original sample of survival times $Y_1, \ldots, Y_n$.

   (b). Determine the distribution of $\hat{\theta}(Y_1, \ldots, Y_n)$ and give formulas for the mean and variance of this distribution.

   (c). Derive the likelihood function of X_1. (For $c < \infty$, the distribution of X_1 is a mixture of a continuous distribution and a discrete distribution).

   (d). Find $\hat{\theta}(X_1, \ldots, X_n)$ the maximum-likelihood estimate of θ based on the sample of censored survival times $X_1, \ldots, X_n$.

   (e). Comment on the degree of similarity/dissimilarity of the estimators in (a) and (d) for the cases:

   - $c >> E[X_1]$
   - $c \approx E[X_1]$
   - $c << E[X[_1]$

2. **Parameters with Order Restrictions.** Let $X_1, \ldots, X_n$ be indpendent random variables with

$$X_i \sim P_{\theta_i}, \text{ for } i = 1, \ldots, n$$

(a). For $P_\theta = N(\theta, 1)$, determine the maximum likelihood estimate of

$$(\theta_1, \ldots, \theta_n)$$

when there are no restrictions on the θ_i.

(b). In (a), for $n = 2$ determine the maximum likelihood estimate of

when (θ_1, θ_2) is restricted to satisfy $\theta_1 \leq \theta_2$.

(c). Repeat (a) and (b) when P_θ is the Laplace distribution with density

$$f(x \mid \theta) = \frac{1}{2} exp\{-|x - \theta|\}, \; -\infty < x < \infty.$$

(d). Compare the answers to (b) in the two cases. Comment on the relative contribution of x_1 and x_2 to the mle's for the two cases.

3. **Gaussian Mixtures.** Let $X_1, \ldots, X_n$ be i.i.d. from a population with density:

$$f(x \mid \theta) = 0.5\tau_1\phi(\tau_1(x - \mu_1)) + 0.5\tau_2\phi(\tau_2(x - \mu_2)).$$

where

- $\phi(\cdot)$ is the density of a standard Normal distribution, and
- $\theta = (\mu_1, \tau_1, \mu_2, \tau_2)$ and
- $\Theta = \{\theta : \mu_1 \in R, \tau_1 > 0, \mu_2 \in R, \tau_2 > 0\}$

The class of Gaussian mixture probability models is

$$\mathcal{P} = \{P_\theta, \text{ distributions with density } f(x \mid \theta), \theta \in \Theta\}.$$

is the class of $50-50$ mixtures of two normal populations with repsective means μ_1, μ_2, respectively, and variances ${\tau_1}^-1$, τ_2^{-1}, respectively,. The parameters τ_j scale the *precision* of the distribution

Consider the *compact* class of Gaussian mixture probability models:

$$\mathcal{P}^* = \{P_\theta, \text{ distributions with density } f(x \mid \theta), \theta \in \Theta^*\}.$$

where

$$\Theta^* = \{\theta \in \Theta : |\theta|^2 \leq C^*\} \text{ for some finite constant } C^* > 0.$$

For each of the following cases, address the problem of specifying the maximum likelihood estimate of θ for when $P \in \mathcal{P}^*$.

(a). $n = 1$.

(b). $n = 2$.

For each case, comment on:

- Existence of a maximum likelihood estimate (MLE).
- If the MLE exists, the uniqueness of the maximum likelihood estimate.

(c). For each case (a) and (b), how are your answers affected by taking the limit: $\Theta^* \to \Theta$, i.e., $C^* \to \infty$.

(d). Comment on the case $n > 2$ when $x_i \neq x_j$ for some $i \neq j$.

4. **Maximum Entropy Distributions.**

Let $\mathcal{P}$ be the class of all continuous distributions on $\mathcal{X} \subset R$.

Consider constraining $\mathcal{P}$ to those distributions satisfying constraints on the expectations of certain statistics of X.

To be precise, let

- $k =$ the number of constraints
- $T_1(X), T_2(X), \ldots, T_K(X)$ are (univariate) statistics of X.
- For $\eta_1, \eta_2, \ldots, \eta_k$, fixed constants, consider
$$\mathcal{P}^* = \{P \in \mathcal{P} : E[T_j(X) \mid P] = \eta_j, j = 1, \ldots, k\}$$
Each constant η_j defines a *parameter* of P.
If $X_1, X_2, \ldots$ are sampled independentlly from P, then
$$\lim_{n\to\infty} \frac{\sum_{i=1}^n T_j(X_i)}{n} = \eta_j.$$

The **entropy** $h(f)$ of a continuous random variable X with density f is defined by:

$$h(f) = E[-\log f(X)] = -\int_{-\infty}^{+\infty} [\log(f(x))] f(x) dx.$$

(a). Show that the canonical $k-$parameter exponential family density

$$f(x \mid \eta) = exp\{\eta_0 + \sum_1^k \eta_j T_j(x) - A(\eta)\},\ x \in \mathcal{X}$$

maximizes $h(f)$ subject to the constraints:

$$\int_{\mathcal{X}} f(x) dx = 1$$
$$\int_{\mathcal{X}} T_1(x) f(x) dx = \eta_1$$
$$\int_{\mathcal{X}} T_2(x) f(x) dx = \eta_2$$
$$\vdots$$
$$\int_{\mathcal{X}} T_k(x) f(x) dx = \eta_k.$$

The entropy measure is used in information theory to (negatively) scale the information content of a random variable. Such content can be associated with the *parameters* of the distribution and the entropy measure scales how much prior information is eliminated from the distribution.

(b). Find the maximum-entropy distribution when

$$k = 1,\ T_1(X) = X, \text{ and } \mathcal{X} = (0, \infty).$$

(c). Find the maximum-entropy distribution when

$$k = 2,\ T_1(X) = X, T_2(X) = X^2, \text{ and } \mathcal{X} = (-\infty, +\infty).$$

18.655 Mathematical Statistics
Spring 2016