STATISTICS QUALIFYING EXAM: JANUARY 22, 1982

<u>INSTRUCTIONS</u>: Answer <u>ANY FOUR</u> problems. Do each in a separate blue book.

. Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables with exponential density given by

$$f_\theta(x) = \theta e^{-\theta x}\ , \quad \theta>0\ ,\ x>0$$

$\frac{\Sigma X_i}{n}$

$\frac{n}{\Sigma X_i}$

(i) Find maximum likelihood estimators for θ and for the mean $\mu = 1/\theta$. Justify your answer. Derive the exact distribution of your estimator of μ. $\int_{-\infty}^{x} \theta^n$ $\frac{1}{n}\Gamma(n, \theta)$

(ii) Derive a uniform minimum variance unbiased estimator for $r(x) = e^{-\theta x}$ and justify your derivation based on appropriate theorems.

. Let $Y_i = \alpha + \beta x_i + \epsilon_i$, $i=1,\ldots,n$, with the ϵ_i being independent identically distributed random variables, and the x_i a sequence of known constants.

(i) Assuming only that $E(\epsilon_i) = 0$ and $Var(\epsilon_i) = \sigma^2 > 0$, derive "good" estimators of α and β.

(a) State and prove <u>any</u> optimality properties you can about the estimators you derive. MVLUE

(b) State and prove that your estimator of β is consistent and asymptotically normal. Be sure to state any sufficient conditions for proving your results.

(c) Give a counter-example under which your estimator is not consistent.

(ii) Assume that ϵ_i is normally distributed with mean 0 and variance σ^2.

(a) Find the exact distribution of your estimator for β.

(b) State and prove <u>any</u> additional optimality properties you can think of under this additional assumption of normality.

3. Let X_i, $i=1,\ldots,n$ be independent random variables where X_i is exponentially distributed with mean μ_i and variance $\mu_i^2\sigma^2$. If $Y_i = \log(X_i)$ and $\mathrm{Var}(Y_i) = \theta_i$, prove that $\theta_i=\theta$ for all i.

$\Rightarrow$ 8:70

$p(x) = \frac{1}{A_i} e^{-\frac{(x-C_i)}{A_i}}$ $x \ge C_i$ $\Rightarrow \mu_i = C_i + A_i$ $\sigma_i^2 = A_i^2 = \mu_i^2\sigma^2$ $\Rightarrow C_i/A_i$ = constant

$EY_i = \log A_i + a$ $EY_i^2 = (\log A_i)^2 + 2(\log A_i)\, a + b$ $\sigma^2(Y_i) = b - a^2$ = constant

4. Let $0<\theta<1$ be an unknown parameter. Suppose we want to estimate θ based on one observation X, where X is binomial with n trials and success probability p and $\theta = .3p^2 + .6p + .2$. Prove that the maximum likelihood estimate of θ, denoted by $\hat{\theta}$, is an inadmissable estimator for a squared error loss function.

5. An experimenter takes repeated observations of an unknown constant μ. According to his model, the measurements are given by

$$X_i = \mu + \epsilon_i \quad ; \quad 1 \le i \le n,$$

where $\epsilon_1,\ldots,\epsilon_n$ are assumed to be independent, identically distributed random variables having a common t distribution with $k=4$ degrees of freedom. Suppose that the number n of measurements that he has performed is large.

(i) What can he say about the distribution of the sample mean, $\overline{X}_n$, of the X's?

$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{2}} \longrightarrow N(0,1)$$

(ii) What can he say about the distribution of the sample median $Y_{[n/2]}$ of the X's?

$$\frac{\sqrt{n}\,(Y_{[\frac{n}{2}]} - \mu)}{\frac{f(\mu)}{2}} \longrightarrow N(0,\ 1)$$

$$= \frac{3}{16}\,\frac{1}{(1+\frac{\mu^2}{4})^{5/2}} \le \frac{3}{16}$$

(iii) Which one of $\overline{X}_n$, $Y_{[n/2]}$ is more reliable as an estimate of μ ?

<u>Reminder:</u> The t-density with k degrees of freedom is given by

$$f(x) = \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{1}{\sqrt{k\pi}} \frac{1}{(1+\frac{x^2}{k})^{\frac{k+1}{2}}}$$

and has a variance equal to $\frac{k}{k-2}$ for $k \geqq 3$. Recall that

$$(i)\ E\varepsilon^2 = \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{2}{\sqrt{k\pi}} \int_0^\infty \frac{x^2}{(1+\frac{x^2}{k})^{\frac{k+1}{2}}} dx$$

$$= \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{1}{\sqrt{k\pi}} \int_0^\infty \frac{x^{1/2}}{(1+\frac{x}{k})^{\frac{k+1}{2}}} dx$$

$$\Gamma(n+\tfrac{1}{2}) = \frac{1.3.5\ldots(2n-1)}{2^n} \sqrt{\pi} .$$

$$z = \frac{1}{1+\frac{x}{k}} \qquad = z^{\frac{k-4}{2}} (1-z)^{1/2}$$

$$= \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})} \frac{k}{\sqrt{k\pi}} \int_0^1 z^{\frac{k-3}{2}} \sqrt{k} \left(\frac{1-z}{z}\right)^{1/2} = 2$$

5. Leaves of a plant are examined for insects, and it is found that X_i leaves have precisely i insects $(i=1,2,\ldots; \Sigma X_i = N)$. The number of insects per leaf is believed to be a Poisson random variable, except that many leaves have no insects because they are unsuitable for feeding and not merely because of the chance variation allowed for by Poisson distribution. The empty leaves are therefore not counted.

Show that: $\sum_{i=2}^{\infty} \frac{iX_i}{N}$

is an unbiased estimator of the Poisson parameter μ, and determine its variance. $Z_{kj} = \begin{cases} k & \text{if } j\text{th leaf has } k \text{ inset} \\ 0 & \text{otherwise} \end{cases}$ then $kX_k = \sum_{j=1}^{N} Z_{kj}$

$$Y_j = \sum_{k=2}^{\infty} Z_{kj} \qquad P_k = P(Z_{kj} = k) = \frac{\frac{\lambda^k}{k!} e^{-\lambda}}{1-e^{-\lambda}}$$

$$\Rightarrow EY_j = \sum_{k=2}^{\infty} kP_k = \lambda \qquad EY_j^2 = E\sum_{k=2}^{\infty} Z_{kj}^2 = \sum_{k=2}^{\infty} k^2 P_k = \frac{1}{1-e^{-\lambda}}(\lambda^2+\lambda-\lambda e^{-\lambda})$$

$$= \frac{\lambda^2}{1-e^{-\lambda}} + \lambda$$

$$\sigma^2(Y_j) = \frac{\lambda^2 e^{-\lambda}}{1-e^{-\lambda}} + \lambda$$

$$E\sum_{k=2}^{\infty} \frac{kX_k}{N} = E\frac{1}{N}\sum_{j=1}^{N} Y_j = EY_1 = \lambda$$

$$\sigma^2\left(\sum_{k=2}^{\infty} \frac{kX_k}{N}\right) = \sigma^2\left(\frac{1}{N}\sum_{j=1}^{N} Y_j\right) = \frac{1}{N}\sigma^2(Y_1) = \frac{1}{N}\left(\lambda + \frac{\lambda^2 e^{-\lambda}}{1-e^{-\lambda}}\right)$$