Indian Statistical Institute, Bangalore
B. Math (II)
Second Semester 2012-13
Mid-Semester Examination : Statistics (II)

Date: 01-03-2013 **Maximum Score 60** **Duration: 3 Hours**

1. Consider a *probability density function (pdf)* f, that is, $f \geq 0$ on $(-\infty, \infty)$ and $\int_{-\infty}^{\infty} f(t)\,dt = 1$. Let $X_1, X_2, \cdots, X_n$, $n \geq 2$, be a random sample from *location-scale family* whose distribution is specified by $f_{\mathrm{X}}(x; \mu, \sigma) = \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)$, $\mu \in (-\infty, \infty)$ and $\sigma > 0$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$. Prove that $\left(\frac{X_1-\overline{X}}{S}, \frac{X_2-\overline{X}}{S}, \cdots, \frac{X_n-\overline{X}}{S}\right)$ is an *ancillary statistic,* that is, its distribution does not depend on the parameter $\theta = (\mu, \sigma)$.

[8]

2. Let $X_1, X_2, \cdots, X_n$ be a random sample from the *distribution with pmf/pdf* $f(x|\theta)$ indexed by $\theta \in \Theta$. Let $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ and $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$. Let $T(\boldsymbol{X})$ be sufficient for θ. Let $e_{\mathrm{MLE}}(\boldsymbol{X})$, the *maximum likelihood estimator* for θ, exist. Then show that $e_{\mathrm{MLE}}(\boldsymbol{X})$ is a function of $T(\boldsymbol{X})$.

[8]

3. An entrepreneur who deals in washing machines is interested in finding a suitable guarantee period for her product. She is willing to set the guarantee period to 3 years provided $\tau(\theta)$, the probability that a washing machine survives for at least 3 years, is not too small. Suppose the lifetime X of her product is known to have *Weibull distribution* with parameter $\theta > 0$ with *pdf* given by $f_{\mathrm{X}}(x, \theta) = \frac{1}{\theta} c x^{c-1} e^{-\frac{1}{\theta}x^c} I_{(0,\,\infty)}(x)$; where $c > 0$ is a constant; $\theta > 0$. Suppose n such machines were put to survival test and that $X_1, X_2, \cdots, X_n$ were their lifetimes. Can you help the entrepreneur to estimate $\tau(\theta)$ based on $X_1, X_2, \cdots, X_n$? Is your estimator unbiased? Obtain *uniformly minimum variance unbiased estimator* (*UMVUE*) for $\tau(\theta)$.

[14]

4. For collecting admission forms for a *kindergarten* school, aspirants queue up well in advance. Let 0 signify the time point at which the school commences issuing forms. Then it may be reasonable to assume that the aspirants arrive independently, randomly and uniformly over a period $(-\theta, 0)$, $\theta > 0$ being unknown. The school and an *NGO* are interested, among other things, in having some idea about θ so as to make arrangements to provide basic amenities to the aspirants. Let $X_1, X_2, \cdots, X_n$ denote the arrival times of n randomly chosen aspirants who arrive at the school. Obtain a sufficient statistic T for θ based on $X_1, X_2, \cdots, X_n$. Is your T minimal as well? If yes, substantiate. If not, obtain minimal sufficient statistic based on $X_1, X_2, \cdots, X_n$. Substantiate. Is your minimal sufficient statistic complete? Substantiate. Hence or otherwise obtain *UMVUE* for θ.

$[(6 + 2) + (8 + 2) = 18]$

[PTO]

5. Let $X_1,\ X_2,\cdots,X_n$ be a random sample from the *distribution with pmf/pdf* $f(x|\theta)$ indexed by $\theta\in\Theta$. Let $\mathbf{X}=(X_1,\ X_2,\cdots,X_n)$ and $\boldsymbol{x}=(x_1,\ x_2,\cdots,x_n)$. Let $T(\mathbf{X})$ be sufficient for θ. Let further the *prior distribution* specified for θ, be $\pi(\theta)$. Show that the *posterior distribution* $\pi(\theta|\boldsymbol{x})$ depends on $\boldsymbol{x}$ only through $T(\boldsymbol{x})$, $\forall\boldsymbol{x}\in\mathcal{X}$.

[8]

6. Let $\pi(\theta)=\frac{1}{\lambda^a\Gamma a}e^{-\frac{\theta}{\lambda}}\theta^{a-1}I_{(0,\infty)}(\theta)$ be the *prior distribution* specified for θ, the parameter that indexes the *Poisson model.* Let $X_1,\ X_2,\cdots,X_n$ be a random sample from the *Poisson distribution with pmf* $f(x|\theta)=\frac{e^{-\theta}\theta^x}{x!}$, $\theta>0$. Find *Bayes estimator.* Recall that *the posterior variance* is a good indicator of the performance of the *Bayes estimator.* Determine the sample size n so that the *posterior variance* is no larger than b. Here $\lambda>0$, $a>0$ and $b>0$ are all known.

$[8+4=12]$