Indian Statistical Institute, Bangalore
B. Math (II)
First semester 2012-2013
End-Semester Examination : Statistics (I)

Date: 27-11-2012 **Maximum Score 75** **Duration: 3 Hours**

1. Let X_1, X_2 be independent $Gamma(\lambda, \alpha_1)$ and $Gamma(\lambda, \alpha_2)$ random variables, $\lambda, \alpha_1, \alpha_2 \in (0, \infty)$. Obtain the distribution of $Y_1 = \delta \frac{X_1}{X_2}$, $\delta > 0$. Can you find values of $\lambda, \alpha_1, \alpha_2$ and δ so that Y_1 would have F distribution with parameters m and n where m and n are positive integers.

   [8 + 4 = 12]

2. A social activist works towards the welfare of blind people. She collects donations on every Monday from different *randomly* chosen people in her town. She stops as soon as she gets donations from 10 persons or she finishes approaching 50 persons for donations, whichever is the earlier. Let $\theta \in (0, 1)$ be the unknown proportion of people that are willing to make a donation. **Build** a probability model for Y, the number of persons she has to request for donations till she stops. Show that $E(Y)$, the *expected* number of persons she requests for donations till she stops, cannot exceed min $\left\{\frac{10}{\theta}, 50\right\}$.

   [8 + 6 = 14]

3. Let $X_1,\ X_2, ..., X_n$ be a random sample from the population with unknown distribution function F. Let $e(\mathbf{X}) = e(X_1,\ X_2, ..., X_n)$ be a *statistic* that is to be used in an inference procedure. To know the *efficacy* of the inference procedure we often need to have an idea about the distribution of $e(\mathbf{X})$, in general and *variance of* $e(\mathbf{X})$, in particular. Explain how you would use the *bootstrap techniques* to achieve that.

   [10]

4. Suppose we draw a random sample of 128 families from a large population of families with exactly three children each. The number of boys per family was recorded. The data are displayed in the following table. The first column of the table gives the number of boys per family, ranging from 0 to 3. The second column gives the number of families in the sample having that number of boys. Thus, for example, 40 of the 128 families, have 2 boys each.

No. of boys per family	Observed frequency
0	26
1	32
2	40
3	30

Formulate and carry out a *chi-square goodness of fit test, at level of significance* $\alpha = 0.05$, to test whether the data come from a *binomial distribution*. Also report the p value.

[16]

[PTO]

5. The Gumbel distribution is used to model extreme values. Let us consider Gumbel *distribution* $G(\mu, \beta)$ with parameters $\mu \in (-\infty, \infty)$ *and* $\beta > 0$ having *probability density function (pdf)* given by

$$f(x|\mu, \beta) = \frac{1}{\beta}\left[\exp\left(-\frac{(x-\mu)}{\beta}\right)\right]\exp\left[-\exp\left(-\frac{(x-\mu)}{\beta}\right)\right];\ x \in (-\infty, \infty). \qquad (1)$$

(a) Check that $f(x|\mu, \beta)$ in (1) is indeed a *density function.*

(b) It is known that $-\int_0^\infty (\log z)\exp(-z)\,dz = \gamma$, where γ is the Euler–Mascheroni constant given by $\gamma = \lim_{n\to\infty}\left(\sum_{k=1}^{n}\frac{1}{k} - \log n\right)$. Hence or otherwise find the mean of the Gumbel *distribution* (1).

(c) Find *pth* quantile, $0 < p < 1$, of the Gumbel *distribution* (1).

(d) Find first quartile, median and third quartile of the Gumbel *distribution* (1).

(e) Find mode of the Gumbel *distribution* (1).

(f) If the *variance* of the Gumbel *distribution* (1) is given by $\frac{\pi^2}{6}\beta^2$ then find *method of moments (mom)* estimators for the parameter μ and β of the Gumbel *distribution* (1) based on a random sample of size n from (1).

(g) Find *maximum likelihood estimator (mle)* for the parameter μ of the Gumbel *distribution* (1) based on a random sample of size n from (1) with known β.

(h) If U is uniform on $(0,1)$ then find the distribution of $W = -\log\log\frac{1}{U}$.

(i) Suppose we can draw observations from uniform distribution on $(0,1)$. *Explain*, using the answer to (5.h), how you would draw observations from the Gumbel *distribution* $G(-10.3, 2.5)$.

$[4+4+4+3+4+6+6+5+4 = 40]$