試科目	數理統計	所別	統計所 441 446	考試時間	5月22日上午第一 星期六 下

1. (20 points) Let $G_j \sim \Gamma(\lambda_j, \alpha)$, $j = 1, ..., r+1$, be independent Gamma random variables, where λ and α are shape and scale parameters, respectively. Define $X_j = G_j / \sum_{i=1}^{r+1} G_i$, for $j = 1, ..., r$.
   (a) Derive the distribution of X_j. (5 points)
   (b) Derive the joint distribution of $(X_1, ..., X_r)$. (15 points)

2. (20 points) Let $X_1, ..., X_n$ be a random sample from $Bernoulli(p)$, where p has a prior distribution $Beta(\alpha, \beta)$.
   (a) Calculate $P(X_{n+1} = 1 | \sum_{i=1}^{n} X_i = r)$. (10 points)
   (b) Calculate $P(X_{n+1} = X_{n+2} = 1 | \sum_{i=1}^{n} X_i = r)$. (10 points)

3. (40 points) Consider the linear regression model

$$Z = \theta_1 + \theta_2 Y + e,$$

   where $e \sim N(0, \sigma^2)$ (σ is known). Suppose that we have observed data $\{(y_i, z_i), i = 1, ..., n.\}$
   (a) Find a sufficient statistic for $\boldsymbol{\theta} = (\theta_1, \theta_2)$. (10 points)
   (b) Suppose that $\boldsymbol{\theta}$ has a non-informative prior $\pi(\boldsymbol{\theta}) \propto 1$. Derive the posterior distribution of $\boldsymbol{\theta}$ given the data. (10 points)
   (c) (continued from part (b)) Suppose that one wants to predict a future Z at a given y. Derive $p(z|\text{data})$, the predictive distribution based on the data. (20 points)

4. (20 points) Consider a sample from a population in which there are three different genotypes. Suppose that the frequencies of the three genotypes are

$$p_1 = \theta^2, p_2 = 2\theta(1-\theta), p_3 = (1-\theta)^2, 0 < \theta < 1.$$

   Let N_i denote the number of individuals of type i in the sample of size n, then (N_1, N_2, N_3) has a multinomial distribution with parameters (n, p_1, p_2, p_3).
   (a) Give two different method of moments estimators for θ. (10 points)
   (b) Suppose we observe a sample of three individuals and obtain $x_1 = 1$, $x_2 = 2$, $x_3 = 1$ (x_i denote the genotype of the ith individual). Write down the likelihood function and derive the MLE for θ. (10 points)