Mid-term Exam: Statistical Inference

1 (10') Suppose that $X_1, \cdots, X_m$ i.i.d. $\sim N(\mu_1, \sigma^2)$, $Y_1, \cdots, Y_n$ i.i.d. $\sim N(\mu_2, \sigma^2)$, and X_i's and Y_j's are independent. Let $\bar{X}, \bar{Y}, S_X^2, S_Y^2$ denote their sample means and sample variances. Determine the distribution of

$$T = \frac{\alpha(\bar{X} - \mu_1) + \beta(\bar{Y} - \mu_2)}{\sqrt{\frac{(m-1)S_X^2 + (n-1)S_Y^2}{m+n-2}\left(\frac{\alpha^2}{m} + \frac{\beta^2}{n}\right)}},$$

where α, β are fixed constants.

2 (20') Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be a random sample from Normal distribution $N(\theta, 1)$.

(i) (5') Derive the moment estimator of θ^2.

(ii) (5') Derive the MLE of θ^2.

(iii) (5') Derive the UMVUE of θ^2.

(iv) (5') Is the UMVUE an efficient estimator of θ^2? Why?

3 (20') Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be a random sample from the distribution with p.d.f.

$$f(x; \alpha, \beta) = \frac{1}{\beta} e^{-(x-\alpha)/\beta}, \ x \geq \alpha, \ \alpha \in R, \ \beta > 0.$$

Find sufficient statistic and MLE of

(i) (5') α when β is known.

(ii) (5') β when α is known.

(iii) (10') α and β when both are unknown.

4 (15') In order to study the height of male students in a university, we took a random sample of size 5 and observed their heights (cm): $174, 171, 168, 175, 170$. For simplicity, suppose that the variance of height is $\sigma^2 = 9$ (cm^2).

(i) (5') Estimate the mean height (μ) of male students in this university.

(ii) (5') Construct a 99% confidence interval for μ. ($z_{0.025} = 1.96, z_{0.005} = 2.58, t_{4,0.005} = 4.60$)

(iii) (5') Determine the sample size n such that the length of confidence interval can be reduced by 80%.

5 (20') Let r.v. X be the number of goals scored by teams during the first round matches of the 2002 World Cup and suppose that X follows a Poisson distribution $P(\lambda)$ with p.d.f.

$$f(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}, \lambda > 0, x = 0, 1, 2, \cdots$$

Let $X_1, \cdots, X_n$ be a random sample.

(i) (5') Determine the distribution of $\sum_{i=1}^n X_i$.

(ii) (10') The observed values of $X_1, \cdots, X_n$ with $n = 95$ are summarized in the following table. Derive the MLE of λ and compute its efficiency.

Goals	0	1	2	3	4	5	6	7	8
Frequency	23	37	20	11	2	1	0	0	1

(iii) (5') Construct an approximately 99% confidence interval for λ. ($z_{0.025} = 1.96, z_{0.005} = 2.58$)

6 (15') The independent random samples $X_i, i = 1, \cdots, 5$ and $Y_i, i = 1, \cdots, 5$ represent resistance measurements taken on two test pieces, and the observed values (in ohms) are as follows:

$$x_1 = 0.3,\ x_2 = 0.2,\ x_3 = 0.1,\ x_4 = 0.2,\ x_5 = 0.1,$$

$$y_1 = 0.2,\ y_2 = 0.1,\ y_3 = 0.3,\ y_4 = 0.2,\ y_5 = 0.2.$$

Assume that $X_i \sim N(\mu_1, \sigma_1^2)$ and $Y_i \sim N(\mu_2, \sigma_2^2)$, $i = 1, \cdots, 5$. ($z_{0.025} = 1.96, t_{4,0.025} = 2.78, t_{4,0.01266} = 3.48, t_{4,0.0125} = 3.50, t_{8,0.025} = 2.31, t_{10,0.025} = 2.23, F_{4,4,0.025} = 9.60, F_{4,4,0.975} = 0.10, F_{5,5,0.025} = 7.15, F_{5,5,0.975} = 0.14$)

(i) (10') Construct a 95% confidence interval for $\mu_1 - \mu_2$.

(ii) (5') Construct a 95% confidence region for (μ_1, μ_2).