Statistical Inference - Homework 1

Nov. 10, 2021

Due date: Nov. 19, 2021

1. Point out the difference between joint distribution and sampling distribution. You may want to use an example to support your statement.

2. Let the population (r.v.) $X \sim B(1, p)$ (i.e., $P(X = 1) = p$, $P(X = 0) = 1 - p$, where p is an unknown parameter. $\boldsymbol{X} = (X_1, X_2, \cdots, X_4)$ is a random sample from the population X,
   - Write out the sample space and the probability distribution of $\boldsymbol{X}$;
   - Point out which of the followings are statistics: X_1+X_2, $\min_{1\le i\le 4} X_i$, X_4+2p, $X_4-E(X_1)$, $(X_4-X_1)^2/Var(X_1)$;

3. Let $X_1, \cdots, X_n$ be a random sample from normal population $X \sim N(\mu, 1)$, let $\bar{X}$ be the sample mean. How large is the sample size n enough to guarantee $P(|\bar{X} - \mu| < 0.5) \ge 0.99$?

4. Suppose X_n from binomial(n, p), where $0 < p < 1$. Find the asymptotic distribution of $g(X_n/n) - g(p)$, where $g(x) = \min\{x, 1- x\}$.

5. Show that the n-dimensional normal family $\{f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma); \boldsymbol{\mu} \in R^n,\ \Sigma \in \mathcal{M}_n\}$ is an exponential family, where $\boldsymbol{x}$ and $\boldsymbol{\mu}$ are n-dimensional column vector, $\mathcal{M}_n$ is a collection of $n \times n$ symmetric positive definite matrices and
$$f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right], \quad \boldsymbol{x} \in R^n.$$

6. Gamma distributions belongs to an exponential family. The p.d.f. of Gamma distribution is
$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-x/\theta}.$$
Please write the natural (canonical) form and specify the corresponding natural parameter space.

7. Is the family of Weibull distributions with two unknown parameters α and β an exponential family? The p.d.f. of Weibull distribution is
$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^\beta}, \quad x > 0 \quad (\alpha, \beta > 0).$$

8. Let r.v.'s $X_1, \cdots, X_n$ *i.i.d.* $\sim N(\theta, \theta^2)$, is $\bar{X}$ a sufficient statistic of θ?

9. Let $X_1, \cdots, X_m$ *i.i.d.* $\sim N(a, \sigma^2)$, $Y_1, \cdots, Y_n$ *i.i.d.* $\sim N(b, \sigma^2)$ and X_i's and Y_j's are independent. Let $\bar{X} = \sum_{i=1}^m X_i/m$, $\bar{Y} = \sum_{j=1}^n Y_j/n$, and
$$S^2 = \frac{1}{n + m - 2}\left[\sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2\right].$$
Show that $(\bar{X}, \bar{Y}, S^2)$ is a sufficient statistic of (a, b, σ^2).