Yau Mathematical Competition 2018
Probability and Statistics Team

Problem 1 (Probability) Let $\{X_n\}$ be a sequence of independent and identically distributed random variables with the distribution $\mathbb{P}\{X_n = 1\} = \mathbb{P}\{X_n = -1\} = 1/2$. Define

$$Z = \sqrt{\frac{1}{2} + \frac{X_1}{2}\sqrt{\frac{1}{2} + \frac{X_2}{2}\sqrt{\frac{1}{2} + \cdots}}}.$$

(1) Let

$$Z_N = \sqrt{\frac{1}{2} + \frac{X_1}{2}\sqrt{\frac{1}{2} + \frac{X_2}{2}\sqrt{\frac{1}{2} + \cdots X_N\sqrt{\frac{1}{2}}}}}$$

be the random variable Z truncated at the nth step. Show that

$$Z_N = \sin\left(\frac{\pi}{4}\sum_{n=0}^{N}\frac{X_1X_2\cdots X_n}{2^n}\right).$$

(2) Let

$$Y_n = X_1X_2\cdots X_n, \quad n = 1, 2, \ldots.$$

What is the joint distribution of the random variables $\{Y_n\}$?

(3) Find the distribution function F_Z of the random variable Z.

Problem 2 (Statistics) For $n \geq 2$, let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be independent, identically distributed random vectors, with a common distribution which is bivariate normal with two component means μ_1 and μ_2 and the variance-covariance elements

$$\mathrm{var}(X_1) = \sigma_1^2, \quad \mathrm{var}(X_2) = \sigma_2^2, \quad \mathrm{cov}(X_1, X_2) = \rho\sigma_1\sigma_2.$$

We assume that σ_1 and σ_2 are both positive. Let $\theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)^T$.

(1) Assuming that the parameter θ is known, show that if one desires to predict Y_1 by using a function $g(X_1, \cdots, X_n)$ that minimizes $\mathbb{E}_\theta(Y_1 - g(X_1, \cdots, X_n))^2$, then the solution is given by

$$g(X_1, \cdots, X_n) = \beta_0 + \beta_1 X_1.$$

Provide expressions for β_0 and β_1 in terms of θ.

(2) Assuming that the parameter θ is unknown, how do you predict Y_1 and how do you measure the uncertainty of your prediction?