S.-T. Yau College Student Mathematics Contests 2018

Probability and Statistics

Individual (5 problems)

Problem 1. A submarine is lost in some ocean. There are two (and only two) possible regions: A and B. Experts estimate the probability of being lost in A is 70%. On the other hand, for each search, the probability of finding this submarine is 40% if it is lost in A. This number is 80% for region B. Now we have independently searched region A 4 times and region B once, but still have not found the submarine yet. Now based on these informations, which region we should search next? And why?

Problem 2. A teacher and 12 students sit around a circle. In the beginning the teach holds a gift, he will randomly pass it to the left person or right person next to him, so as the other students each time. (For the gift, It is like a random walk between these people) The rule is that the gift will be eventually given to some student (not teacher) if he/she

is the last student who ever touches the gift.

Which student(s) have the highest probability to get this gift (i.e., win) ?

Problem 3. In a party, N people attend, each of them brings k gifts. When they leave, each of them randomly picks k gifts. Let X be the total number of gifts which are taken back by their owners. Let's fix k, please find the limiting distribution of X when $N \to \infty$.

Problem 4. Suppose that a random vector $\mathbf{x} = (x_1, ..., x_n)' \in R^n (n \geq 2)$ is distributed as a multivariate normal distribution $N(\mathbf{0}, \Sigma)$ with the following joint probability density function

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}} det(\Sigma)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}\mathbf{x}'\Sigma^{-1}\mathbf{x}\right\}, \ \mathbf{x} \in R^n,$$

where Σ is an $n \times n$ positive definite matrix. Let the (i, j) element of $\Omega = \Sigma^{-1}$ be ω_{ij} $(1 \leq i, j \leq n)$. For $1 \leq i \neq j \leq n$, show that if $\omega_{ij} = 0$, then x_i and x_j are conditionally independent when the other elements of $\mathbf{x}$ are given.

Problem 5. Let$\mathbf{x}, \mathbf{y}$ be two independent random vectors in R^n $(n \geq 3)$. Assume that $P(\mathbf{y} = \mathbf{0}) = 0$ and $\mathbf{x}$ has a standard multivariate normal distribution, i.e., $\mathbf{x} \sim N(0, I_n)$.

(a) For any nonzero constant vector $\mathbf{a} \in R^n$ satisfying $||\mathbf{a}|| = (\mathbf{a}'\mathbf{a})^{1/2} = 1$, prove that

$$\sqrt{n-1}\frac{\mathbf{a}'\mathbf{x}}{\sqrt{||\mathbf{x}||^2 - (\mathbf{a}'\mathbf{x})^2}} \sim t_{n-1},$$

here t_{n-1} stands for a t distribution with $n - 1$ degrees of freedom.

(b) The sample correlation coefficient between $\mathbf{x} = (x_1, ..., x_n)'$ and $\mathbf{y} = (y_1, ..., y_n)'$ is defined as

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}.$$

where $\bar{x} = \sum_{i=1}^n x_i/n, \bar{y} = \sum_{i=1}^n y_i/n$. Show that $\sqrt{n-2}\frac{r}{\sqrt{1-r^2}} \sim t_{n-2}$.