Yau College Math Competition 2020 Final Probability and Statistics

Individual Exam Problem Set 1 (Saturday, October 24, 2020)

PROBLEM 1. Suppose that $\{X_n\}$ is a sequence of independent, identically distributed random variables with the uniform distribution on the unit interval $[0,1]$. For each $x \in [0,1]$, define

$$X_n^x = \begin{cases} 1, & X_n \le x; \\ 0, & X_n > x. \end{cases}$$

Let $f : [0,1] \to \mathbb{R}$ be an nondecreasing continuous function on $[0,1]$ and

$$B_n(x; f) = \mathbb{E}\left[f\left(\frac{X_1^x + \cdots + X_n^x}{n} \right) \right].$$

Show that

(1) $B_n(x; f)$ is a polynomial in x of degree n; (2) $B_n(x; f)$ is nondecreasing in x;

(3) $B_n(x; f) \to f(x)$ uniformly on $[0,1]$.

PROBLEM 2. An urn contains N balls marked $1, 2, ..., N$. A ball is drawn from the urn repeatedly and independently with replacement. Let T_N be the first time every ball in the turn has been drawn at least once. Show that $T_N / N \log N$ converges to 1 in probability.

PROBLEM 3. Suppose $\{X_1, \ldots, X_n\}$ is a random sample from an unknown probability distribution with finite mean, variance, and third central moment, denoted by μ, σ^2, and $\mu_3 = \mathbb{E}(X_1 - \mu)^3$, respectively. It is of interest to study the relationship between

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \quad \text{and} \quad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

(1) Show that they are independent when the underlying distribution is Gaussian.
(2) For a general distribution, what is $\operatorname{cov}(\bar{X}, S^2)$? Find an expression.
(3) Suppose the random sample is from Bernoulli(1/2). Show that $\bar{X}$ and S^2 are uncorrelated, but are not independent by showing that

$$\mathbb{P}(S^2 = 0 \mid \bar{X} = 1) \neq \mathbb{P}(S^2 = 0)$$

Yau College Math Competition 2020 Final Probability and Statistics

Individual Exam Problem Set 2 (Sunday, October 25, 2020)

PROBLEM 1. Suppose that $\{X_n\}$ is a sequence of real valued, independent, identically distributed random variables and B is a Borel set in $\mathbb{R}$. Assume that $\mathbb{P}(X_1 \in B) > 0$. Let $T = \inf\{n : X_n \in B\}$ be the first time the sequence is in the set B.

(1) Show that $\mathbb{P}(T < \infty) = 1$.

(2) Suppose $\mathbb{E}|X_1| < \infty$. Show that $\mathbb{E}X_T = \mathbb{E}[X_1 I_B(X_1)]\mathbb{E}T$.

PROBLEM 2. We flip a fair coin repeatedly and independently. Let N_n be the number of consecutive heads beginning from the n^{th} flip. (For example, $N_n = 0$ if the n^{th} flip is a tail, and $N_n = 2$ if the n^{th} and $(n+1)^{\text{th}}$ flips are heads but the $(n+2)^{\text{th}}$ flip is a tail. Show that

$$\limsup_{n\to\infty} \frac{N_n}{\log n} = \frac{1}{\log 2}.$$

PROBLEM 3. Let $\{X_1, \ldots, X_n\}$ be independent and identically distributed from the uniform distribution on the interval $(-\theta, \theta)$ with $\theta > 0$.

(1) Find a minimal sufficient statistic T for θ.

(2) Define $V = \bar{X}/|X|_{(n)}$, where $|X|_{(n)} = \max(|X_1|, \ldots, |X_n|)$. Show that V is independent of T.

Yau College Math Competition 2020 Final Probability and Statistics

Overal Exam Problem Set (Saturday, October 24, 2020)

PROBLEM 1. Let $\{X_n\}$ be a sequence of independent, identically distributed random variables taking values in $\mathbb{N}$, the set of positive integers. Define

$$R_n = \mathrm{Card}\{X_1, \cdots, X_n\},$$

i.e., R_n is the number of distinct integers in the set $\{X_1, \cdots, X_n\}$. Suppose that $\mathbb{E}[X_1] < \infty$. Prove that $R_n/\sqrt{n} \to 0$ in probability.

PROBLEM 2. Consider the mixture experiment whose components are E_1 and E_2, taken with equal probabilities and each with the sample space $\Omega = \{0, 1, 2, ...\}$. It is postulated that

(1) the outcome of E_1 follows the Binomial model Binomial(n, θ) with the total number of trials $n > 0$ and the unknown probability of success parameter $\theta \in [0, 1]$, and
(2) the outcome of E_2 follows the Negative-Binomial model Binomial(r, θ) with the target for number of successful trials $r > 0$ and the unknown probability of success parameter $\theta \in [0, 1]$.

Let H be the observed index to the experiment that is actually conducted, and let X denote the observed outcome of the conducted experiment.

Find a minimal sufficient statistic for θ and prove your claim.