S.-T. Yau College Student Mathematics Contests 2014

Probability and Statistics Problems

Individual

Please solve the following 5 problems.

Problem 1. Let X be a real valued random variable such that for all smooth functions $f : R \to R$ with compact support we have $E[Xf(X)] = E[f'(X)]$. Show that X has the standard normal distribution.

Problem 2. Let (X_n) be a sequence of uncorrelated random variables of mean zero such that

$$\sum_{n=1}^{\infty} nE|X_n|^2 < \infty.$$

Show that $S_n = \sum_{i=1}^n X_i$ converges almost surely.

Problem 3. Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. Let P and Q be two probabilities which are mutually absolutely continuous on $\mathcal{F}$. We denote by X_0 the Radon-Nikodym density of Q with respect to P on $\mathcal{F}$. Show that the following two properties are satisfied:

(a) $0 < E_P[X_0|\mathcal{G}] < +\infty$, P-a.s.;

(b) for every $\mathcal{F}$-measurable non-negative random variable f,

$$E_P[fX_0|\mathcal{G}] = E_Q[f|\mathcal{G}]E_P[X_0|\mathcal{G}].$$

Problem 4. Suppose $X_1, \ldots, X_n, \ldots$ is a sequence of random numbers drawn from the uniform distribution $U(0,1)$. One observes these numbers sequentially. At time n, one keeps a record of $Y_n \overset{def}{=} X_{(n)} = \max_{i=1}^n X_i = \max\{Y_{n-1}, X_n\}$ and $Z_n \overset{def}{=} \bar{X}_n = \sum_{i=1}^n X_i/n = (n-1)/nZ_{n-1} + 1/nX_n$ and discards all previous recordings.

(a) What is the best guess of X_1 if one only observes Y_n?

(b) What is the best guess of X_1 if one only observes Z_n?

(c) Comparing the two guesses of X_1, which one is better (and in what sense)?

Give good reasoning to justify your answers.

Problem 5. Suppose we take a random sample of size n from a bag of colored balls (red, blue and yellow balls) with replacement. Let X_1 denote the number of red balls, X_2 denote the number of blue balls, and X_3 denote the number of yellow balls in the sample. Assuming we know that the total number of yellow balls is triple the total number of red balls in the bag. Or in other words, the red, blue and yellow balls occur with probability p_1, p_2 and $p_3 = 3p_1$, respectively in the bag.

1. Find the aymptotic distribution (after appropriate normalization) for the MLE of p_2.

2. Construct the likelihood ratio test statistic for the null hypothesis that $p_1 = p_2 = p_3/3$ (the alternative is that $p_1 = p_2 = p_3/3$ is not true). What is the asymptotic distribution of your test statistic under null?