**Statistics 930: Probability Theory**
**Homework No. 9 —**

PROBLEM 1. Suppose that $\vec{a} = (a_1, a_2, ..., a_d)$ is a point chosen at random from the sphere in d-dimensions with radius one:

$$S_d = \{\vec{a} : |\vec{a}| = 1\}.$$

Show that as $d \to \infty$ the random variable $a_1\sqrt{d}$ converges in distribution to a normal with mean zero and variance one.
HINT: You could go nuts trying all sorts of complicated ways of proving this very easy fact. Keep calm. Don't jump to the first idea that comes into your head. Instead, think how you could use $d$ independent random $Y_j \sim N(0,1)$ to simulate the random vector $\vec{a}$. You may then see that the problem is an almost immediate consequence of the LLN! No characteristic functions or any other funny business is needed. Isn't it curious that we get a non-trivial CLT from the LLN? I find it lovely.

It's not part of this problem, but sometime you should ponder that the ball $B_d = \{\vec{a} : |\vec{a}| \leq 1\}$ in high dimensions behaves more like an "annulus" than like a ball in lower dimension. For example, asymptotically essentially all of the mass of $B_d$ is in the "annulus" $A_d(\epsilon) = \{\vec{a} : 1 - \epsilon \leq |\vec{a}| \leq 1\}$. Properly viewed, this is just a version of the weak law of large numbers; no Gamma function calculations are required.

PROBLEM 2. Suppose that a random variable $X$ has characteristic function

$$Ee^{itX} = p(e^{it})$$

where $p$ is real polynomial of degree $n$. Note that this implies that the coefficients of $p$ are nonnegative and that these coefficients sum to 1. Now, suppose that all of the roots of the polynomial are real.

Show that $X$ is equal in distribution to $Y_1 + Y_2 + \cdots + Y_n$ where the random variables $Y_j$, $j = 1, 2, ..., n$ are independent (but not necessarily identically distributed) Bernoulli random variables. Hint: Stay calm and write out in symbols what the hypothesis gives you. Remind yourself what the characteristic function of a general Bernoulli random variable must look like. Now, fit the pieces of the puzzle together.

PROBLEM 3. Suppose that $X_j$, $j = 1, 2, ...$ are i.i.d. with the "tent" characteristic function

$$\phi(t) = (1 - |t|)_+$$

- Let $S_n = X_1 + X_2 + \cdots + X_n$ and prove that $S_n/n$ converges in distribution to the standard Cauchy distribution. Hints: Note we normalize by $n$ not $\sqrt{n}$. This problem is super-easy, but you will have to look up the characteristic function of the Cauchy if you don't know it.
- What is the median of $X_1$? What is the value of $E|X_1|$? Justify your answers. Be sure of your logic. It may help to argue by contradiction.
- Show that if $Y$ is *any* random variable, then

$$\psi(t) = E(\{1 - |tY|\}_+)$$

is the characteristic function of some random variable $Z$. What experiment (or simulation) would give us an observation with this characteristic function?

PROBLEM 4. If $(p_1, p_2, \ldots, p_M)$ describes a probability mass function on the set $S = \{1, 2, \ldots, M\}$ then the associated entropy is defined by

$$H(p_1, p_2, \ldots, p_M) = -\sum_{j=1}^{M} p_j \log p_j.$$

We also know (or you can assume) that for all $(p_1, p_2, \ldots, p_M)$ we have

$$H(p_1, p_2, \ldots, p_M) \leq \log M$$

and if $H(p_1, p_2, \ldots, p_M) = \log M$ then $p_j = 1/M$ for all $j = 1, 2, \ldots, M$. That is, the entropy of an $S$ valued random variable is uniquely maximized by the uniform distribution on $S$.

If $X$ is an $S$ valued random variable and $P(X = j) = p_j$, then by a natural abuse of notation we also write $H(X)$ for $H(p_1, p_2, \ldots, p_M)$. Show that if $X_n$ is a sequence of random variables such that $H(X_n)$ converges to $\log M$, then $X_n$ converges in distribution to the uniform distribution on $S$.

You may want to pattern your argument on the proof of Levy's continuity theorem. This is our fourth example of this argument, and we may see two more.

PROBLEM 5. (a) If $f$ and $g$ are monotone increasing functions and $X$ is any random variable then we have

$$E[f(X)]E[g(X)] \leq E[f(X)g(X)]$$

whenever both sides of the equation make sense.

(b) Suppose $f : \mathbb{R} \to \mathbb{R}$ satisfies $|f(x) - f(y)| \leq |x - y|$. Show that

$$\mathrm{Var}[f(X)] \leq \mathrm{Var}[X].$$

TOO GIANT SIZED HINTS: For part (a) introduce another random variable $X'$ that is independent of $X$ and has the same distribution as $X$. Now consider the random variable $Z = \{f(X) - f(X')\}\{g(X) - g(X')\}$. For part (b) use a similar trick after recalling the relationship between $E[(X - X')]^2$ and $\mathrm{Var}[X]$.

There are many problems where one benefits by the introduction of "independent copies" of some random variable. To wax philosophical for a moment, this is one of the unique features of the "category of probability spaces."