Probability Theory

1. *Characteristic Functions:* Calculate the characteristic function, the mean, and the variance of the following probability density functions:

(a) *Uniform* $p(x) = \frac{1}{2a}$ for $-a < x < a$, and $p(x) = 0$ otherwise;

(b) *Laplace* $p(x) = \frac{1}{2a}\exp\left(-\frac{|x|}{a}\right)$;

(c) *Cauchy* $p(x) = \frac{a}{\pi(x^2+a^2)}$.

The following two probability density functions are defined for $x \geq 0$. Compute only the mean and variance for each.

(d) *Rayleigh* $p(x) = \frac{x}{a^2}\exp(-\frac{x^2}{2a^2})$,

(e) *Maxwell* $p(x) = \sqrt{\frac{2}{\pi}}\frac{x^2}{a^3}\exp(-\frac{x^2}{2a^2})$.

****

2. *Directed Random Walk:* The motion of a particle in three dimensions is a series of independent steps of length ℓ. Each step makes an angle θ with the z axis, with a probability density $p(\theta) = 2\cos^2(\theta/2)/\pi$; while the polar angle ϕ is uniformly distributed between 0 and 2π. (Note that the solid angle factor of $\sin\theta$ is already included in the definition of $p(\theta)$, which is correctly normalized to unity.) The particle (walker) starts at the origin and makes a large number of steps N.

(a) Calculate the expectation values $\langle z\rangle$, $\langle x\rangle$, $\langle y\rangle$, $\langle z^2\rangle$, $\langle x^2\rangle$, and $\langle y^2\rangle$, and the covariances $\langle xy\rangle$, $\langle xz\rangle$, and $\langle yz\rangle$.

(b) Use the central limit theorem to estimate the probability density $p(x, y, z)$ for the particle to end up at the point (x, y, z).

****

3. *Tchebycheff's Inequality:* Consider any probability density $p(x)$ for $(-\infty < x < \infty)$, with mean λ, and variance σ^2. Show that the total probability of outcomes that are more than $n\sigma$ away from λ is less than $1/n^2$, i.e.

$$\int_{|x-\lambda|\geq n\sigma} dx p(x) \leq \frac{1}{n^2}.$$

Hint: Start with the integral defining σ^2, and break it up into parts corresponding to $|x-\lambda| > n\sigma$, and $|x-\lambda| < n\sigma$.

****

4. *Optimal Selections:* In many specialized populations, there is little variability among the members (e.g. in the GRE scores of the 8.333 students compared to GRE scores of a random group.) Is this a natural consequence of optimal selection?

(a) Let $\{r_\alpha\}$ be n random numbers, each independently chosen from a probability density $p(r)$, with $r \in [0, 1]$. Calculate the probability density $p_n(x)$ for the largest value of this set, i.e. for $x = \max\{r_1, \cdots, r_n\}$.

(b) If each r_α is uniformly distributed between 0 and 1, calculate the mean and variance of x as a function of n, and comment on their behavior at large n.

****

5. *Information:* Consider the velocity of a gas particle in one dimension ($-\infty < v < \infty$).

(a) Find the unbiased probability density $p_1(v)$, subject only to the constraint that the average *speed* is c, i.e. $\langle |v| \rangle = c$.

(b) Now find the probability density $p_2(v)$, given only the constraint of average kinetic energy, $\left\langle mv^2/2 \right\rangle = mc^2/2$.

(c) Which of the above statements provides more information on the velocity? Quantify the difference in information in terms of $I_2 - I_1 \equiv \left(\langle \ln p_2 \rangle - \langle \ln p_1 \rangle\right) / \ln 2$.

****

(Optional) 6. *Benford's Law* describes the observed probabilities of the *first digit* in a great variety of data sets, such as stock prices. Rather counter-intuitively, the digits 1 through 9 occur with probabilities 0.301, .176, .125, .097, .079, .067, .058, .051, .046 respectively. The key observation is that this distribution is invariant under a change of scale, e.g. if the stock prices were converted from dollars to persian rials! Find a formula that fits the above probabilities on the basis of this observation. (*Hint:* Think about random multiplicative processes.)

****

Suggested Reading: S.-K. Ma, Statistical Mechanics, Part III.