**Temas avanzados de termodinámica y física estadística – 1er. cuatrimestre de 2013**

**Guía 1:** *Probabilidad*

1. (Van Kampen §1.1, ec. 1.5.) Dos volúmenes $V_1$ y $V_2$ se comunican a través de una abertura y contienen en total $N$ moléculas distinguibles y no interactuantes. Postule probabilidades razonables, $p_1$ y $p_2$, de que una dada molécula esté en $V_1$ o $V_2$, respectivamente. Demuestre que la probabilidad de que en $V_1$ haya exactamente $n$ moléculas está dada por una binomial,

$$P(n) = \binom{N}{n} p_1^n \, p_2^{N-n}.$$

2. (Jaynes §8.12.4.) Dos personas, $A$ y $B$, lanzan alternativamente una moneda. Gana el primero que obtiene cara. Si $A$ hace el primer lanzamiento, calcule las probabilidades que tiene cada uno de ganar. (*Sugerencia:* hay infinitos caminos que llevan a uno u otro ganador y la probabilidad correspondiente puede obtenerse sumando la probabilidad de cada camino independiente; trate de dar a esta proposición una forma rigurosa. Al margen de lo anterior, la simetría del problema permite llegar más rápidamente al resultado: note que luego del primer lanzamiento, si $A$ no gana, a partir del segundo lanzamiento los roles de $A$ y $B$ se intercambian; formalice esto último y obtenga, en un par de pasos, el resultado.)

3. **Problème des rencontres.** (Van Kampen §1.1, "Matching problem".) Hay $n$ objetos distintos, dispuestos según un cierto orden inicial en $n$ lugares diferentes. Tiene lugar una permutación al azar de estos objetos. Se trata de encontrar la probabilidad $p(n)$ de que ningún objeto vuelva a su posición inicial, asumiendo que todas las permutaciones tienen igual probabilidad $1/n!$. Para eso demuestre que $p(n)$ satisface la siguiente relación de recurrencia,

$$np(n) - (n-1)p(n-1) = p(n-2).$$

Encuentre $p(n)$ y muestre que $p(n) \to e^{-1}$ cuando $n \to \infty$. (*Sugerencia:* defina la función generatriz $F(x) = \sum_n x^n p(n)$ y transforme la relación de recurrencia para $p$ en una ecuación diferencial para $F$. Este método es de uso muy extendido.

4. **El problema del cumpleaños.** En un aula hay $n$ personas. Considerando, para simplificar, que la probabilidad de que una persona cumpla años un determinado día es $1/365$, la misma para todos los días del año, y que las fechas de los cumpleaños de las $n$ personas son estadísticamente independientes:

   (a) ¿Cuántas personas debe haber en el aula para que la probabilidad $p(n)$ de que al menos dos cumplan años el mismo día supere el 50%? Grafique $p(n)$ con la ayuda de algún programa.

   (b) Una expresión aproximada de $p(n)$ puede obtenerse si analizan las personas de a pares. i) Calcule la probabilidad de que un dado par de personas no cumpla años el mismo día. ii) Calcule cuántos pares hay en un aula con $n$ personas. iii) Si se asume que cada evento de la forma "$A$ y $B$ no cumplen años el mismo día" (al que notaremos como $AB$), es independiente de todos los otros, encuentre la probabilidad de que ningún par de personas cumpla años el mismo día, y de ahí obtenga $p_{\text{aprox}}(n)$. ¿En dónde entra la aproximación? Grafique $p_{\text{aprox}}(n)$ y compare con la solución exacta.

(c) En verdad, no todos los eventos de la forma "$A$ y $B$ no cumplen años el mismo día" son independientes entre sí. Para ver eso en un caso sencillo, suponga que hay 3 personas, $M$, $L$ y $C$, y calcule $P(ML, MC, LC)$, $P(ML, MC)$, $P(ML|MC)$ y $P(ML|MC, LC)$.

5. **Falso positivo.** Aparece una nueva enfermedad, 100% fatal aunque extremadamente rara, estimándose que se da en 1 de cada 1000 millones de personas. Sus síntomas son imperceptibles, hasta que eventualmente la cabeza explota. Por suerte, se descubre un test de diagnóstico prácticamente infalible: la probabilidad de que el test falle y dé positivo al ser ensayado en una persona sana es de 1 en un millón, lo que se conoce como un *falso positivo*. Existe una probabilidad igual de que el test falle al ser aplicado a una persona que sí tiene la enfermedad.

   (a) Una persona se hace el test y le da positivo. Teniendo en cuenta que el test falla en un caso de cada un millón, ¿debe desahuciarse a la persona o darle alguna esperanza? Concretamente, ¿cuál es la probabilidad de que no tenga la enfermedad? (*Sugerencia:* el hecho de formular esta pregunta indica que la respuesta no es "1 en un millón".)

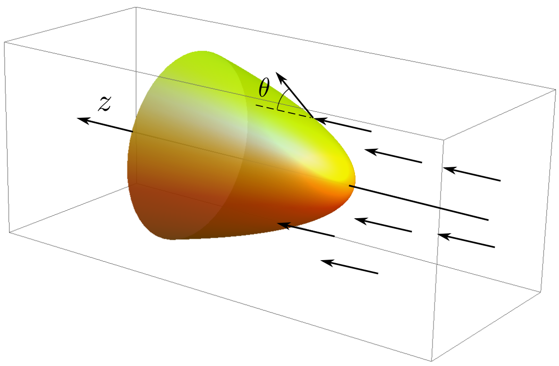
   (b) Si a una persona el test le da negativo, ¿cuál es la probabilidad de que sí tenga la enfermedad?

   (c) Generalice sus resultados para valores arbitrarios de las probabilidades que aparecen como dato en el enunciado.

   (Ver `http://www.pagina12.com.ar/diario/contratapa/13-187588-2012-02-15.html`)

**Funciones de variables aleatorias.**

6) **Dispersión de partículas.**




Considere un haz de partículas independientes que se mueven en la dirección $\hat{z}$. El haz tiene simetría cilíndrica respecto del eje $z$. Cualquier sección transversal al haz es atravesada por una partícula por unidad de tiempo, con una distribución $I$ que, debido a la simetría, sólo depende de la coordenada $\rho$:

$$\int dA\ I(\mathbf{r}) = 2\pi \int_0^\infty \rho d\rho\ I(\rho) = 1,$$

donde

$$I(\mathbf{r})dA = \left\{ \begin{array}{l} \text{probabilidad de que durante un intervalo de tiempo unidad} \\ \text{una partícula pase a través del elemento de área } \hat{z}\, dA \text{ en } \mathbf{r} \end{array} \right\}.$$

El haz es dispersado elásticamente por un paraboloide de revolución, $z(\rho) = \alpha\rho^2$ con $\alpha > 0$, como muestra la figura. Encuentre la distribución de la intensidad de las partículas dispersadas, $I_d(\theta)$, en función del ángulo de dispersión. Verifique que esta distribución integra a 1.

7) (Reichl, §4.) Las variables aleatorias $X$ e $Y$ son independientes, con distribución gaussiana centrada en el cero y desviación estándar $\sigma_X = \sigma_Y = 1$. Encuentre la función característica para la variable $Z = X^2 + Y^2$ y calcule sus tres primeros momentos.

8) (Reichl, §4.) Las variables aleatorias $X_1, \ldots, X_N$ son independientes y tienen la misma densidad $p_{X_i}(x) \equiv p(x)$. La suma $S = X_1 + \ldots + X_N$ puede representar, por ejemplo, el desplazamiento de una caminata al azar luego de $N$ pasos. Encuentre la densidad de $S$ si:

   (a) Los desplazamientos elementales son discretos, $p(x) = p\delta(x-a) + q\delta(x+a)$, con $q = 1-p$.

   (b) Los desplazamientos elementales $X_i$ son gaussianos, $p(x) = (2\pi\sigma^2)^{-1/2} \exp\left[-(x-a)^2/(2\sigma^2)\right]$.

   (c) Los desplazamientos elementales siguen una distribución de Cauchy, $p(x) = a/[\pi(a^2 + x^2)]$.

   *Sugerencia:* calcule la función característica de $S$ o use el desarrollo de Fourier de la delta de Dirac $\delta(s - x_1 - \ldots - x_N)$.

9) Sea $X$ una variable aleatoria con distribución acumulativa $F(x)$, es decir, $\mathrm{Prob}(X \leq x) = F(x)$. Demuestre que $X$ puede generarse a partir de una variable aleatoria $R$ con distribución uniforme en el intervalo $\overline{0,1}$, definiendo

$$X = F^{-1}(R).$$

10) Este problema es el inverso del anterior: igual que antes sea $F(x) = \mathrm{Prob}(X \leq x)$. Encuentre la función de distribución acumulativa y de densidad de probabilidad para la variable aleatoria $Y$ definida como $Y = F(X)$.

**Direcciones y vectores aleatorios** (Feller § I.10.)

11) En $\Re^3$ la dirección de un vector unitario $\mathbf{\Omega}$ puede darse mediante los ángulos esféricos $\varphi$ y $\theta$, con $\varphi \in \overline{0, 2\pi}$ y $\theta \in \overline{0, \pi}$. Estos vectores pueden representar, por ejemplo, estrellas en el cielo o la orientación de los planos de galaxias espirales en un cúmulo. Se dice que la variable aleatoria $\mathbf{\Omega}$ está distribuida uniformemente si su densidad es independiente de la dirección, $P(\mathbf{\Omega}) = \text{cte}$. Es decir, la probabilidad de que el vector unitario apunte en la dirección $\mathbf{\Omega}$ dentro de un ángulo sólido $d\Omega$ es independiente de $\mathbf{\Omega}$. Bajo estas condiciones:

   (a) Calcule la constante $P(\mathbf{\Omega})$.

   (b) Escriba la densidad correspondiente para las variables angulares esféricas, $p(\varphi, \theta)$.

   (c) En esféricas, suele ser más cómodo trabajar con $\xi = \cos\theta$. Encuentre la densidad $f(\varphi, \xi)$.

   (d) Demuestre que la densidad de probabilidad para la proyección de $\mathbf{\Omega}$ sobre cualquier eje fijo es uniforme en $\overline{-1, 1}$.

   (e) Esta uniformidad de la proyección sobre un eje no tiene mayor relación con la uniformidad de la distribución de $\mathbf{\Omega}$, sino con la dimensionalidad del espacio. Para ver que no vale en otras dimensiones, encuentre la densidad de la proyección sobre un eje de un vector uniformemente distribuido en un círculo en $\Re^2$.

(f) Para hacer en la computadora, como aplicación de lo anterior y del problema 9): use un generador de números aleatorios, uniforme en $\overline{0,1}$, para simular una muestra de $n$ direcciones aleatorias uniformemente distribuidas en la esfera. Hágalo en términos de las variables $\varphi$ y $\theta$, es decir, genere valores al azar de $\varphi$ y de $\theta$ de acuerdo a la densidad $p(\varphi,\theta)$ calculada antes. Grafique varias de estas muestras. Grafique también un histograma para las proyecciones sobre alguno de los ejes; tome $n \sim 10000$ y compruebe cualitativamente que la distribución es uniforme en $\overline{-1,1}$.

12) Siguiendo con el problema anterior, un vector aleatorio en $\Re^3$ queda definido por una dirección aleatoria $\mathbf{\Omega}$ y un módulo $L$, que es otra variable aleatoria en $\overline{0,\infty}$ con densidad $v(L)$.

(a) Qué densidad $v(L)$ debe asociarse a una distribución de vectores isótropa y homogénea, acotada por un radio máximo $R$ (por ejemplo, la posición de una molécula de agua dentro de una gota esférica y homogénea). Escriba la densidad completa $p(\varphi,\xi,L)$.

(b) Para hacer en la computadora: use un generador de números aleatorios, uniforme en $\overline{0,1}$, para simular una muestra de $n$ vectores aleatorios uniformemente distribuidos en el interior de una esfera de radio 1. Grafique varias muestras.

(c) En la práctica, si la distribución es isótropa, basta analizar la proyección sobre cualquier eje fijo para deducir $v(L)$. Es decir, $v(L)$ puede obtenerse conociendo, por ejemplo, la densidad $f(L_z)$ asociada a la longitud de la proyección del vector aleatorio sobre el eje $z$, evitando así tener que hacer un análisis en 3 dimensiones. Concretamente, demuestre que

$$v(L) = -Lf'(L).$$

*Sugerencia:* deduzca primero una relación entre $v(L)$ y la distribución acumulativa de $L_z$; derive para encontrar $f$ e invierta la relación.

**Estadística de orden.**

13) $X_1$ y $X_2$ son dos variables aleatorias independientes, con distribución uniforme en el intervalo $\overline{0,1}$.

(a) Escriba las densidades de probabilidad de $(X_1,X_2)$, $X_1+X_2$ y $X_2-X_1$.

(b) ¿Cuál es la probabilidad de que sea $X_2 > X_1$? (*Sugerencia:* use la simetría del problema.)

(c) Se definen $Z_1 = \min(X_1,X_2)$ y $Z_2 = \max(X_1,X_2)$. ¿Cuáles son las densidades de probabilidad de $(Z_1,Z_2)$, $Z_1$, $Z_2$ y $Z_2-Z_1$.

(d) ¿Cuál es el valor medio de la distancia $L = Z_2 - Z_1$ entre $X_1$ y $X_2$.

14) El problema anterior puede llevarse a $n$ variables aleatorias idénticas e independientes, $(X_1,\ldots,X_n)$, cada una con densidad de probabilidad $p(x)$. En función de estas variables se define otro conjunto de $n$ variables aleatorias, cuyas realizaciones consisten en los valores de los $X_i$ ordenados de menor a mayor, $(Z_1,Z_2,\ldots,Z_n) = \mathcal{O}(X_1,X_2,\ldots,X_n)$. Por ejemplo, si en una realización $(X_1,X_2,X_3) = (2,3,1)$ entonces $(Z_1,Z_2,Z_3) = (1,2,3)$. (Si las densidades son continuas, no tiene importancia considerar los casos en que dos variables toman el mismo valor.)

(a) Demuestre que la densidad de probabilidad conjunta es

$$P(Z_1, Z_2, \ldots, Z_n) = n!\, \Theta(Z_n, Z_{n-1}, \ldots, Z_1) \prod_{i=1}^{n} p(Z_i),$$

donde $\Theta(Z_n, \ldots, Z_1)$ es 1 si $Z_n \geq Z_{n-1} \geq \ldots \geq Z_1$ y cero en caso contrario. (*Sugerencia:* con la solución a la vista, no es difícil llegar a una justificación intuitiva en términos de la simetría del problema, pues las $n$ variables $X_i$ son equivalentes. Sin embargo, es instructivo tomar un camino más formal y fácilmente generalizable a otras situaciones menos simétricas: primero escriba la densidad conjunta $p(z_1, \ldots, z_n; x_1, \ldots, x_n)$ en términos de una densidad condicional y de la densidad conjunta de las variables $X_i$, y luego integre sobre estas últimas para obtener la densidad marginal $p(z_1, \ldots, z_n)$. De manera alternativa, mediante funciones escalón como la definida más arriba, no es difícil escribir explícitamente el vector $(Z_1, \ldots, Z_n)$ como una función de $(X_1, \ldots, X_n)$. Luego es posible aplicar la fórmula para la densidad de una función de una variable aleatoria, que en este caso es un vector.)

(b) (Feller §I.6.) Como aplicación de lo anterior, suponga que las $n$ variables $X_i$ tienen una densidad de probabilidad exponencial, $p(t) = \alpha e^{-\alpha t}$, con $t > 0$. Por ejemplo, cada $X_i$ puede representar el tiempo que logra sobrevivir una partícula inestable antes de decaer, o el tiempo de espera en una fila. A partir de la densidad conjunta de las $Z_i$, encuentre $\mathrm{Prob}(Z_1 > t)$ y $\mathrm{Prob}(Z_2 - Z_1 > t)$. (No es el método más sencillo pero es el de aplicación más general. Con los resultados a la vista, trate reobtenerlos con argumentos que no requieran mayores cálculos.)

(c) (Feller §I.13.) Si las $X_i$ tienen una distribución uniforme en el intervalo $\overline{0,1}$, demuestre que la densidad de probabilidad de la amplitud $L = Z_n - Z_1$ es $n(n-1)x^{n-2}(1-x)$. (*Sugerencia:* A partir de la densidad conjunta de las $n$ variables $Z_i$, encuentre la densidad conjunta de $Z_1$ y $Z_n$ y de ahí la densidad de probabilidad de $L$, teniendo en cuenta que $L$ es la resta de dos variables aleatorias. De manera alternativa, con muy pocos cálculos el problema puede resolverse por simetría y combinatoria.) Demuestre que el valor medio de $L$ es $(n-1)/(n+1)$.

**Extra.**
'[...] El hombre que escribió esa nota conocía todos los datos. Nunca podría haberlos dado tan mal si no los conociera. Uno tiene que saber mucho para equivocarse en todos los puntos, como el diablo. Un hombre que miente al azar, habría dicho alguna verdad. Imagine que alguien le envía a usted a buscar una casa con una puerta verde y una persiana azul, con jardín delantero pero sin jardín trasero, con perro pero sin gato y donde sus ocupantes beben café pero no té. Usted diría, si no encontrara una casa así, que todo era inventado. Pero yo digo que no. Digo que si usted encontrara una casa donde la puerta fuerza azul y la persiana verde, donde hubiera un jardín en el fondo pero no hubiera un jardín delantero, donde abundaran los gatos pero mataran inmediatamente a los perros, donde bebieran litros de té pero el café estuviera prohibido, entonces usted habría encontrado la casa. El informador debería haber conocido esa casa concreta para ser tan exactamente inexacto.' [De "El duelo del Dr. Hirsch", en *La sabiduría del padre Brown*, de G.K. Chesterton.]
■ Calcule la probabilidad de haber encontrado la casa.


Referencias:

W. Feller, *An introduction to probability theory and its applications*, vol II.

E. T. Jaynes, G. L. Bretthorst, *Probability theory. The logic of science*.

L. E. Reichl, *A modern course in statistical physics*.

N. Van Kampen, *Stochastic Processes in Physics and Chemistry*.