Temas avanzados de termodinámica y física estadística – 1er. cuatrimestre de 2015

Guía 1: *Probabilidades*

1. (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 el resultado en un par de pasos.)

2. **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.

3. (Pécseli §1.) **El problema de los pistoleros daneses**. Después de eliminar a todos sus rivales, el último pistolero danés con vida decide, como es natural, celebrar su triunfo con una fiesta. En fin, ¡ahora tiene tantos amigos! (pues esto suele acontecerle a los ganadores) que, eligiendo sólo a los más pasables, decide reunir a 111 comensales. Tal como ha escuchado que es de uso en las recepciones de los embajadores de los mejores países del mundo, frente a cada silla dispone una tarjeta con el nombre del invitado correspondiente. Desafortunadamente, hombre de poco roce diplomático, el primer invitado en llegar no advierte este detalle y se sienta en un lugar al azar (es decir, podría incluso ocupar el lugar correcto). Resignados y corteses (pues son hijos de Odín al fin y al cabo), los otros invitados se sientan en sus lugares correctos, siempre y cuando los encuentren disponibles; caso contrario toman un lugar desocupado al azar. La pregunta, entonces, es la siguiente: ¿cuál es la probabilidad de que el último invitado en llegar se siente en el lugar que le fue originalmente asignado? La respuesta es completamente anti intuitiva.

   Sugerencia: Los casos simples de 2, 3 y hasta 4 invitados pueden analizarse directamente y servir como base al problema de los 111 invitados. Es interesante también hacer el experimento numérico en la computadora: en muy pocas líneas puede escribirse un programa que genere el ordenamiento final de los invitados, cuyo número puede variarse; ejecutándolo muchas veces y contando aquellas en que el último invitado se sienta en el lugar que le corresponde, puede estimarse la probabilidad buscada y su dependencia con el número de invitados.

4. **El problema del coleccionista.** Un coleccionista de figuritas compra de a una figurita por vez. Hay N figuritas distintas y todas aparecen con la misma probabilidad. Es decir, cada vez que compra una

figurita la probabilidad de que sea una figurita determinada es $1/N$. La variable que interesa es el número n de figuritas que tiene que comprar para reunir las N figuritas distintas.

(a) Calcule el número medio de figuritas que tiene que comprar, $\langle n \rangle$, para reunir las N figuritas distintas. (No es necesario en este punto calcular la distribución de probabilidad.) En particular, dé $\langle n \rangle$ para los casos $N = 10$, $N = 100$ y $N = 1000$.

(b) Calcule explícitamente $p(n)$ y grafique para $N = 10$, $N = 100$ y $N = 1000$.

(Ver `http://www.pagina12.com.ar/diario/contratapa/13-250187-2014-07-06.html`)

5. **El problema de las rachas de suerte.** Una moneda se tira sucesivamente n veces. En cada tiro, cara y ceca tienen probabilidad $1/2$.

   (a) ¿Cuál es la probabilidad de que en n tiros aparezcan al menos dos caras seguidas? O, equivalentemente, pero más fácil de plantear, ¿cuál es la probabilidad de que en n tiros no aparezcan nunca dos caras seguidas? El resultado involucra los números de Fibonacci.

   (b) ¿Cuál es la probabilidad de que en n tiros aparezcan al menos j caras seguidas? (Igual que antes, puede ser más fácil encontrar la probabilidad complementaria.) La solución involucra una generalización de los números de Fibonacci.

   (c) ¿Cuál es la probabilidad de que en n tiros aparezcan al menos j caras o j cecas seguidas?

   (d) Revise críticamente las afirmaciones cuantitativas que aparecen en el siguiente artículo: `http://www.pagina12.com.ar/diario/contratapa/13-149696-2010-07-18.html`

6. **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 del enfermo 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, dándola por sana.

   (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, ¿la persona tiene alguna esperanza de estar sana? Concretamente, ¿cuál es la probabilidad de que no tenga la enfermedad si el test le dio positivo?

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

   (c) Generalice sus resultados para valores arbitrarios de las probabilidades que aparecen como datos (probabilidad p de tener la enfermedad, probabilidad q de que el test falle). En especial, analice el caso en que la enfermedad en realidad no exista y todo sea un fraude para vender el test.

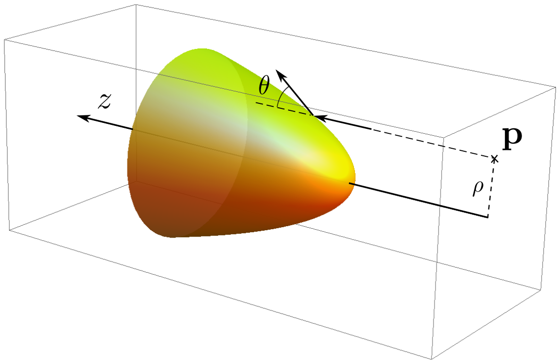
(Ver `http://www.pagina12.com.ar/diario/contratapa/13-187588-2012-02-15.html`)

7. **Dispersión de partículas.** Una partícula es dispersada elásticamente por un paraboloide de revolución, como muestra la figura. La partícula se mueve inicialmente paralela al eje del paraboloide, según una

recta que atraviesa al plano xy en el punto $\mathbf{p}$. Este punto no se conoce de antemano. La densidad de probabilidad de que las coordenadas polares de $\mathbf{p}$ sean ρ y ϕ es $I(\rho)$, independiente de ϕ. De modo que

$$\int_0^{2\pi} d\phi \int_0^{\infty} d\rho I(\rho) = 2\pi \int_0^{\infty} d\rho I(\rho) = 1.$$

¿Cuál es la densidad de probabilidad marginal $p(\rho)$ asociada a ρ? ¿Cuál es la densidad $\Theta(\theta)$ del ángulo de dispersión θ? Esto ya lo calcularon en Mecánica Clásica sin hablar de probabilidades.



8. (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.

9. (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 $\delta(s - x_1 - \ldots - x_N)$.

10. Sea X una variable aleatoria continua 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 $X = F^{-1}(R)$.

11. 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 la densidad para la variable Y definida como $Y = F(X)$.

12. Este problema combina los dos anteriores. Sean X e Y dos variables aleatorias continuas con distribuciones acumulativas F_X y F_Y, respectivamente. Encuentre la función f tal que $Y = f(X)$.

13. El objetivo de este problema es aplicar la transformación que relaciona dos densidades de probabilidad para disminuir las variaciones accidentales o indeseables entre cuadros vecinos de un film.

Cada punto de una imagen puede asociarse con una variable aleatoria. El ejemplo más simple lo ofrecen las imágenes monocromáticas, pues en tal caso la variable aleatoria asociada a cada punto puede tomar valores reales en un intervalo acotado. Más aún, si las imágenes se almacenan digitalmente estos valores pertenecen a un conjunto discreto. Por ejemplo, para una imagen de 8 bits en escala de grises es convencional asignar a cada punto un valor entero entre 0 y 255, donde cero representa el negro y 255 el blanco. Hay $2^8 = 256$ valores posibles. Para trabajar siempre en el mismo intervalo, independientemente del número de valores permitidos, es más cómodo asignar a cada punto de la imagen un valor entre 0 y 1. En el caso de las imágenes de 8 bits en escala de grises los valores posibles son de la forma $n/255$ con n entero entre 0 y 255.

En una primera aproximación, desentendiéndonos de la posición en que aparecen los píxeles, una imagen rectangular de w píxeles de ancho y h de altura puede considerarse como una muestra de $w \times h$ realizaciones de una misma variable aleatoria. Por lo tanto, si w y h son suficientemente grandes la imagen en cuestión permite una estimación experimental de la distribución de probabilidad, del mismo modo en que $w \times h$ lanzamientos de un dado permitirían una estimación de la probabilidad de cada uno de los 6 resultados posibles.

Cuadros vecinos en un film pueden tener asociadas distribuciones de probabilidad diferentes. Las diferencias pueden corresponder a una modificación en la escena (un cambio de luz, el movimiento de un objeto, el desplazamiento de la cámara), pero también pueden ser producto de variaciones no intencionales, por ejemplo en el tiempo o las aperturas de exposición de cada cuadro, tanto durante la filmación como durante el proceso posterior de revelado y transferencia. Es común observar este tipo de artefactos en los films antiguos, pero también en filmaciones hechas con cámaras digitales con funciones de autoajuste. La figura muestra 7 cuadros consecutivos de la película "Sherlock, Jr.", de 1924. El quinto cuadro es el que más se diferencia de sus vecinos, sin que la escena se haya modificado apreciablemente.

La idea para mitigar este tipo de defectos consiste en asumir que cuadros vecinos de un film, con poca variación intrínseca de la escena, corresponden en principio a la misma distribución de probabilidad, y que la diferencia observada se debe a la aplicación de una misma función f a cada punto de la imagen. Luego, es necesario encontrar esta función y aplicar la transformación inversa a los puntos de la imagen que se quiere corregir. Tomaremos como ejemplos los cuadros 1 y 5 de la figura anterior:

(Se adjuntan los archivos.) La primera es la imagen de referencia. La segunda es la que se quiere corregir, para eliminar el cambio de luminosidad entre una y otra imagen.

(a) Encuentre los histogramas de las dos imágenes adjuntas. Esto da una estimación de las distribuciones de probabilidad de cada imagen. Son distribuciones discretas (cada punto toma uno de 256 valores posibles), así que no es posible aplicar aún los resultados del problema 12.

(b) Encuentre las funciones acumulativas $F_1(x)$ y $F_2(x)$ de cada imagen. Estas son funciones escalonadas, con saltos a intervalos de $1/255$.

(c) Aproxime, de la manera más sencilla que se le ocurra, las funciones acumulativas escalonadas F_1 y F_2 mediante un par de funciones $\tilde{F}_1$ y $\tilde{F}_2$ que sean 1 a 1.

(d) Encuentre la función f que relaciona las variables aleatorias asociadas a $\tilde{F}_1$ y $\tilde{F}_2$. Aplique la transformación correspondiente a la segunda imagen. Compare el resultado con las imágenes iniciales. Encuentre el histograma de la imagen corregida y compárelo con los otros dos.

14. **Direcciones aleatorias** (Feller § I.10.) En $\Re^3$ la dirección de un vector unitario $\mathbf{\Omega}$ puede darse mediante los ángulos esféricos φ y θ, 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.

(a) ¿Cuál es la densidad de probabilidad p en la variables θ y φ que corresponde a una distribución isótropa de versores aleatorios? ¿Seguro?

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

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

(d) 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 $\Re^2$.

(e) Para hacer en la computadora: como chequeo de lo anterior y como aplicación del problema 10, 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 φ y θ, es decir, genere valores al azar de φ y de θ 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 $n \sim 10^4$ y compruebe cualitativamente que la distribución es uniforme en $\overline{-1, 1}$.

Referencias:

Feller, *An introduction to probability theory and its applications*, vol II.
Jaynes, G. L. Bretthorst, *Probability theory. The logic of science*.
Pécseli, *Fluctuations in Physical Systems*.
Reichl, *A modern course in statistical physics*.
Van Kampen, *Stochastic Processes in Physics and Chemistry*.