Estadística en Física Experimental (1er cuatrimestre de 2015)

Guía de Problemas N° 3 | Variables aleatorias contínuas – Cambio de variables

Distribución Normal, Exponencial, Cauchy

1. La altura de la población masculina tiene distribución normal N(175 cm, 8 cm). Sabiendo que los colectivos tienen una altura de 1.95 m, ¿qué fracción puede viajar parada sin necesidad de reclinarse? [Rta: 0.9938]

2. Se mide una magnitud A con un detector bien calibrado cuya resolución es gaussiana con ancho σ, obteniéndose el resultado x.
   (a) ¿Qué rango $x \pm a$ debe informarse si se desea que haya 90% de probabilidad de que éste incluya el verdadero valor de A? [Rta: $x \pm 1.645\sigma$]
   (b) Si el detector tuviera una respuesta no gaussiana y desconocida, pero se sabe que tiene una resolución σ, ¿qué rango puede informar en este caso? [Rta: $x \pm 3.162\sigma$]

3. *Distribución exponencial.* Una fuente radiactiva de constante λ (decaimientos por unidad de tiempo) se ubica frente a un detector, de modo que en promedio, el número de partículas que no decayeron es $N(t) = N_0 \exp(-\lambda t)$
   (a) Muestre que X, el tiempo transcurrido hasta la detección de la primera partícula, tiene distribución exponencial $f_X(t) = \lambda e^{-\lambda t}$. Proceda de dos maneras alternativas, discutiendo la validez de cada paso:
      i. Obtenga primero $F_X(t) = P(X \leq t)$ como el complemento del suceso "el primer decaimiento ocurre a tiempo mayor que t", y luego por derivación, la densidad de probabilidad $f_X(t)$.
      ii. Escriba directamente $f_X(t)$ como el producto de la probabilidad de que no haya ningún decaimiento entre 0 y t, y que haya un decaimiento justo en el dt siguiente.
   (b) Se define que una distribución no tiene memoria si cumple $P(X > t + s|X > t) = P(X > s)$.
      i. ¿Por qué se llama a ésto "no tener memoria"?
      ii. Convénzase que la exponencial es la única distribución de variable contínua que posee esta propiedad (intente usando otras distribuciones). Relaciónelo con las hipótesis que llevan a la distribución de Poisson.

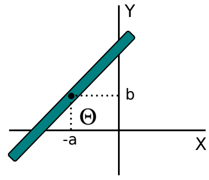
      Nota: para variables aleatorias discretas se define la condición como: $P(X > n+m|X \geq n) = P(X > m)$
   (c) Se tiene una fuente de 1 Bq (Becquerel $\equiv$ 1 decaimiento s^{-1}).
      i. ¿Cuál es la probabilidad de que el primer decaimiento ocurra recién después de 5 segundos? [Rta: 0.0067]
      ii. Calcule la probabilidad de que el primer decaimiento ocurra dentro del primer segundo. Repita el cálculo para el tercer segundo (i.e. el primer decaimiento ocurre entre el 2^{do} y 3^{er} segundo). ¿Contradice ésto la "falta de memoria"? [Rta: 0.6321 y 0.0855]
      iii. ¿Cuánto tiempo debe esperarse para que haya al menos un decaimiento, con una probabilidad del 99%? [Rta: 4.61 s]
   (d) En realidad, el detector cubre un ángulo sólido Ω visto desde la fuente, ¿como se modifica $f_X(t)$ si las partículas son emitidas en todas direcciones con igual probabilidad?
   Nota: relacionarlo con el ejercicio 21 de la guía 2.
   (e) Mencione otros tres ejemplos de procesos que tengan distribución exponencial.

4. Las distribuciones de Cauchy y Gauss tienen ambas forma de campana invertida.
   (a) Usando la computadora dibuje conjuntamente la distribución de Cauchy standard con densidad de probabilidad $f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$, y la $N(0, 0.75)$, normalizada de modo que ambas curvas coincidan en $x = 0$. Note que en la zona central ambas tienen la misma forma pero que difieren notablemente en el comportamiento en las colas ("colas no gausianas").
   (b) **Ejercicio para entregar**. Es común representar colas no gausianas usando la suma de dos gaussianas, $f(x) = aN(0, \sigma_1)+bN(0, \sigma_2)$ (con $a+b = 1$, ¿por qué?). Represente sobre el gráfico anterior, normalizado en $x = 0$ a la de Cauchy, la densidad $f(x)$ para $-6 < x < 6$, con $a = 1/2$, $\sigma_1 = 0.75$ y $\sigma_2 = 3$. Note cómo la suma de dos gausianas ha logrado representar colas marcadamente no gausianas que, para los parámetros y rango elegidos, corresponden con bastante precisión a la forma de la distribución de Cauchy. ¿Qué ocurre si se incrementa el rango del gráfico mucho más allá de $|x| = 6$?

Cambio de variables

5. Si X tiene distribución normal estándar,
   (a) ¿qué distribución tiene $Y = aX + b$?
   (b) ¿cómo implementaría un generador de números al azar N(μ,σ) a partir de uno N(0,1)?

6. Una barra gira alrededor del punto (-a, b), hasta detenerse al azar formando un ángulo Θ con el eje x, ($\Theta \in [-\pi/2, \pi/2]$), como muestra la figura. La recta que contiene a la barra corta al eje de ordenadas en un punto Y. Encuentre la función de distribución $F_\Theta(t)$, y a partir de ésta muestre que Y tiene densidad de probabilidad de Cauchy, $f_Y(t) = \dfrac{1}{\pi}\dfrac{a}{a^2 + (t-b)^2}$. Encuentre su esperanza.
   Notar que la esperanza de una variable con distribución de Cauchy es cero por simetría (su *valor principal* está definido), pero la primitiva de $x/(1+x^2)$ es divergente.



7. Utilice el resultado del problema anterior para generar con la computadora 10000 números al azar que sigan la distribución de Cauchy, a partir de una uniforme [0,1]. Presente los datos en un histograma y grafique sobre éstos la predicción teórica.

8. Muestre que la densidad de probabilidad de $Y{=}X^2$ se obtiene a partir de la de X como $f_Y(t) = [f_X(\sqrt{t}) + f_X(-\sqrt{t})]/2\sqrt{t}$. Halle $f_Y(t)$ si
   (a) X es uniforme en [0,1] [Rta: $f_Y = (2\sqrt{y})^{-1}(0 < t < 1)$]
   (b) X es normal N(0,1) (esta f_Y se denomina χ_1^2, chi^2 con un grado de libertad). [Rta: $f_Y = (1/\sqrt{2\pi t})\exp(-t/2)$, con $t > 0$]

9. La variable aleatoria X tiene distribución uniforme en [0,1].
   (a) Muestre que $Y = e^X$ tiene distribución $f_Y(t) = 1/t, 1 \leq t \leq e$. Note que ésto permite escribir una rutina que genere números con distribución $1/x$.
   (b) **Ejercicio para entregar**. ¿Cómo haría para generar números al azar con distribución exponencial? Impleméntelo en la computadora, construya un histograma con 500 números generados y dibuje sobre éste la distribución teórica. Use $\lambda = 0.25$. Hacer esto para dos 2 histogramas: uno con bines de igual ancho en todo el rango (un ancho apropiado para resolver la forma de la distribución) y otro con bines más gruesos a partir de $Y \geq 15$. ¿Cómo debe escalear la altura de los bines para que se superponga con la distribución?

Valores extremos

10. Se realiza una serie de mediciones *independientes* $\{x_1, \ldots, x_n\}$, de una variable aleatoria X con distribución acumulativa $F_X(x)$ y densidad de probabilidad $F'_X(x) = f_X(x)$.
    (a) Demuestre que la densidad de probabilidad del máximo de todas las mediciones U=max$\{x_1, \ldots, x_n\}$ es $f_{\mathrm{U}}(x) = n f_X(x)\,[F_X(x)]^{n-1}$.
    (b) Probar que la densidad de probabilidad del mínimo de todas las mediciones V=min$\{x_1, \ldots, x_n\}$ es $f_{\mathrm{V}}(x) = n f_X(x)\,[1 - F_X(x)]^{n-1}$.
    (c) Encuentre la expresión formal de la esperanza del rango de una muestra $E(R)$, donde $R = U - V$.

11. Un jugador inexperto juega al pool y quiere dejar la bola blanca lo más cerca posible de la banda opuesta haciendo 2 rebotes, como se muestra en la figura. Si la posición final de la bola blanca en la dirección x, es una variable aleatoria X que tiene distribución uniforme en $[0, L]$, donde L=200 cm es el largo de la mesa. ¿Cuántos tiros debería hacer para que haya alguno en el que la bola quede a menos de 10 cm con probabilidad igual o mayor que p=0.9. [Rta: más de 44]
    Ayuda: hallar primero la distribución acumulativa del valor máximo de la muestra.