

# Estadística en Física Experimental (Verano 2019)

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

### Distribución Normal, Exponencial, Cauchy

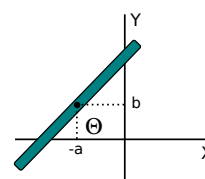
- La altura de la población masculina tiene distribución normal  $N(175 \text{ cm}, 8 \text{ cm})$ . Sabiendo que los colectivos tienen una altura de 1.95 m, ¿qué fracción puede viajar parada sin necesidad de reclinarse? [Rta: 0.9938]
- Se mide una magnitud  $A$  con un detector bien calibrado cuya resolución es gaussiana con ancho  $\sigma$ , obteniéndose el resultado  $x$ .
  - ¿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:  $x \pm 1.645\sigma$ ]
  - Si el detector tuviera una respuesta no gaussiana y desconocida, pero se sabe que tiene una resolución  $\sigma$ , ¿qué rango puede informar en este caso? [Rta:  $x \pm 3.162\sigma$ ]
- Distribución exponencial.** Una fuente radiactiva de actividad  $\lambda$  (decaimientos por unidad de tiempo) se ubica frente a un detector de eficiencia perfecta. Asuma que en cada decaimiento se emite una sola partícula.
  - 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:
    - 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  $f_X(t)$ .
    - 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.
  - Se define que una distribución no tiene memoria si cumple  $P(X > t + s | X > t) = P(X > s)$ .
    - ¿Por qué se llama a esto “no tener memoria”?
    - Convéncase que la exponencial es la única distribución de variable continua 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:  $P(X > n+m | X \geq n) = P(X > m)$

  - Se tiene una fuente de 1 Bq (Becquerel  $\equiv$  1 decaimiento  $s^{-1}$ ).
    - ¿Cuál es la probabilidad de que el primer decaimiento ocurra recién después de 5 segundos? [Rta: 0.0067]
    - 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<sup>do</sup> y 3<sup>er</sup> segundo). ¿Contradice esto la “falta de memoria”? [Rta: 0.6321 y 0.0855]
    - ¿Cuánto tiempo debe esperarse para que haya al menos un decaimiento, con una probabilidad del 99%? [Rta: 4.61 s]
  - Si el detector tuviera perfecta eficiencia intrínseca (mide todo lo que le llega) pero cubre un ángulo sólido  $\Omega$  visto desde la fuente, ¿cómo se modifica  $f_X(t)$  si las partículas son emitidas en todas direcciones con igual probabilidad?
  - Mencione otros tres ejemplos de procesos que tengan distribución exponencial.

### Cambio de variables

- Si  $X$  tiene distribución normal estándar,
  - ¿qué distribución tiene  $Y = aX + b$ ?
  - ¿cómo implementaría un generador de números al azar  $N(\mu, \sigma)$  a partir de uno  $N(0,1)$ ?
- Una barra gira alrededor del punto  $(-a, b)$ , hasta detenerse al azar formando un ángulo  $\Theta$  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,  $f_Y(t) = \frac{1}{\pi} \frac{a}{a^2 + (t - b)^2}$ . 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.

6. 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
- $X$  es uniforme en  $[0,1]$  [Rta:  $f_Y = (2\sqrt{y})^{-1}(0 < t < 1)$ ]
  - $X$  es normal  $N(0,1)$  (esta  $f_Y$  se denomina  $\chi_1^2$ ,  $\chi^2$  con un grado de libertad). [Rta:  $f_Y = (1/\sqrt{2\pi t}) \exp(-t/2)$ , con  $t > 0$ ]
7. Muestre que  $Y = e^X$  (siendo  $X$  una variable aleatoria con distribución uniforme) tiene distribución  $f_Y(t) = 1/t, 1 \leq t \leq e$ . Note que ésto permite escribir una rutina que genere números con distribución  $1/x$ . ¿Cómo haría para generar números al azar con distribución exponencial?
8. Se realiza una serie de mediciones *independientes*  $\{x_1, \dots, x_n\}$ , de una variable aleatoria  $X$  con distribución acumulativa  $F_X(x)$  y densidad de probabilidad  $f'_X(x) = f_X(x)$ .
- Demuestre que la densidad de probabilidad del máximo de todas las mediciones  $U=\max\{x_1, \dots, x_n\}$  es  $f_U(x) = nf_X(x) [F_X(x)]^{n-1}$ .
  - Probar que la densidad de probabilidad del mínimo de todas las mediciones  $V=\min\{x_1, \dots, x_n\}$  es  $f_V(x) = nf_X(x) [1 - F_X(x)]^{n-1}$ .
  - Ciertas lámparas en uso continuo tienen una vida útil con distribución exponencial de 20 días. Una habitación usa para su iluminación tres de estas lámparas. Utilice el resultado de alguno de los ítems anteriores para determinar la distribución del tiempo transcurrido hasta que se deba cambiar una lámpara.
  - Proponga una pregunta que requiera emplear la otra fórmula, es decir, la que no usó para resolver el ítem anterior.

### Problemas computacionales

9. Las distribuciones de Cauchy y Gauss tienen ambas forma de campana invertida.
- Usando la computadora dibuje conjuntamente la distribución de Cauchy estándar con densidad de probabilidad  $f(x) = \frac{1}{\pi} \frac{1}{1+x^2}$ , 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 gaussianas”).
  - Es común representar colas no gaussianas 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 gaussianas ha logrado representar colas marcadamente no gaussianas 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  $|x| = 6$ ?
10. *Método de Monte Carlo*. Para generar números pseudoaleatorios con una distribución arbitraria  $f(t)$ , en un dominio  $[a, b]$  en el que  $f$  está acotada (siendo  $f_m$  su valor máximo) se puede proceder de la siguiente manera:
- Se generan dos números al azar  $y$  y  $z$  con distribución uniforme en  $[0, 1]$ .
  - A partir de  $y$  y  $z$  se determinan  $u = a + (b - a)y$  y  $v = f_m z$ .
  - Si  $v \leq f(u)$ , se incluye  $x = u$  en la muestra de números generados, en otro caso se descarta.
  - Se repiten los pasos I-III hasta obtener la cantidad deseada de números aceptados.
- Utilizando el resultado del problema 5, genere 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.
  - Repita el ítem anterior pero ahora usando el Método Monte Carlo (conocido como aceptación-rechazo).
  - Discuta los pros y contras de cada método.