Tarea # 2

Jonathan Quirola Vásquez
Estadística para Astrónomos

1. **Velocidades superlumínicas de quásares**

    (a) Un quásar emite un jet de partículas con velocidad $\vec{v}$ con un ángulo de inclinación θ respecto a la línea de observación. Además, emite dos fotones, uno en el punto A en un tiempo t_1 y otro en el punto B en un tiempo t_2. Un observador se encuentra en O y recibe los fotones en los tiempos t_1' y t_2' respectivamente. La diferencia en el tiempo de emisión de los fotones es Δt, mientras que la diferencia en el tiempo de arribo al observador es $\Delta t'$, ambos se definen como:

$$\Delta t = t_2 - t_1 \quad \Delta t' = t_2' - t_1'.$$

En la Figura 1 se observa el esquema que se va a considerar. Las partículas recorren una distancia AB con una velocidad v durante un tiempo Δt, por lo tanto:

$$AB = v\Delta t.$$

Por otro lado, se puede definir las distancias AC y BC como:

$$AC = v\Delta t\cos\theta \;\text{ y }\; BC = v\Delta t\sin\theta.$$

El tiempo de arribo del primer fotón al observador es:

$$t_1' = t_1 + \frac{D_L + AC}{c} = t_1 + \frac{D_L + v\Delta t\cos\theta}{c},$$

mientras que el tiempo de arribo del segundo fotón al observador será

$$t_2' = t_2 + \frac{D_L}{c},$$

entonces,

$$\begin{aligned}\Delta t' &= t_2' - t_1' = t_2 - t_1 + \frac{D_L}{c} - \frac{D_L + v\Delta t\cos\theta}{c},\\ \Delta t' &= \Delta t\left(1 - \frac{v}{c}\cos\theta\right).\end{aligned}$$

Por lo tanto, la diferencia en el tiempo de arribo en el sistema de referencia del quásar será

$$\Delta t = \frac{\Delta t'}{1 - \frac{v}{c}\cos\theta}.$$

Si suponemos que los quásares se encuentran a grandes distancias talque $\phi \to 0$, entonces tenemos que

$$BC = D_L\sin\phi \approx D_L\phi = v\Delta t\sin\theta.$$

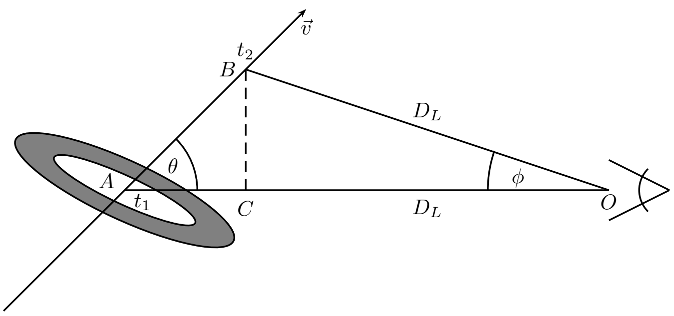

Figure 1: Esquema de la emisión del quásar.

Por lo tanto,

$$\phi D_L = v\sin\theta \frac{\Delta t'}{1-\dfrac{v}{c}\cos\theta},$$

podemos escribir la expresión anterior como:

$$\frac{\phi D_L}{\Delta t'} = \frac{v\sin\theta}{1-\dfrac{v}{c}\cos\theta}.$$

La relación $\phi/\Delta t'$ es la velocidad angular ω del quásar respecto al observador en O. La velocidad transversal es $v' = D_L\omega$, por lo tanto la velocidad transversal aparente del quásar será:

$$v' = \frac{v\sin\theta}{1-\dfrac{v}{c}\cos\theta}.\blacksquare$$

(b) Si consideramos que los quásares están orientados de forma aleatoria entonces θ va entre 0 y $\pi/2$, sin embargo no se distribuyen uniformemente, sino que se distribuyen de acuerdo a su ángulo sólido del quásar, es decir,

$$f(\theta) = \sin\theta \text{ con } 0 < \theta < \pi/2.$$

Si realizamos los siguientes cambios de variables: $\beta = v/c$ y $\beta' = v'/c$, tendremos que el resultado anterior se puede escribir como

$$\beta' = \frac{\beta\sin\theta}{1-\beta\cos\theta}.$$

Como las variables β' y θ' se distribuyen de manera similar, entonces podemos suponer que:

$$f(\theta)d\theta = f(\beta')d\beta' \quad \rightarrow \quad f(\beta') = f(\theta)\left|\frac{d\theta}{d\beta'}\right|,$$

donde tendremos que

$$\frac{d\beta'}{d\theta} = \frac{\left|\beta\cos\theta - \beta^2\right|}{(1-\beta\cos\theta)^2},$$

por lo tanto la distribución de β' tendría la forma

$$f(\beta') = \sin\theta\frac{(1-\beta\cos\theta)^2}{\left|\beta\cos\theta - \beta^2\right|}.$$

Suponiendo que el quásar emite partículas a velocidades cercanas a las de la luz $\beta \rightarrow 1$, la distribución de β' sería

$$f(\beta') = \sin\theta(1-\cos\theta).$$

Considerando la expresión de la velocidad transversal aparente se puede demostrar que (la demostración es larga y tiene algunos pasos poco obvios):

$$1-\cos\theta = \frac{2}{1+\beta'^2} \quad \text{y} \quad \sin\theta = \frac{2\beta'}{1+\beta'^2},$$

reemplazando en la distribución de β' vamos a tener que:

$$f(\beta') = \frac{4\beta'}{(1+\beta'^2)^2}.$$

Para obtener la probabilidad de observar un quásar superlumínico, es decir $\beta' > c$, tenemos

$$\mathbb{P}(\beta' > 1) = \int_1^\infty f(\beta') = \int_1^\infty \frac{4\beta'}{(1+\beta'^2)^2}d\beta.$$

Realizando el cambio de variable $u = 1+\beta'^2$, tendremos que la probabilidad sera:

$$\mathbb{P}(\beta' > 1) = \int_2^\infty \frac{2du}{u^2} = -\left.\frac{2}{u}\right|_2^\infty = 1.$$

Por lo tanto, la probabilidad de observar un quásar con velocidad superlumínicas dado que el jet emitido es cercano a la velocidad de la luz es 1.$\blacksquare$

2. **Generando variables aleatorias de una distribución normal multivariada**

(a) Sea X una variable aleatoria que se distribuye como F, talque $X \sim F$. Si Y es una variable aleatoria es definida como $Y = F(X)$, entonces la CDF de Y será

$$F_Y(y) = p(Y \leq y) = p(F(X) \leq y),$$

entonces el conjunto $A_y = \{x : F(X) \leq y\}$ de y satisface

$$F_Y(y) = p(Y \leq y) = p(X \leq F^{-1}(y)),$$

donde esto es por definición la CDF de X (osea F) en $F^{-1}(y)$,

$$F_Y(y) = p(Y \leq y) = p(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y.$$

Por lo tanto la PDF de Y será:

$$f_Y(y) = F'_Y(y) = 1,$$

este resultado nos quiere decir que Y se distribuye uniformemente, es decir

$$Y \sim \text{Uniforme}(0,1).\blacksquare$$

(b) Suponga que $U \sim \text{Uniforme}(0,1)$ y $X = F^{-1}(U)$, por lo tanto la CDF de X se define como:

$$F_X(X) = p(X \leq x) = p(F^{-1}(U) \leq x) = p(U \leq F(x)),$$

como U se distribuye uniformemente, esto quiere decir que

$$p(U \leq u) = \int_0^u 1du = u,$$

es decir, la CDF de X será

$$F_X(X) = p(X \leq x) = p(F^{-1}(U) \leq x) = p(U \leq F(x)) = F(x).$$

Por lo tanto, X se distribuye como F

$$X \sim F.$$

El resultado anterior se puede interpretar como: Si se considera una distribución cualquiera F y definimos su función inversa $X = F^{-1}$, y además evaluamos F^{-1} con números aleatorios U con distribución uniforme entre 0 y 1 ($X = F^{-1}(U)$), entonces los números aleatorios generados X van a distribuirse como F, es decir,

$$X \sim F.$$

Para probar la idea anterior, primeramente se generó 10000 números aleatorios entre 0 y 1 distribuidos uniformemente, esto se puede observar en la Figura 2 *izquierda*. Después, se aplicó los números aleatorios generados a la función inversa de la CDF de la distribución normal estándar, la cual tiene la siguiente forma:

$$F^{-1}(z) = \sqrt{2}\text{erf}^{-1}(2z - 1),$$

donde z es el número generado aleatoriamente distribuido uniformemente y erf^{-1} es la función error inversa. La ecuación anterior también se denomina función probit y representa la función inversa de la CDF de la distribución normal estándar.

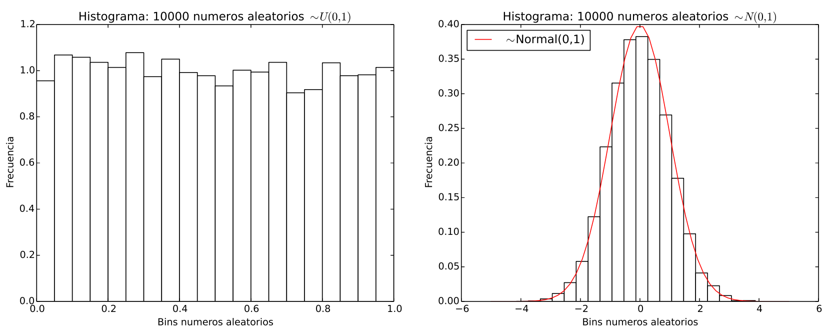

Figure 2: *Izquierda:* Histograma de números aleatorios entre 0 y 1 ([0, 1)) distribuidos uniformemente. *Derecha:* Histograma de números aleatorios obtenidos a partir de números aleatorios distribuidos uniformemente entre 0 y 1 evaluados en la función inversa de la distribución normal estándar, y distribución de ajuste normal estándar (*curva roja*). Ambas gráficas están normalizadas.

El código generado para obtener nos números aleatorios con ambos distribuciones se lo encuentra en el siguiente link: Github.

En la Figura 2 se observa el resultado de generar números aleatorios distribuidos uniformemente (*izquierda*) y al aplicar la inversa de la CDF de la distribución normal estándar. Se aprecia que los números uniformes cambien su distribución al aplicar el resultados obtenido en este literal. Por lo tanto, se puede generar números aleatorios con cualquier tipo de distribución solamente conociendo su CDF inversa.■

(c) Si Σ es una matriz simétrica definida positiva que tiene una descomposición de Cholesky:

$$\Sigma = LL^T,$$

donde L es una matriz triangular inferior. Se define al vector $\vec{S} \sim N(\vec{0}, I)$, donde $\vec{0}$ es un vector de ceros, y I es la matriz identidad, además sea $\vec{X} = \vec{\mu} + L\vec{S}$. Es importante aclarar que las dimensiones de $\vec{X}$ y $\vec{S}$ son las mismas, es decir k.
Podemos definir la transformación inversa de $\vec{X}$ como:

$$\vec{S} = L^{-1}\left(\vec{X} - \vec{\mu}\right),$$

además, tenemos que definir el jacobiano para transformar del espacio de $\vec{S}$ al espacio de $\vec{X}$, el cual va a depender L como:

$$\det(L^{-1}) = \frac{1}{\det(L)},$$

donde $\det(L^{-1})$ y $\det(L)$ son las determinantes de L^{-1} y L respectivamente. Como $\vec{S}$ se distribuye como $\sim \mathrm{N}(\vec{0}, I)$, es decir

$$f_k(\vec{S}) = \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{1}{2}\vec{S}^T\vec{S}\right\}.$$

Por lo tanto, debemos evaluar $\vec{S} = L^{-1}\left(\vec{X} - \vec{\mu}\right)$ en la expresión anterior, donde tendremos

$$f_k(\vec{X}) = \frac{1}{\det(L)} f_k\left[L^{-1}\left(\vec{X} - \vec{\mu}\right)\right]$$

es decir,

$$f_k(\vec{X}) = \frac{1}{(2\pi)^{k/2}\det(L)} \exp\left\{-\frac{1}{2}\left[\left(L^{-1}(\vec{X} - \vec{\mu})\right)^T \cdot \left(L^{-1}(\vec{X} - \vec{\mu})\right)\right]\right\}.$$

Ya que $\Sigma = LL^T$ entonces $\det(\Sigma) = \det(LL^T) = \det(L)\det(L^T) = |\det(L)|^2$ entonces tendremos

$$|\det(L)| = |\det(\Sigma)|^{1/2}.$$

Por otro lado, tenemos que

$$\begin{aligned}
\left(L^{-1}(\vec{X} - \vec{\mu})\right)^T \cdot \left(L^{-1}(\vec{X} - \vec{\mu})\right) &= \left(\vec{X} - \vec{\mu}\right)^T L^{-1T}L^{-1}\left(\vec{X} - \vec{\mu}\right) \\
&= \left(\vec{X} - \vec{\mu}\right)^T L^{T-1}L^{-1}\left(\vec{X} - \vec{\mu}\right), \\
&= \left(\vec{X} - \vec{\mu}\right)^T (L^TL)^{-1}\left(\vec{X} - \vec{\mu}\right)
\end{aligned}$$

como $\Sigma = LL^T$ entonces tendremos que

$$\left(L^{-1}(\vec{X} - \vec{\mu})\right)^T \cdot \left(L^{-1}(\vec{X} - \vec{\mu})\right) = \left(\vec{X} - \vec{\mu}\right)^T \Sigma^{-1}\left(\vec{X} - \vec{\mu}\right).$$

Por lo tanto, la distribución de $\vec{X}$ será:

$$f_k(\vec{X}) = \frac{1}{(2\pi)^{k/2}|\det(\Sigma)|^{1/2}} \exp\left\{-\frac{1}{2}\left(\vec{X} - \vec{\mu}\right)^T \Sigma^{-1}\left(\vec{X} - \vec{\mu}\right)\right\}.$$

El resultado anterior tiene la forma de una normal multivariada, por lo tanto $\vec{X}$ se distribuye como una distribución normal multivariada con media $\vec{\mu}$ y matriz de covarianza Σ.■

(d) Se realizó un código en python para obtener la distribución normal estándar considerando la siguiente matriz de covarianza:

$$\Sigma_{i,j} = \exp\left(-\frac{1}{2}\left|\vec{x}_i - \vec{x}_j\right|^2\right)$$

donde $\vec{x} = [-5, -4.9, \ldots, 4.9, 5]^T$, por lo tanto la dimensión será 100. Se utilizó el hecho que la matriz Σ al poder ser descompuesta con el método de Cholesky ($\Sigma = LL^T$) y ser multiplicada por un vector que se distribuye como una normal multivariada ($\vec{S} \sim N(\vec{0}, I)$), entonces $\vec{X} = \vec{\mu} + \vec{S}L$ estará distribuida como una normal multivariada $\vec{X} \sim N(\vec{\mu}, \Sigma)$.
Los resultados obtenidos se encuentran en la Figura 3, tanto la distribución multivariada obtenida con la descomposición de Cholesky (*izquierda*) así como con la función de `numpy` (*derecha*).

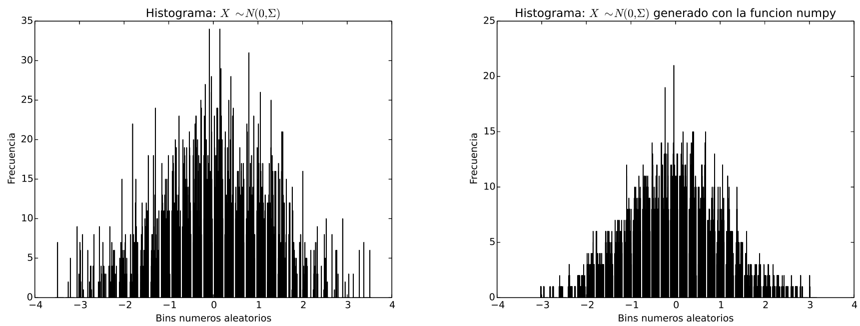

Figure 3: *Izquierda:* Histograma de números aleatorios con una distribuidos multivariada normal considerando el resultado del numeral anterior (descomposición de Cholesky) con dimensión 100. *Derecha:* Histograma de números aleatorios con una distribución multivariada normal considerando la función de numpy `multivariate_normal` con dimensión 100.

En el siguiente link se encuentra el código en `python` de la distribución normal multivariada: Github. Por otro lado, las gráficas de las distribuciones multivariadas estan en: Github.■

3. **Funciones generadoras de momentos**

(a) La función generadora de momentos se define como:

$$\psi_X(t) = \mathbb{E}(e^{tX}) = \int e^{tx} dF(x).$$

Para una variable aleatorio X que se distribuye como una chi-cuadrado con ν grados de libertad, entonces la función generadora de momentos serán:

$$\psi_X(t) = \int f(x)e^{tx}dx = \int_0^\infty e^{tx}\frac{1}{\Gamma(\nu/2)2^{\nu/2}}x^{\frac{\nu}{2}-1}e^{-\frac{x}{2}}dx = \frac{1}{\Gamma(\nu/2)2^{\nu/2}}\int_0^\infty x^{\frac{\nu}{2}-1}e^{-x(\frac{1}{2}-t)}dx.$$

Realizando el cambio de variable $y = x(\frac{1}{2} - t)$ y $dy = (\frac{1}{2} - t)dx$ tenemos que

$$\psi_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}}\int_0^\infty \frac{y^{\frac{\nu}{2}-1}}{\left(\frac{1}{2}-t\right)^{\frac{\nu}{2}-1}}e^{-y}\frac{dy}{(\frac{1}{2}-t)} = \frac{1}{\Gamma(\nu/2)2^{\nu/2}}\frac{1}{\left(\frac{1}{2}-t\right)^{\nu/2}}\int_0^\infty y^{\frac{\nu}{2}-1}e^{-y}dy,$$

donde la integral anterior es la función gamma evaluada en $\nu/2$, es decir,

$$\psi_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}}\frac{1}{\left(\frac{1}{2}-t\right)^{\nu/2}}\Gamma(\nu/2) = \frac{1}{2^{\nu/2}}\frac{1}{\left(\frac{1}{2}-t\right)^{\nu/2}}.$$

Por lo tanto, la función generadora de momentos de la variable aleatorio X distribuida como chi-cuadrada será:

$$\psi_X(t) = \frac{1}{2^{\nu/2}\left(\frac{1}{2}-t\right)^{\nu/2}}.■$$

(b) Si la variable aleatoria Y_1 se distribuye como chi-cuadrada con ν_1 grados de libertad, y la variable aleatoria Y_2 se distribuye como una chi-cuadrado con ν_2 grados de libertad, es decir,

$$\begin{array}{l} Y_1 \sim \chi^2_{\nu_1} \\ Y_2 \sim \chi_{\nu_2} \end{array},$$

y además se define $W = Y_1 + Y_2$, entonces la función generadora de momentos de W sería:

$$\psi_W(t) = \psi_{Y_1}(t)\psi_{Y_2}(t) = \frac{1}{2^{\nu_1/2}\left(\frac{1}{2}-t\right)^{\nu_1/2}} \frac{1}{2^{\nu_2/2}\left(\frac{1}{2}-t\right)^{\nu_2/2}},$$

entonces,

$$\psi_W(t) = \frac{1}{2^{(\nu_1+\nu_2)/2}\left(\frac{1}{2}-t\right)^{(\nu_1+\nu_2)/2}} = \frac{1}{2^{\nu/2}\left(\frac{1}{2}-t\right)^{\nu/2}},$$

donde $\nu = \nu_1 + \nu_2$. Por lo tanto, W se distribuye como una chi-cuadrado con $\nu_1 + \nu_2$ grados de libertad:

$$W \sim \chi^2_\nu \blacksquare$$

(c) Si $Z_i \sim$ N(0,1) (distribución normal estándar). Sea $X = Z_i^2$, entonces la distribución de X se define como:

$$p(X \leq x) = p(Z_i^2 \leq x) = p(-\sqrt{x} \leq Z_i \leq \sqrt{x}),$$

como $Z_i \sim$ N(0,1), tenemos

$$p(-\sqrt{x} \leq Z_i \leq \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$$

donde Φ es la CDF de la distribución normal estándar, entonces

$$\begin{array}{l} p(-\sqrt{x} \leq Z_i \leq \sqrt{x}) = \Phi(\sqrt{x}) - (1 - \Phi(\sqrt{x})) \\ p(-\sqrt{x} \leq Z_i \leq \sqrt{x}) = 2\Phi(\sqrt{x}) - 1 \end{array}.$$

La expresión anterior es la CDF de X, es decir,

$$X \sim 2\Phi(\sqrt{x}) - 1.$$

La PDF se define como $f(x) = F'(x)$, entonces la PDF de X será

$$f(x) = 2\Phi'(\sqrt{x})\frac{1}{2}x^{-1/2} = \phi(\sqrt{x})x^{-1/2},$$

donde $\phi(\sqrt{x})$ es la PDF de la distribución normal estándar evaluada en $\sqrt{x}$. Evaluando $\sqrt{x}$ en la PDF de la normal estándar, tendremos que la PDF de X es:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x} x^{-1/2}.$$

La función generadora de momentos de X será entonces:

$$\psi_X(t) = \mathbb{E}(e^{tx}) = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x} x^{-1/2} e^{tx} dx.$$

Si consideramos el siguiente cambio de variable: $z = x\left(\frac{1}{2}-t\right)$ y $dz = dx\left(\frac{1}{2}-t\right)$, entonces

$$\psi_X(t) = \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{z^{-1/2}}{\left(\frac{1}{2}-t\right)^{1/2}} e^{-z} dz = \frac{1}{\sqrt{2\pi}\left(\frac{1}{2}-t\right)^{1/2}} \int_0^\infty z^{\frac{1}{2}-1} e^{-z} dz,$$

donde la última integral es $\Gamma(1/2)$. Como $\Gamma(1/2) = \sqrt{\pi}$, entonces

$$\psi_X(t) = \frac{\sqrt{\pi}}{\sqrt{2\pi}\left(\frac{1}{2}-t\right)^{1/2}} = \frac{1}{2^{1/2}\left(\frac{1}{2}-t\right)^{1/2}}.$$

La expresión anterior coincide con la función generadora de momentos para una distribución chi-cuadrado con 1 grado de libertad. Por lo tanto $X = Z_i^2$ se distribuye como una chi-cuadrado con 1 grado de libertad

$$Z_i^2 \sim \chi^2_1.$$

Sea $Y = \sum_{i=1}^n Z_i^2$, entonces la función generadora de momentos de Y será:

$$\psi_Y(t) = \psi_{Z_1^2}(t)\psi_{Z_2^2}(t)\ldots\psi_{Z_n^2}(t),$$

como $Z_i^2 \sim \chi^2_1$ entonces,

$$\psi_Y(t) = \underbrace{\frac{1}{2^{1/2}\left(\frac{1}{2}-t\right)^{1/2}}}_{Z_1^2}\underbrace{\frac{1}{2^{1/2}\left(\frac{1}{2}-t\right)^{1/2}}}_{Z_2^2}\cdots\underbrace{\frac{1}{2^{1/2}\left(\frac{1}{2}-t\right)^{1/2}}}_{Z_n^2} = \frac{1}{2^{n/2}\left(\frac{1}{2}-t\right)^{n/2}}.$$

La expresión anterior es la función generadora de momentos de la distribución χ^2 para n grados de libertad. Por lo tanto, Y se distribuye como una chi-cuadrado de n grados de libertad

$$Y \sim \chi_n^2. \blacksquare$$