Tarea 2

Rodrigo Carvajal Pizarro (rcarvaja@astro.puc.cl)
https://github.com/racarvajal/Tarea2_Astro_Stats.git

Resolución

Problema 1

a) Utilizando el diagrama del enunciado del problema, se pueden obtener expresiones para las distancias existentes entre el observador y el quasar y el material eyectado.

Si el material viaja a una velocidad v y fue eyectado en un tiempo t_0, en un tiempo t_1 se encontrará a una distancia $r = vt$ del quasar.

En ese mismo tiempo, la luz que sale directamente del quasar habrá recorrido una distancia $x_1 - x_0 = vt\cos\theta = v(t_1 - t_0)\cos\theta$ y la distancia entre esa luz y un jet emitido en el mismo momento será $vt\sin\theta = v(t_1 - t_0)\sin\theta$.

Ahora, se requiere conocer el tiempo que le tomará al fotón del quasar y al del jet llegar hasta el observador. Para simplificar los cálculos, se puede asumir que el ángulo θ es pequeño. Sea la distancia entre x_1 y el observador D.

El fotón del quasar tomará un tiempo D/c en llegar hasta el observador desde x_1, lo mismo que el fotón que proviene del jet. De este modo, el fotón del quasar demorará $\frac{vt\cos\theta+D}{c} = \frac{v(t_1-t_0)\cos\theta+D}{c}$ en llegar al observador. Por su parte, el fotón del jet demorará $t + D/c = (t_1 - t_0) + D/c$

Durante la expulsión del jet, el observador solo podrá inferir una velocidad transversal en que el material se habrá desplazado $vt\sin\theta = v(t_1 - t_0)\sin\theta$. Con estos valores, se puede inferir esta velocidad.

$$v^{'} = \frac{v(t_1 - t_0)\sin\theta}{(t_1 - t_0) + D/c - \frac{v(t_1-t_0)\cos\theta+D}{c}}$$

$$v^{'} = \frac{v(t_1 - t_0)\sin\theta}{(t_1 - t_0) - \frac{v(t_1-t_0)\cos\theta}{c}}$$

$$v^{'} = \frac{v\sin\theta}{1 - \frac{v}{c}\cos\theta}$$

Esta velocidad corresponde al valor con que el observador verá que se desplaza el jet.

b) Lo primero que debe considerarse para la resolución del problema es la distribución de ángulos de orientación. Se indica que están distribuidos aleatoriamente. Esto, y sin la incorporación de información extra, corresponde a una distribución uniforme. Pero, como se trata de ángulos, debe tomarse esta distribución con cuidado. Para ello, se calculará la distribución.

Se considerará, en primer lugar, una distribución constante en dos tres dimensiones (ángulo sólido). Esta distribución debe ser normalizada:

$$\int_{\Omega} p(\Omega)d\Omega = 1$$

$$\int_0^{\pi/2}\int_0^{2\pi} p(\Omega)d\Omega = 1$$

$$\int_0^{\pi/2}\int_0^{2\pi} p(\Omega)\sin\theta d\theta d\phi = 1$$

$p(\Omega)$ es constante, por lo que puede salir de la integral.

$$2\pi p(\Omega) = 1$$

Entonces,

$$p(\Omega) = \frac{1}{2\pi}$$

Ahora, se marginalizará esta distribución para obtener aquella relacionada solo al ángulo que interesa, θ.

$$p(\theta) = \int_0^{2\pi} p(\theta, \phi) d\phi$$

Por otro lado, se tiene que

$$p(\Omega)d\Omega = p(\theta, \phi)d\theta d\phi$$
$$\frac{1}{2\pi}\sin\theta d\theta d\phi = p(\theta, \phi)d\theta d\phi$$

Entonces,

$$p(\theta, \phi) = \frac{1}{2\pi}\sin\theta$$

Volviendo a la marginalización,

$$p(\theta) = \int_0^{2\pi} p(\theta, \phi)d\phi = \int_0^{2\pi} \frac{1}{2\pi}\sin\theta d\phi$$
$$p(\theta) = \sin\theta$$

Ahora, ya se cuenta con una distribución uniforme para los ángulos en que están orientados los quasares. Los ángulos a considerar para los siguientes cálculos, $0 < \theta < \pi/2$, tienen que ver con la simetría en que se puede ver el jet. Como se expulsa material en dos direcciones opuestas, no es necesario utilizar toda el rango de ángulos pues el efecto será visible o con uno o con el otro jet. También, vale la pena recordar que, para derivar las expresiones de la sección anterior, se consideró que el ángulo era pequeño.

También, es importante considerar que las variaciones en las probabilidades (y sus variables) para la velocidad transversal y el ángulo son pequeñas. Con esto, se tiene que:

$$p(v^{'})dv^{'} = p(\theta)d\theta$$
$$p(v^{'}) = p(\theta)\left|\frac{d\theta}{dv^{'}}\right|$$
$$p(v^{'}) = p(\theta) \times 1/\left|\frac{dv^{'}}{d\theta}\right|$$

Entonces, se puede realizar un cambio de variable en la probabilidad calculando la derivada de la velocidad transversal con respecto al ángulo θ.

$$\left|\frac{dv^{'}}{d\theta}\right| = \frac{|v(\cos\theta - v/c)|}{(1 - v/c\cos\theta)^2}$$

Así,

$$p(v^{'}) = \sin\theta \frac{(1 - v/c\cos\theta)^2}{|v(\cos\theta - v/c)|}$$

También, cambian los límites en que esta distribución es válida. Con el ángulo, correspondía a $0 < \theta < \pi/2$, por lo que con la componente transversal de la velocidad será:

$$0 < v^{'} < v^2/c$$

Ahora, se necesita la distribución de probabilidad en el límite en que la velocidad del jet se acerca a la velocidad de la luz, $v/c \to 1$. Reemplazando y tomando el rango deseado para θ, se tiene:

$$v^{'} = \frac{v\sin\theta}{1 - \cos\theta}$$

$$p(v^{'}) = \sin\theta \frac{(1 - \cos\theta)^2}{|v(\cos\theta - 1)|} = \frac{\sin\theta}{v}(1 - \cos\theta)^2$$

Elevando la primera de las dos últimas ecuaciones al cuadrado y reemplazando $\sin^2\theta = (1 - \cos\theta)(1 + \cos\theta)$, se llega a:

$$\cos\theta = \frac{1 - (v/v^{'})^2}{1 + (v/v^{'})^2}$$

$$1 - \cos\theta = \frac{2}{1 + (v/v^{'})^2}$$

Tomando esta última expresión y reemplazándola, junto con la primera que se manipuló, en la probabilidad para la velocidad transversal, se ve que:

$$p(v^{'}) = \frac{\sin\theta}{v}(1 - \cos\theta)^2 = \frac{v^{'}}{v^2}\frac{4}{1 + (v/v^{'})^2} \qquad \text{para} \quad v/c \to 1$$

Ahora, se necesita la probabilidad para velocidades transversales mayores a la velocidad de la luz ($v^{'} > c$).

$$\mathbb{P}(v^{'} > c) = 1 - \mathbb{P}(v^{'} < c) = 1 - \int_0^c p(v^{'})dv^{'}$$

$$\mathbb{P}(v^{'} > c) = 1 - 2\log{(v^2 + v^{'2})}|_0^c$$

$$\mathbb{P}(v^{'} > c) = 1 - 2[\log{(v^2 + c^2)} - \log{(v^2)}] = 1 - 2\log\left(\frac{v^2 + c^2}{v^2}\right)$$

$$\mathbb{P}(v^{'} > c) = 1 - 2\log{(1 - (c/v)^2)}$$

En que $\log{(\cdot)}$ corresponde al logaritmo natural.

Problema 2

1. Para conocer la PDF de Y, se tiene:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(F(X) \leq y)$$
$$F_Y(y) = \mathbb{P}(X \leq F^{-1}(y)) = F_X(F^{-1}(y)) = y$$

2. Para demostrar que ambas variables aleatorias tienen la misma CDF, se puede considerar que,

$$F_X(u) = \mathbb{P}(X \leq u) = \mathbb{P}(F^{-1}(U) \leq u)$$
$$F_X(u) = \mathbb{P}(U \leq F(u)) = F(u)$$

Si se necesitan números generados aleatoriamente con cualquier distribución y solo se tienen números que siguen una distribución uniforme, se puede utilizar lo demostrado anteriormente. Esto, llevado a pasos a seguir, significa que, primero, debe calcularse la inversa de la distribución deseada ($F^{-1}(X)$) y evaluar esta función con los números generados uniformemente.

Un ejemplo de este procedimiento puede hacerse al obtener números que distribuyan como una Normal Estándar a partir de valores generados uniformemente ($X \sim \text{Unif}(0,1)$).

La CDF de la distribución Normal Estándar es:

$$F(x) = \frac{1}{2}\left[1 + \text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

Con $\mu = 0$ y $\sigma = 1$. Ahora lo que debe hacerse es invertir esta función, con lo que se llega a:

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

Así, cada número que sea generado de una distribución uniforme, al ser evaluado en esta función inversa, corresponderá a una distribución Normal Estándar. Para corroborar esta idea, se realizó una simulación con un millón de números generados desde una distribución uniforme. Estos valores se ven en el histograma de la Figura 1.

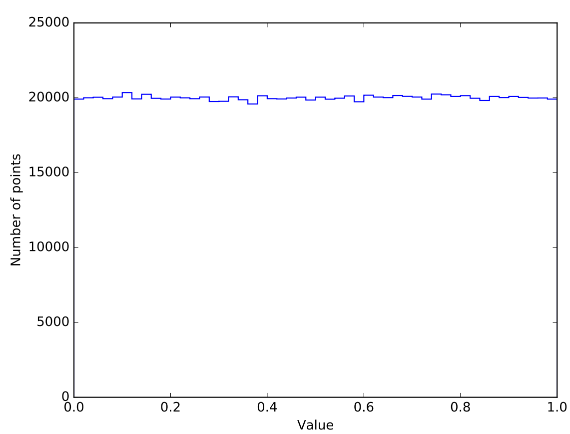

Figura 1: Histograma para un millón de números generados uniformemente

Si se grafican los valores una vez evaluados en la CDF inversa ya señalada, se verán como en la Figura 2. Con una simple inspección, puede notarse que, ahora, los puntos siguen una distribución estándar; están centrados en 0 y su desviación estándar es cercana a 1.

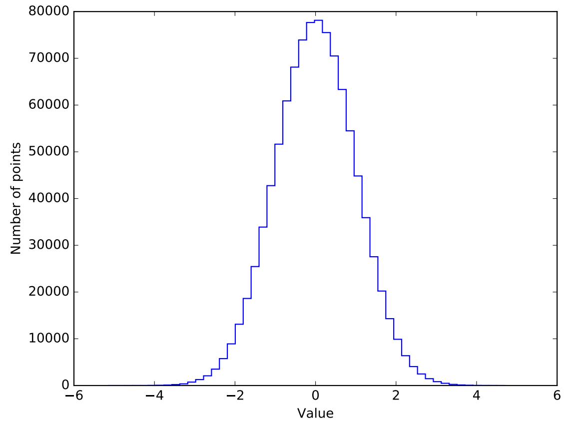

Figura 2: Histograma para un millón de números en Distribución Estándar

Con esto, se comprueba que el método para obtener números que tengan una distribución cualquiera a partir de de números que distribuyen uniformemente sí funciona.

3. Considerando que la matriz **L** es triangular inferior, la expresión para $\vec{X}$ se puede representar como:

$$\vec{X} = \vec{\mu} + \mathbf{L}\vec{S} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix} + \begin{pmatrix} 0 & 0 & \dots & 0 \\ L_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \dots & 0 \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{pmatrix}$$

$$\vec{X} = \vec{\mu} + \mathbf{L}\vec{S} = \begin{pmatrix} \mu_1 \\ \mu_2 + L_{21}S_2 \\ \vdots \\ \mu_n + \sum_{i=1}^{n} L_{ni}S_i \end{pmatrix}$$

Como puede notarse, cada uno de los términos del vector $\vec{X}$ corresponde a una combinación lineal de elementos de $\vec{S}$.

Por otro lado, se conoce que una combinación lineal de variables aleatorias normales será, también, una variable aleatoria normal con parámetros:

$$\mu = \sum_{i=1}^{k} \mu_i L_{ki}$$

y

$$\sigma^2 = \sum_{i=1}^{k} \sigma_i^2 L_{ki}^2$$

Así, y cuando se tiene un vector cuyos elementos distribuyen como una Normal, este vector distribuirá como una Normal Multivariada. Sus parámetros serán:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}$$

y

$$\sigma^2 = \Sigma$$

Este último parámetro se obtiene al conocer que, por sus propiedades, la matriz $\mathbf{L}$ cumplirá que:

$$\mathbf{L}^2 = \mathbf{L} \times \mathbf{L}^{-1}$$

Por lo tanto,

$$\vec{X} = \vec{\mu} + \mathbf{L}\vec{S} \sim \mathcal{N}(\vec{\mu}, \mathbf{\Sigma})$$

4. La sección anterior indica que el vector que se desea (Normal multivariado $\vec{X}$ con media $\vec{\mu} = \vec{0}$ y matriz de covarianza Σ) puede escribirse como:

$$\vec{X} = \vec{\mu} + \mathbf{L}\vec{S}$$

En que $\mathbf{L}$ corresponde a la matriz inferior de la descomposición de Cholesky para la matriz de covarianza.

Al programar el llenado de la matriz de covarianza y su descomposión, surge un problema importante. Σ es una matriz no definida positiva, por lo que no puede descomponerse en dos matrices triangulares.

Para solucionar este problema, se intenta 'parchar' la matriz sumándole en su diagonal un valor positivo que sea suficiente para hacer que todos los valores propios de la matriz se vuelvan positivos. En este caso y por inspección, el valor utilizado es $\epsilon = 2{,}843557 \cdot 10^{-15}$. Además, este valor tiene que ser pequeño para no alterar los resultados finales de forma significativa.

Una vez solucionado este inconveniente, se obtiene un vector de valores distribuidos como una normal multivariada estándar ($\mathcal{N}(\vec{0}, \mathbf{1})$). Los valores pueden verse en el histograma de la Figura 3.

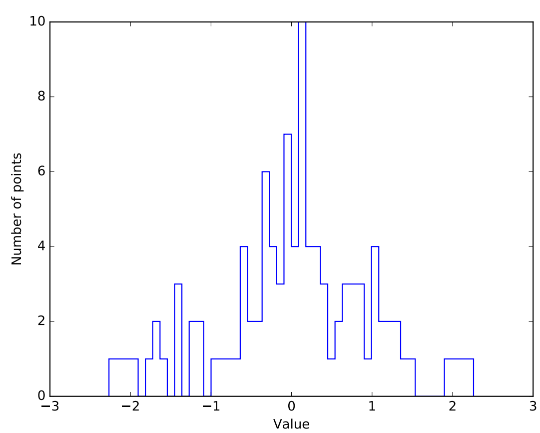

Figura 3: Histograma para valores distribuidos como una Normal Multivariada Estándar

Ahora, estos valores pueden ser llevados a la distribución deseada y mencionada anteriormente. para ello se evalúan en $\vec{X} = \vec{\mu} + \mathbf{L}\vec{S}$. El resultado de este proceso puede verse en el histograma de la Figura 4 en donde puede comprobarse que los datos se encuentran centrados en $X = 0$ tal como se requería.

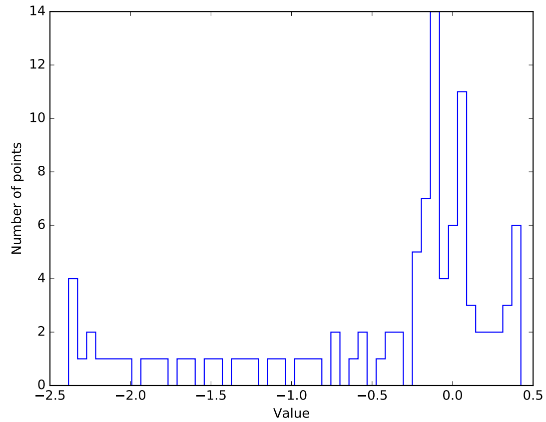

Figura 4: Histograma para valores distribuidos como una Normal Multivariada centrada en cero y con matriz de covarianza Σ

Problema 3

a) En general, la Función Generadora de Momentos, $M_X(t)$, se puede obtener a través de la siguiente expresión:

$$M_X(t) = \int_{-\infty}^{\infty} \exp{(tx)} f(x) dx$$

En el caso de la distribución Chi-cuadrado, los valores de x serán no negativos.

$$M_X(t) = \int_{0}^{\infty} \exp{(tx)} \frac{1}{\Gamma(\nu/2)2^{\nu/2}} x^{\nu/2-1} \exp{(-x/2)} dx$$

$$M_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}} \int_{0}^{\infty} \exp{(x(t-1/2))} x^{\nu/2-1} dx$$

Para resolver la integral, se puede hacer un cambio de variables: $\frac{du}{(t-1/2)} = dx$. Ante este cambio, además, deben reconsiderarse los límites de la integral. Para ello, lo primero que debe notarse es que hay una restricción para el valor de t. El cálculo de la Función Generadora de Momentos implica que no debe haber restricciones para t en todo su rango. Como hay problemas para $t = 1/2$, se debe acotar su intervalo. Quedará como $t < 1/2$.

Con los valores posibles para t, pueden obtenerse los nuevos límites para la integral. Y quedarán:

$$M_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}} \int_{0}^{-\infty} \exp{(u)} \left(\frac{u}{t-1/2}\right)^{\nu/2-1} \frac{du}{t-1/2}$$

$$M_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}(t-1/2)^{\nu/2}} \int_{0}^{-\infty} \exp{(u)} u^{\nu/2-1} du$$

Haciendo un nuevo cambio de variables, $dw = -du$, se tiene,

$$M_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}(t-1/2)^{\nu/2}} \int_{0}^{\infty} \exp{(-w)} (-w)^{\nu/2-1} (-dw)$$

Como x y u son no negativas, cualquier variación de w será negativa. Por lo tanto, el signo negativo que acompaña a la variable en la integral anterior hará que se vuelvan no negativas.

$$M_X(t) = \frac{1}{\Gamma(\nu/2)2^{\nu/2}(t-1/2)^{\nu/2}} \int_0^\infty \exp(-w) w^{\nu/2-1} dw$$

De este modo, puede reconocerse que la integral corresponde a la función Gamma, $\Gamma(\nu/2)$. Este valor se cancela con la función Gamma ya presente y se tiene:

$$M_X(t) = \frac{1}{2^{\nu/2}(t-1/2)^{\nu/2}} = \frac{1}{(2t-1)^{\nu/2}}, \qquad t < 1/2, x > 0$$

b) La Distribución Chi-cuadrado puede representarse aquella que tiene la suma del cuadrado de ν variables aleatorias que, por sí solas, tienen una distribución normal estándar.

$$X = \sum_{i=1}^{\nu} Z_i^2 \sim \chi^2(\nu)$$

Por lo tanto, si se tienen dos variables, Y_1 y Y_2, con una distribución Chi-cuadrado con ν_1 y ν_2 grados de libertad respectivamente, la suma de estas variables tendrá el siguiente comportamiento:

$$W = Y_1 + Y_2 = \sum_{i=1}^{\nu_1} Z_i^2 + \sum_{i=1}^{\nu_2} Z_i^2$$

$$W = Y_1 + Y_2 = \sum_{i=1}^{\nu_1+\nu_2} Z_i^2 \sim \chi^2(\nu_1 + \nu_2)$$

Así, la suma de dos variables aleatorias que tienen una distribución Chi-cuadrado también tendrá una distribución Chi-cuadrado con $\nu_1 + \nu_2$ grados de libertad.

c) Si se tiene que $Y = Z_i^2$ y utilizando la distribución de probabilidad de Z_i, se llega a:

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(Z^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq Z \leq \sqrt{y})$$

Y,

$$F_Y(y) = F_Z(-\sqrt{y}) - F_Z(\sqrt{y})$$

Por lo tanto,

$$f_Y(y) = \frac{1}{2} y^{-1/2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} + \frac{1}{2} y^{-1/2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y} = \frac{1}{2^{1/2}\sqrt{\pi}} y^{-1/2} e^{-\frac{1}{2}y}$$

$$f_Y(y) = \frac{y^{-1/2} e^{-y/2}}{2^{1/2}\Gamma(\frac{1}{2})}$$

Esta función corresponde a la distribución de probabilidad Chi-cuadrado con un grado de libertad.
Para $Y = \sum_{i=1}^{n} Z_i^2$, el desarrollo será similar al obtenido en la sección anterior (b).