Tarea 3.

Fecha de entrega 2 de abril

1. Simule lo siguiente:

   a) $\{(X_i, Y_i)\}_{i=1}^n$ una muestra aleatoria de normales bivariadas con coeficiente de correlación 0.0 ($\rho = 0$), de medias 0 y varianza $\mathbf{I}$ (la identidad en dos dimensiones).

   b) $\{(X_i, Y_i)\}_{i=1}^n$ una muestra aleatoria de normales bivariadas con coeficiente de correlación 0.5 ($\rho = 0.5$), $\mu = (1,1)^T$ y varianza 1.

   c) De un ejemplos de v.a. bivariadas con marginales normales con media 0, varianza 1 y su simulación.

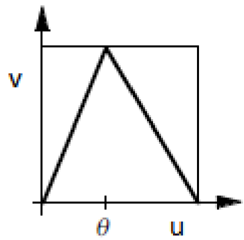
2. Simule una muestra de la siguiente cópula y utilice el teorema de Sklar para dar la simulación de una v.a. con marginales $\mathcal{N}(0,1)$ y $Gama(1.5, 3)$ para $\theta = 2$.

$$C(u, v) = \max\left(1 + \frac{\theta}{\log(e^{\theta/(u-1)} + e^{\theta/(v-1)})}, 0\right)$$

3. Sea $\theta \in [0,1]$, y suponga que la masa de probabilidad θ se distribuye uniformemente en el segmento de linea que va de (0,0) a (θ,1) y la masa de probabilidad $(1-\theta)$ se distribuye uniforme en el segmento de linea de (θ,1) a (1,0).

   - Encuentre la función de distribución del vector aleatorio y exhiba que es una cópula.
   - Utilice el teorema de Sklar para dar la simulación de una v.a. con marginales $\mathcal{N}(0,1)$ y $Gama(1.5, 3)$ que tenga asociada la cópula del punto anterior.

   Sugerencias: La función de distribución está definida por partes. El siguiente dibujo de los valores que toma la variable aleatoria puede ser útil.



   analice cuanto vale $C(u, v)$ para cada una de las tres regiones.

4. Respecto al algoritmo de Marsaglia, surge un planteamiento intuitivo. Ya que el algoritmos nos pide generar U_1 y U_2 cada vez que proponemos un punto, considere el caso de que rechacemos (U_1, U_2) ¿ Es más eficiente si en vez de simular una nueva pareja (U_3, U_4) simulamos solamente una nueva uniforme y proponemoso el punto (U_2, U_3) y si lo rechazamos proponer entonces (U_3, U_4)? La idea que inspira esta pregunta es que estamos ahorrando el tiempo de la simulación de una uniforme en cada iteración, que como pudo constatar en ejercicios anteriores puede ser un proceso tardado.

   La respuesta a esta pregunta puede ser una demostración o una comparación de la eficiencia de los algoritmos de Marsaglia y la modificación planteada.

5. La función de densidad y de distribución de la v.a. Laplace asimétrica están dadas por

$$f(x; m, \lambda, \kappa) = \frac{\lambda}{\kappa + 1/\kappa} \begin{cases} \exp\left((\lambda/\kappa)(x-m)\right) & \text{if } x < m \\ \exp(-\lambda\kappa(x-m)) & \text{if } x \geq m \end{cases}$$

$$F(x; m, \lambda, \kappa) = \begin{cases} \frac{\kappa^2}{1+\kappa^2} \exp((\lambda/\kappa)(x-m)) & \text{if } x \leq m \\ 1 - \frac{1}{1+\kappa^2} \exp(-\lambda\kappa(x-m)) & \text{if } x > m \end{cases}$$

Para $\mu = 0, \lambda = 2, \kappa = 1.5$ realice lo siguiente.

- Estime la mediana. Sugerencia utilice cuantiles empíricos
- Estime la media
- Estime la media utilizando variables control usando el siguiente resultado: La expresión de la mediana está dada por

$$q(0.5) = m + (\kappa/\lambda) \log\left(\frac{1+\kappa^2}{2\kappa^2}\right)$$

6. **Contexto**

Sea $X \sim N_m(\mathbf{0}, \sigma^2 I)$ con σ^2 conocida.

Un problema habitual en estadística consiste en estimar el vector de medias μ, una de las estadísticas más usadas es $\hat{\mu}$ que es el estimador máximo verosimil (EMV), cuya expresión está dada por

$$\hat{\mu} = \frac{1}{n} \sum x_i,$$

donde cada x_i es una muestra de $X \sim N_m(\mathbf{0}, \sigma^2 I)$, este estimador se puede verificar que es consistente e insesgado.

Además de la máxima verosimilitud existen otro tipos de estimadores, entre ellos están los que surgen de la teoría de decisiones; para estos estimadores se requiere una función de pérdida $\mathcal{L}$, que mide el costo de elegir el estimador $\tilde{\mu}$ con respecto a un valor μ que se supone cierto, en el caso de X podemos tomar

$$\mathcal{L}(\tilde{\mu}, \mu) = ||\tilde{\mu} - \mu||_2^2 = (\tilde{\mu} - \mu)^T (\tilde{\mu} - \mu)$$

Es decir nuestra pérdida es la distancia euclideana al cuadrado, de la μ real y de nuestro estimador $\hat{\mu}$ ¿Te parece adecuada esta manera de cuantificar el error?

Los estimadores en teoría de decisiones son los que surgen de encontrar $\hat{\mu}$ que minimiza

$$\mathbb{E}(\mathcal{L}(\tilde{\mu}, \mu))$$

es decir aquellos que te dan la menor pérdida esperada. En el caso de $\mathcal{L}(\tilde{\mu}, \mu) = ||\tilde{\mu} - \mu||_2^2$ estamos buscando el estimador que tenga el menor error cuadrático medio.

En 1956 Stein demostró que el uso del EMV era inadmisible para dimensiones grandes y en 1961 propuso un estimador, Â¡El cuál revolucionó la estimación de la media de la normal!, se le conocec como el estimador de James-Stein y su expresión está dada por

$$\tilde{\mu} = \Big(1 - \frac{(m-2)\frac{\sigma^2}{n}}{||\hat{\mu}||^2}\Big)\hat{\mu}.$$

para $m \geq 2$.

$\tilde{\mu}$ puede ser interpretado como una corrección de $\hat{\mu}$, que depende de la dimensión y de la norma de $\hat{\mu}$.

Si está interesadod en leer más al respecto puede consultar: Efron, B., & Morris, C. (1972). Limiting the risk of Bayes and empirical Bayes estimators-Part II: The empirical Bayes case. Journal of the American Statistical Association, 67(337), 130-139.

Ejercicio

para $m \in \{2, 100, 10000\}$, tomando $X \sim N_m(\mathbf{0}, \sigma^2 I)$. En este ejercicio sientase con total libertad de usar generadores de normales multivariadas en cualquier extensión. Resuelva lo siguiente

- Estime la esperanza de $\tilde{\mu}$ y de $\hat{\mu}$
- Estime $\mathbb{E}(\mathcal{L}(\tilde{\mu}, \mu))$, $\mathbb{E}(\mathcal{L}(\hat{\mu}, \mu))$ y $\mathbb{E}(||\hat{\mu} - \tilde{\mu}||)$, ¿ Qué estimador es mejor y por qué?
- Estime la varianza de sus estimadores y diga que método de muestreo fue el más efectivo

7. Realizar lo siguiente

   - Grafique en un cuadrado $(0,1) \times (0,1)$ los puntos que surgen de particionar ambos ejes usando la secuencia de Van Der Corput para la base 10.
   - Grafique en un cuadrado $(0,1) \times (0,1)$ los puntos que surgen de particionar el eje x usando la secuencia de Van Der Corput para la base 7 y el eje y usando la secuencia de Van Der Corput para la base 11. Con base en este ejercicio responda ¿Por qué cree que se sugiere usar secuencias de Halton para métodos Quasi-Montecarlo?
   - Grafique en un cuadrado $(0,1) \times (0,1)$ los puntos que surgen de usar la secuencia de Hammersley (7,11).