



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES ACATLÁN

MODELACIÓN MULTIVARIADA POR MEDIO DE
CÓPULAS DE DEPENDENCIAS NO LINEALES PARA
LA SIMULACIÓN ESPACIAL CONJUNTA DE
PROPIEDADES GEO-REFERENCIADAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

Lic. en Matemáticas Aplicadas y Computación

PRESENTA:

Jorge Enrique Ingle González

F
E
S
ACATLÁN

ASESOR:

Dr. Arturo Erdely Ruiz

ABRIL 2016

Santa Cruz Acatlán, Naucalpan, Estado de México



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Al Dr. Arturo Erdely Ruiz, a la Mtra. Jeanett López García, a la Profesora Norma Hernández Hernández y al Mtro. Robert Hernández Martínez por aguantarme, compartirme su conocimiento y brindarme su apoyo incondicional.

A mis padres, mis hermanos, mi abuela, mis tías, tíos y primos por todo su amor y siempre estar conmigo aún cuando quería tirar la toalla.

A mis amigos Enrique, Gerardo, Mauricio y Danny Boy que siempre escucharon todas mis burradas cuando nadie más quería.

A la memoria del Prof. José Calvillo García y de mi primo Aristóteles.

Al profesor Gregorio Paniagua por sus consejos.

Al profesor Raúl Gallardo, el físico nuclear más "greñudo".

A Zeus.

Índice general

Índice general	I
Lista de imágenes	III
Índice de tablas	VI
Introducción	1
1. Marco Teórico	3
1.1. Preámbulo	3
1.2. Cópulas bivariadas y propiedades básicas	4
1.3. Teorema de Sklar y vectores aleatorios bivariados	6
1.4. Simulación de vectores aleatorios bivariados	7
1.5. Medidas de dependencia	9
2. Metodología de Cópulas en Racimo	13
2.1. Cópulas Trivariadas	13
2.2. Independencia Condicional	15
2.3. Un modelo teórico	16
2.4. Cópula Trivariada en Racimo	24
2.4.1. Antecedentes	24
2.4.2. Justificación matemática	25
2.4.3. Generalización del modelo teórico	26
2.4.4. Supuesto de simplificación	42

2.4.5. Prueba de Rémillard y Scaillet	51
3. Aplicación a un caso de estudio	64
3.1. Análisis descriptivo	64
3.2. Supuesto simplificador	69
3.3. Construcción del modelo trivariado	72
Conclusiones	78
Apéndice	80
A.1. Código en R	80
A.1.1. Instalación y carga de paquetes	80
A.1.2. Funciones base creadas por el Dr. Arturo Erdely Ruiz	81
A.1.3. Funciones empleadas en el desarrollo de la tesis	84
A.2. Código para generar imágenes y de las tablas del documento	110
A.2.1. Desarrollo de la tesis	131

Lista de imágenes

1.1. Cotas Frechet Hoeffding y Cópula producto	5
2.1. Histogramas contra funciones marginales de densidad	18
2.2. Gráficos 3D de función densidad bivariada	19
2.3. Curvas de nivel a color de función de densidad bivariada	19
2.4. Diagramas de dispersión de observaciones	19
2.5. Diagramas de dispersión de observaciones simuladas	20
2.6. $\varphi(x_2, x_3)$ versus $\ell_1(x_2, x_3)$	21
2.7. $\psi(x_2, x_3) := \mathbb{V}(X_1 X_2 = x_2, X_3 = x_3)$ contra $\hat{\sigma}^2$ del ajuste de regresión lineal múltiple	22
2.8. Prueba Shapiro-Wilk	23
2.9. Histograma de residuales y Normal ajustada	23
2.10. $f_{1 23}(x_1 0.9, 0.8)$ contra Normal $(\hat{\beta}_1 + 0.9\hat{\beta}_2 + 0.8\hat{\beta}_3, \hat{\sigma}^2)$	24
2.11. Histogramas de X_1, X_2 y X_3 con $\theta = \{0.01, 0.49, 0.51\}$	31
2.12. Histogramas de X_1, X_2 y X_3 con $\theta = \{0.99, 1.01, 1.98\}$	32
2.13. Histogramas de X_1, X_2 y X_3 con $\theta = 2.03$	33
2.14. Densidad de función bivariada (X_1, X_3) con $\theta = \{0.01, 0.49, 0.51\}$	34
2.15. Densidad de función bivariada (X_1, X_3) con $\theta = \{0.99, 1.01, 1.98\}$	35
2.16. Densidad de función bivariada (X_1, X_3) con $\theta = 2.03$	36
2.17. Densidad de función bivariada (X_2, X_3) con $\theta = \{0.01, 0.49\}$	36
2.18. Densidad de función bivariada (X_2, X_3) con $\theta = \{0.51, 0.99, 1.01\}$	37
2.19. Densidad de función bivariada (X_2, X_3) con $\theta = \{1.98, 2.03\}$	38

2.20. Comparación entre correlaciones	41
2.21. Comparación entre correlaciones condicionales con $\theta = 0.01$	44
2.22. Comparación entre densidades de cópula con $\theta = 0.01$	44
2.23. Comparación entre correlaciones condicionales con $\theta = 0.49$	45
2.24. Comparación entre densidades de cópula con $\theta = 0.49$	45
2.25. Comparación entre correlaciones condicionales con $\theta = 0.51$	46
2.26. Comparación entre densidades de cópula con $\theta = 0.51$	46
2.27. Comparación entre correlaciones condicionales con $\theta = 0.99$	47
2.28. Comparación entre densidades de cópula con $\theta = 0.99$	47
2.29. Comparación entre correlaciones condicionales con $\theta = 1.01$	48
2.30. Comparación entre densidades de cópula con $\theta = 1.01$	48
2.31. Comparación entre correlaciones condicionales con con $\theta = 1.98$	49
2.32. Comparación entre densidades de cópula con $\theta = 1.98$	49
2.33. Comparación entre correlaciones condicionales con con $\theta = 2.03$	50
2.34. Comparación entre densidades de cópula con $\theta = 2.03$	50
2.35. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.01$.	52
2.36. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.49$.	52
2.37. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.51$.	53
2.38. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.99$.	53
2.39. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 1.01$.	54
2.40. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 1.98$.	54
2.41. Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 2.03$.	55
2.42. Pseudo Observaciones Condicionales divididas para $\theta = 0.01$	56
2.43. Pseudo Observaciones Condicionales divididas para $\theta = 0.49$	57
2.44. Pseudo Observaciones Condicionales divididas para $\theta = 0.51$	58
2.45. Pseudo Observaciones Condicionales divididas para $\theta = 0.99$	59
2.46. Pseudo Observaciones Condicionales divididas para $\theta = 1.01$	60
2.47. Pseudo Observaciones Condicionales divididas para $\theta = 1.98$	61
2.48. Pseudo Observaciones Condicionales divididas para $\theta = 2.03$	62

3.1. Diagramas de caja y brazos	65
3.2. Diagramas de dispersion de la muestra original en 2D y 3D	66
3.3. Histograma de cada variable aleatoria de la muestra original	67
3.4. Pseudo observaciones de la muestra original	67
3.5. Pseudo observaciones de la muestra original	68
3.6. Pseudo observaciones condicionales	69
3.7. Pseudo observaciones condicionales clasificadas en dos grupos	70
3.8. Pseudo-observaciones de simulación	72
3.9. Datos de muestra originales vs Simulación	73
3.10. Pseudo-observaciones condicionales de las observaciones simuladas	74
3.11. Datos de muestra originales vs Simulación	76
3.12. Datos de muestra originales vs Simulación en 3D	77

Índice de tablas

2.I. Resumen de estadística descriptiva	18
2.II. Coeficiente de correlación teórico v.s muestral	20
2.III. Resumen estadística descriptiva (parte I)	29
2.IV. Resumen estadística descriptiva (parte II)	29
2.V. Resumen estadística descriptiva (parte III)	30
2.VI. Resumen estadística descriptiva (parte IV)	30
2.VII. Coeficiente de correlación teórico v.s muestral	41
2.VIII. Prueba de Rémillard y Scalliet, p-values	63
3.I. Resumen de Estadística descriptiva	65
3.II. Comparación entre medidas de ρ de Spearman vs σ de Wolff-Schweizer de las pseudo-observaciones	68
3.III. <i>p-values</i> de pruebas de Rémillard-Scaillet	71
3.IV. Comparación entre medidas de ρ de Spearman vs σ de Wolff-Schweizer de las pseudo-observaciones condicionales	71
3.V. <i>p-values</i> de pruebas de Rémillard-Scaillet de comparación datos originales vs simulación	74

Introducción

El presente trabajo fue realizado como parte del proyecto PAPIIT IN115914 dirigido por el Dr. Arturo Erdely Ruiz.

Objetivo General

El objetivo general de la tesis es conocer si es factible el empleo de la metodología de Cópulas en Racimo para la simulación de datos petrofísicos y conocer la eficiencia del análisis de regresión y de correlación tradicionales.

Resumen de cada capítulo

En el capítulo uno se presenta una breve introducción a la teoría de Cópulas, el algoritmo para simular observaciones a partir de un conjunto de datos bivariados empleando funciones cópula y la definición de medidas de dependencia. También se definen los conceptos de cópula empírica y la cópula Bernstein, las cuáles permiten la aproximación, de forma no paramétrica, de la cópula subyacente de un conjunto de datos.

En el capítulo dos se muestra la metodología de Cópulas en Racimo, la cual permite capturar la estructura de dependencia de un vector, que en este caso pertenece al caso trivariado. Además se propuso un modelo teórico, a partir del cual se simularon observaciones, se realizó el análisis de correlación y el ajuste de regresión lineal. Aparte de todo lo anterior se aplicó a las observaciones simuladas de este modelo la metodología de Cópulas en Racimo. Se definió el concepto de supuesto simplificador, vital para desarrollo de la tesis. Así mismo se propone la prueba de Rémillard y Scaillet (2009), para determinar si n pares de observaciones de las mismas dimensiones poseen una cópula subyacente

similar. Con dicha prueba se demostró que es posible aplicar el supuesto simplificador a pseudo-observaciones condicionales pertenecientes al modelo teórico.

En el capítulo tres se aplicó la metodología de Cópulas en Racimo a un conjunto de datos petrofísicos con el objetivo de simular observaciones y compararlas con otras observaciones simuladas a partir del ajuste de regresión lineal. También se empleó con éxito la prueba de Rémillard y Scaillet (2009) para conocer si los datos muestrales y los simulados poseían una cópula subyacente parecida.

1.1. Preámbulo

El presente preámbulo y las secciones 1.2 y 1.3 están basadas en Nelsen (2006).

Se dará a conocer la noción de función *2-creciente*, la cual es el análogo en dos dimensiones de una función no decreciente de una variable.

Para entender mejor dicho concepto, introduciremos alguna notación.

Sea $\overline{\mathbb{R}}$ la representación del conjunto extendido de los números reales. Un rectángulo en $\overline{\mathbb{R}}^2$ es el producto cartesiano de dos intervalos cerrados:

$$B = [x_1, x_2] \times [y_1, y_2]$$

Siendo los vértices del rectángulo B los puntos $\{(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)\}$.

El cuadrado unitario I^2 es el producto $I \times I$ donde $I = [0, 1]$.

Definición 1.1.1. Sean S_1 y S_2 dos subconjuntos no vacíos de $\overline{\mathbb{R}}$, y sea H una función tal que $Dom H \subseteq \overline{\mathbb{R}}^2$. Sea $B = [x_1, x_2] \times [y_1, y_2]$ un rectángulo cuyos vértices se encuentran en $Dom H$. Entonces el H -volumen de B se encuentra dado por:

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1)$$

La función H es *2-creciente* si $V_H(B) \geq 0$ para todos los rectángulos B cuyos vértices se encuentren en $Dom H$.

1.2. Cópulas bivariadas y propiedades básicas

Definición 1.2.1. Una cópula es una función C de I^2 a I que posee las siguientes propiedades:

Para cada u y v en I :

$$\begin{aligned} C(u, 0) = 0, \quad C(0, v) = 0 \\ C(u, 1) = u, \quad C(1, v) = v \end{aligned} \tag{1.a}$$

En (1.a) se puede observar que si alguno de los argumentos de la función cópula es cero, entonces su valor es 0 y si el valor de alguno de los argumentos de la función cópula es uno, entonces esta funge como función identidad.

Así mismo las expresiones contenidas en (1.a) se consideran como las condiciones de frontera de una función cópula.

Para cada u_1, u_2, v_1, v_2 en I tales que $u_1 \leq u_2$ y $v_1 \leq v_2$:

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0 \tag{1.b}$$

Lo anterior se denomina propiedad *2-creciente* de una función cópula.

Debido a que $C(u, v) = V_c([0, u] \times [0, v])$, uno puede pensar en $C(u, v)$ como una asignación de un número en I al rectángulo $[0, u] \times [0, v]$. Entonces (1.b) nos proporciona una fórmula de “inclusión-exclusión” para C en cada rectángulo $[u_1, u_2] \times [v_1, v_2]$ en I^2 y enuncia que el número a asignarse debe ser no negativo.

Teorema 1.2.2. Sea C una cópula, entonces para cada (u, v) en $\text{Dom } C$:

$$\text{máx}(u + v - 1, 0) \leq C(u, v) \leq \text{mín}(u, v) \tag{1.c}$$

Es posible referirnos también a $\text{máx}(u + v - 1, 0)$ como W o *cota inferior de Fréchet-Hoeffding* y a $\text{mín}(u, v)$ como M o *cota superior de Fréchet-Hoeffding*. Tomando en cuenta esto, la expresión (1.c) se reescribiría de la siguiente forma:

$$W(u, v) \leq C(u, v) \leq M(u, v) \tag{1.d}$$

Cualquier función cópula C está contenida dentro de las cotas de Fréchet-Hoeffding, las cuales también son cópulas.

Otra cópula que encontraremos frecuentemente es la cópula producto $\Pi(u, v) = uv$. A continuación se muestran las gráficas en tres dimensiones y las curvas de nivel de las cotas de Fréchet-Hoeffding y de la función cópula producto:

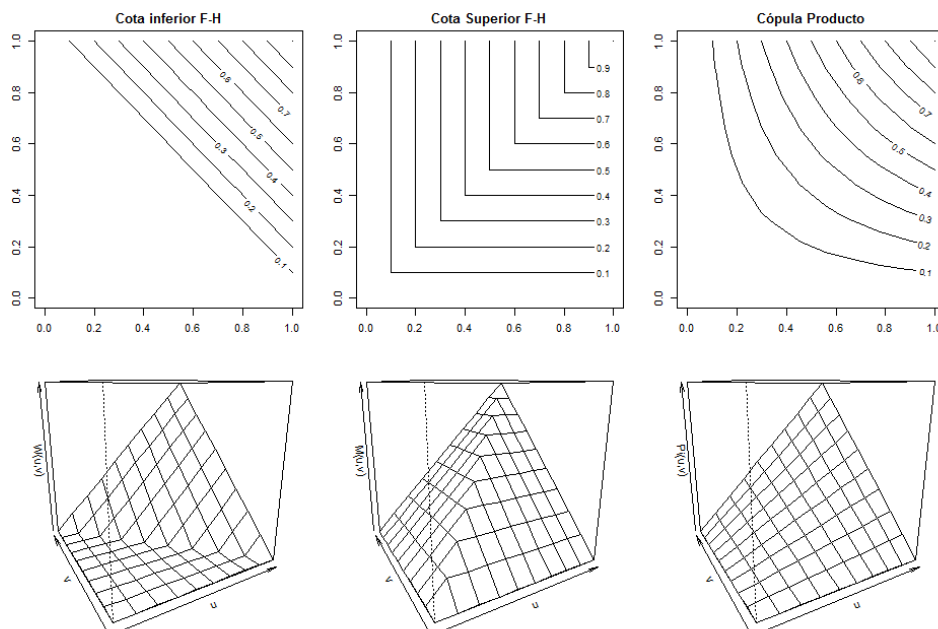


Ilustración 1.1: Cotitas Fréchet Hoeffding y Cópula producto

Teorema 1.2.3. Sea C una función cópula, y sea $\mathbf{a} \in [0, 1]$. Cuando se fija \mathbf{a} como segundo argumento de C se obtiene la sección horizontal de C dada por $t \rightarrow C(t, \mathbf{a})$; así mismo la sección vertical de C se obtiene fijando \mathbf{a} como primer argumento de C , esto es $t \rightarrow C(\mathbf{a}, t)$; y la sección diagonal de C es la función δ_C definida por $\delta_C = C(t, t)$.

Dichas funciones son monótonas crecientes y uniformemente continuas en el intervalo cerrado $[0, 1]$.

Ahora veremos un teorema concerniente a derivadas parciales de una función cópula.

Teorema 1.2.4. Sea C una función cópula. Para cualquier v en $I := [0, 1]$, la derivada parcial $\frac{\partial C(u, v)}{\partial u}$ existe casi para cada u , y para dichos valores de u y v :

$$0 \leq \frac{\partial C(u, v)}{\partial u} \leq 1 \tag{1.e}$$

Similarmente, para cualquier u en $I := [0, 1]$, la derivada parcial $\frac{\partial C(u,v)}{\partial v}$ existe para casi toda v , y para dichos valores de u y v :

$$0 \leq \frac{\partial C(u,v)}{\partial v} \leq 1 \tag{1.f}$$

Las funciones $\frac{\partial C(u,v)}{\partial v}$ y $\frac{\partial C(u,v)}{\partial u}$ están definidas como no decrecientes y están casi en donde sea en I , es decir, excepto posiblemente sobre algún conjunto de medida de Lebesgue cero.

Definición 1.2.5. Sea $c_{12}(u,v)$ la función de densidad de la cópula $C_{12}(u,v)$, entonces:

$$c_{12}(u,v) = \frac{\partial^2}{\partial u \partial v} C_{12}(u,v)$$

1.3. Teorema de Sklar y vectores aleatorios bivariados

De Nelsen (2006):

Teorema 1.3.1. Sean X y Y dos variables aleatorias con funciones de distribución marginal F y G respectivamente y función de distribución conjunta H . Entonces existe una función cópula C , tal que:

$$H(x,y) = C(F(x), G(y)) \tag{1.g}$$

Si X y Y son continuas, entonces C es única, de otra forma, C está determinada de forma única en $\text{Ran } F \times \text{Ran } G$.

Más aún, si C es una función cópula y F y G son funciones de distribución univariadas, entonces la función H definida en (1.g) es una función de distribución conjunta con marginales F y G .

De aquí en adelante emplearemos C_{XY} para denotar la cópula subyacente al vector aleatorio (X, Y) .

Teorema 1.3.2. Sean X y Y variables aleatorias continuas. Entonces X y Y son independientes si y sólo si $C_{XY} = \Pi$.

Teorema 1.3.3. Sean X y Y variables aleatorias continuas con función cópula C_{XY} . Si α y β son funciones estrictamente crecientes en $\text{Ran } X$ y en $\text{Ran } Y$ respectivamente, entonces $C_{\alpha(X)\beta(Y)} = C_{XY}$. Es decir, C_{XY} es invariante bajo transformaciones estrictamente crecientes de X y Y .

Si la función cópula de X y Y es la cota superior de Frechét-Hoeffding (M), entonces Y es una función creciente de X .

Y si la función cópula de X y Y es la cota inferior de Frechét-Hoeffding (W), entonces Y es una función decreciente de X .

1.4. Simulación de vectores aleatorios bivariados

En esta sección abordaremos los conceptos necesarios para poder simular observaciones de un vector aleatorio bivariado (X, Y) .

Se mostrará una función tomada de Nelsen (2006), que será la base del algoritmo para la simulación:

$$\mathbb{P}(V \leq v | U = u) = \left. \frac{\partial C(u, v)}{\partial u} \right|_{u=F(x), v=G(y)} \quad (1.h)$$

Es preciso comentar que si C es la función cópula de (X, Y) , con F y G como sus funciones de distribución marginales, al definir las variables aleatorias $U = F(X)$ y $V = G(Y)$, entonces por teorema (1.3.3) la cópula de (U, V) es la misma que la de (X, Y) y por lo tanto $P(Y \leq y | X = x)$ es igual a (1.h).

Podemos apreciar que en (1.h) la función de distribución condicional de V condicionada en un valor u es igual a la primera derivada parcial de la cópula con respecto a u . Por conveniencia nos referiremos a (1.h) como $c_u(v)$.

Se deben inicializar las variables aleatorias $U \sim Unif(0, 1)$ y $T \sim Unif(0, 1)$ y después para simular observaciones de la variable aleatoria V , basta con obtener el valor de la función $c_u^{-1}(t)$.

Una vez realizado lo anterior se obtendrán pares de pseudo-observaciones de un vector bivariado (U, V) .

Más adelante se mostrará que se deben transformar dichas pseudo-observaciones mediante las cuasi-inversas F^{-1} y G^{-1} para concluir la simulación de (X, Y) .

A continuación se muestran algunas ideas basadas en el trabajo de Erdely y Díaz-Viera (2010):

Las funciones marginales de distribución empírica de un conjunto de observaciones $\{(x_1, y_1), \dots, (x_n, y_n)\}$ se obtienen de la siguiente forma:

$$\begin{aligned} F_n(x) &= \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{x_k \leq x\}} \\ G_n(y) &= \frac{1}{n} \sum_{k=1}^n \mathbf{1}_{\{y_k \leq y\}} \end{aligned} \tag{1.i}$$

Es posible demostrar que las expresiones contenidas en (1.i) son estimadores consistentes de F y G , es decir $F_n(t)$ converge a $F(t)$ cuando $n \rightarrow \infty$ para todo t y el mismo caso aplica para $G_n(t)$ y $G(t)$. También es importante hacer notar que las funciones de distribución empíricas son funciones escalonadas y por ende discontinuas por lo cual no son apropiadas para variables aleatorias continuas sin antes suavizarlas.

Las funciones que nos permitirán transformar los valores de las pseudo-observaciones en una simulación del vector aleatorio bivariado (X, Y) son las siguientes:

$$\begin{aligned} Q(u) &= F^{-1}(u) \text{ para } u := [0, 1] \\ R(v) &= G^{-1}(v) \text{ para } v := [0, 1] \end{aligned} \tag{1.j}$$

Ambas expresiones en (1.j) se refieren a una estimación de la función cuantílica marginal o cuasi-inversa de F y G para valores de u y v respectivamente.

Para relizar la estimación de $Q(u)$ y de $R(v)$ se aplicarán los polinomios de Bernstein-Kantorovic propuestos por Perez y Palacín (1987), citados por Erdely (2009):

$$\begin{aligned} \hat{Q}_n(u) &= \sum_{k=0}^n \frac{x_{(k)} + x_{(k+1)}}{2} \binom{n}{k} u^k (1-u)^{n-k} \\ \hat{R}_n(v) &= \sum_{k=0}^n \frac{y_{(k)} + y_{(k+1)}}{2} \binom{n}{k} v^k (1-v)^{n-k} \end{aligned} \tag{1.k}$$

Donde $x_{(k)}$ y $y_{(k)}$ son los respectivos estadísticos de orden.

Para concluir la sección se mostrará el algoritmo para simular observaciones de un vector bivariado (X, Y) basado en el trabajo de Nelsen (2006) y de Erdely y Díaz-Viera (2010):

1. Inicialice dos variables aleatorias continuas independientes con distribución uniforme u y t .
2. Asigne $v = c_u^{-1}(t)$, donde c_u^{-1} denota la cuasi-inversa de c_u .
3. Para obtener los pares deseados (x, y) aplicamos la función \hat{Q} a todos los valores de u y la función \hat{R} a todos los valores de v respectivamente.

1.5. Medidas de dependencia

Mostraremos ahora una definición propuesta por Schweizer y Wolff (1981), citados por Erdely (2009):

Definición 1.5.1. Sean X, Y cualesquiera variables aleatorias continuas con cópula subyacente C . Una medida numérica de asociación para este par de variables aleatorias es *medida de dependencia*, denotada mediante $\mu_{X,Y}$ o μ_C , si satisface las siguientes propiedades:

1. $\mu_{X,Y}$ está definida para cualquier par de variables aleatorias continuas.
2. $0 \leq \mu_{X,Y} \leq 1$.
3. $\mu_{X,Y} = \mu_{Y,X}$.
4. X, Y son independientes si y sólo si $\mu_{X,Y} = \mu_{\Pi} = 0$.
5. $\mu_{XY} = 1$ si y sólo si X, Y son cada una casi seguramente función estrictamente monótona de la otra; es decir, si C_{XY} es igual a M o W .
6. Si α y β son casi seguramente funciones estrictamente monótonas en $Ran X$ y $Ran Y$, respectivamente, entonces $\mu_{\alpha(X)\beta(Y)} = \mu_{X,Y}$.

7. Si $\{(X_n, Y_n)\}$ es una sucesión de variables aleatorias continuas con cópulas subyacentes C_n , respectivamente, y si $\{C_n\}$ converge puntualmente a C , entonces $\lim_{n \rightarrow \infty} \mu_{C_n} = \mu_C$.

Agregan Schweizer y Wolff (1981), citados por Erdely (2009):

[...] cualquier medida de distancia, adecuadamente estandarizada, entre las superficies $z = C(u, v)$ y $z = uv$, por ejemplo, cualquier distancia L_p , debiera proporcionar una medida simétrica y no paramétrica de dependencia.

Para cualquier $p \in [1, \infty[$ la distancia L_p entre una cópula C y Π está dada por:

$$\mathfrak{U}_C(p) := \left(k_p \int \int_{I^2} |C(u, v) - uv|^p dudv \right)^{\frac{1}{p}} \quad (1.1)$$

en donde k_p es una constante escogida de modo que (1.1) asigne el valor $\mathfrak{U}_C(p) = 1$ cuando C sea igual a M o W .

Erdely (2009) comenta que cuando $\mathfrak{U}_C(1)$ a (1.1) se le denomina como la medida de dependencia de Schweizer y Wolff (1981) dada por:

$$\sigma_C := 12 \int \int_{I^2} |C(u, v) - uv| dudv \quad (1.m)$$

podemos interpretarla como una medida del “valor absoluto promedio” entre la distribución conjunta de un par de variables aleatorias continuas (representada por su cópula C) y la distribución conjunta de un par de variables aleatorias continuas independientes (representada por $\Pi(u, v) = uv$).

A continuación se mostrará una medida de concordancia tomada de Nelsen (2006), popularmente conocida como ρ de Spearman la cual es posible expresar en términos de su cópula subyacente:

$$\rho_{XY} = 12 \int \int_{I^2} [C(u, v) - uv] dudv \quad (1.n)$$

Si existe dependencia positiva o dependencia negativa entre las variables aleatorias X y Y , entonces $|\rho_{X,Y}| = \sigma_C$. Para variables aleatorias donde no exista dependencia positiva ni negativa, $|\rho_{X,Y}| \neq \sigma_C$.

Es importante también mencionar el coeficiente de correlación lineal (o de Pearson):

$$r_{XY} = \frac{Cov(X, Y)}{\sqrt{V(X) V(Y)}} \quad (1.\tilde{n})$$

que de igual manera es expresable en términos de su cópula subyacente:

$$r_{XY} = \frac{1}{\sqrt{V(X) V(Y)}} \int \int_{I^2} [C(u, v) - uv] dF^{-1}(u) dG^{-1}(v) \quad (1.o)$$

Es menester decir que ρ_{XY} (1.n) es mucho mejor que r_{XY} (1.o) porque no depende del comportamiento marginal de las variables involucradas. Sin embargo si $\rho_{XY} = 0$ no implica que las variables aleatorias implicadas sean independientes.

Embrechts, McNeil y Straumann (1999) citados por Erdely (2009) enuncian un conjunto de elementos para considerar al momento de utilizar correlación lineal:

1. La correlación lineal es una medida escalar de dependencia y por lo tanto al utilizarse se pierde información sobre la estructura de dependencia original.
2. Debido a que el valor del coeficiente correlación lineal depende de las distribuciones marginales de las variables aleatorias involucradas, los valores de -1 y 1 no son siempre alcanzables.
3. Variables aleatorias con dependencia positiva perfecta no siempre tienen correlaciones iguales a 1 y variables aleatorias con dependencia negativa perfecta no tienen correlaciones iguales a -1 .
4. La correlación lineal no es invariable bajo transformaciones crecientes de las variables aleatorias. Por ejemplo, $\log X$ y $\log Y$ por lo general no tienen la misma correlación lineal de X y Y , a pesar de que la cópula de (X, Y) es la misma que la de $(\log X, \log Y)$ y que por tanto su estructura de dependencia es equivalente.
5. La correlación sólo está definida cuando las varianzas de las variables aleatorias son finitas por lo que no es una medida de dependencia recomendada para variables aleatorias cuyas varianzas sean infinitas.

No contamos con información previa acerca de la estructura de dependencia a la que están sujetos los datos que se emplearan durante la esta tesis, por eso nos ayudaremos de la función cópula empírica para obtener una estimación de la cópula subyacente.

Definición 1.5.2. La cópula empírica es una función C_n dada por:

$$C_n\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{\left(\text{numero de pares } (x, y) \text{ en la muestra con } x \leq x_{(i)}, y \leq y_{(j)}\right)}{n} \quad (1.p)$$

Donde $x_{(i)}$ y $y_{(i)}$, $1 \leq i, j \leq n$, denotan los respectivos estadísticos de orden.

Erdely y Díaz-Viera (2010) mencionan:

La cópula empírica no es una cópula, desde que solo esta definida en una rejilla finita, no en el cuadrado unitario $[0, 1]^2$ pero por el teorema de Sklar C_n puede ser extendida a una cópula. Dicha extensión es posible por una cópula Bernstein.

Una consecuencia de lo anterior es que la cópula Bernstein *convierte* en continua y *suaviza* la superficie de la cópula empírica. A continuación se mostrará la definición formal de función cópula Bernstein propuesta por Sancetta y Satchell (2004).

Definición 1.5.3. Una cópula Bernstein $\hat{C}(u, v)$ se define de la siguiente manera:

$$\hat{C}(u, v) = \sum_{u=0}^n \sum_{v=0}^n C_n\left(\frac{i}{n}, \frac{j}{n}\right) \binom{n}{i} u^i (1-u)^{n-i} \binom{n}{j} v^j (1-v)^{n-j}. \quad (1.q)$$

donde (u, v) está definido en todo el cuadrado unitario $[0, 1]^2$ y C_n es como en (1.p).

Metodología de Cópulas en Racimo

Las definiciones, los teoremas y las demostraciones presentes en este capítulo están basados en el trabajo de Nelsen (2006), Czado (2011), Bedford y Cooke (2001) y en Joe (1996).

2.1. Cópulas Trivariadas

Definición 2.1.1. Una 3-cópula es una función C que va de I^3 a I que posee las siguientes propiedades:

- Para cada u en I^3 , $C(u) = 0$, si al menos una coordenada de u es 0. Y si todas las coordenadas de u son 1 excepto u_k , entonces $C(u) = u_k$.
- Para cada \mathbf{a} y \mathbf{b} en I^3 , tal que $a_1 \leq b_1, a_2 \leq b_2, a_3 \leq b_3$ se tiene que $V_c([\mathbf{a}, \mathbf{b}]) \geq 0$.

Teorema 2.1.2. Sea H una función de distribución trivariada con marginales F_1, F_2, F_3 . Entonces existe una 3-cópula C tal que para todo $x \in \overline{\mathbb{R}}^3$:

$$H(x_1, x_2, x_3) = C(F_1(x_1), F_2(x_2), F_3(x_3)) \quad (2.a)$$

De forma inversa, si C es una 3-cópula y F_1, F_2, F_3 son funciones de distribución, entonces la función H definida por (2.a) es una función de distribución de tres dimensiones. Si F_1, F_2 y F_3 son todas continuas, C es única. De otra forma, C está determinada únicamente en $\text{Ran}F_1 \times \text{Ran}F_2 \times \text{Ran}F_3$.

Las versiones en tres dimensiones de las 2-cópulas M, Π y W están denotadas como M^3, Π^3 y W^3 , y están dadas por:

$$\begin{aligned} M^3(u) &= \min(u_1, u_2, u_3) \\ \Pi^3(u) &= u_1 \cdot u_2 \cdot u_3 \\ W^3(u) &= \max(u_1 + u_2 + u_3 - 2, 0) \end{aligned}$$

Las funciones M^3 y Π^3 son 3-cópulas mientras que W^3 es lo más cercano posible a una función cópula.

Teorema 2.1.3. Sea C una cópula de tres dimensiones, entonces para cada u en $\text{Dom } C$,

$$W^3(u) \leq C(u) \leq M^3(u) \quad (2.b)$$

Teorema 2.1.4. Sean X_1, X_2, X_3 variables aleatorias continuas, entonces:

1. X_1, X_2 y X_3 son independientes entre si, sí y solo sí la 3-cópula de X_1, X_2 y X_3 es Π^3 .
2. Con probabilidad uno, las variables aleatorias X_1, X_2 y X_3 son funciones estrictamente crecientes unas de las otras sí y solo sí su función cópula es M^3 .

Teorema 2.1.5. Sea C una función cópula trivariada y sea

$$K(u_1, u_2) = C(u_1, u_2, 1) \quad (2.c)$$

Entonces $K(u_1, u_2)$ es función cópula y de hecho una de las marginales bivariadas de C .

Demostración del teorema 2.1.5. Por la definición (2.1.1):

$$\begin{aligned} C(0, v, 1) &= C(u, 0, 1) = 0 \\ C(u, 1, 1) &= u \\ C(1, v, 1) &= v \end{aligned}$$

Y sea \mathbf{u} y \mathbf{v} tal que $u_1 \leq v_1, u_2 \leq v_2, u_3 = v_3 = 1$, entonces:

$$V_C(\mathbf{u} \times \mathbf{v}) = C(u_2, v_2, 1) - C(u_2, v_1, 1) - C(u_1, v_2, 1) + C(u_1, v_1, 1) \geq 0$$

Por lo tanto $K(u, v)$ es función cópula y de hecho una de las marginales bivariadas de C .

Definición 2.1.6. Sea $c_{123}(u, v, w)$ la función de densidad de la cópula $C_{123}(u, v, w)$, entonces:

$$c_{123}(u, v, w) = \frac{\partial^3}{\partial u \partial v \partial w} C_{123}(u, v, w)$$

2.2. Independencia Condicional

Definición 2.2.1. Sea (X_1, X_2, X_3) un vector aleatorio con función de densidad conjunta $f(x_1, x_2, x_3)$ tal que:

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_2, x_3 | x_1) \cdot f(x_1) \\ &= f(x_2 | x_1) \cdot f(x_3 | x_1) \cdot f(x_1) \end{aligned} \quad (2.d)$$

Se dice entonces que variables aleatorias X_2 y X_3 son condicionalmente independientes en X_1 .

Es posible reexpresar (2.d) de dos formas diferentes:

$$\begin{aligned} f(x_1, x_2, x_3) &= f(x_1 | x_2) \cdot f(x_3 | x_2) \cdot f(x_2) \\ f(x_1, x_2, x_3) &= f(x_1 | x_3) \cdot f(x_2 | x_3) \cdot f(x_3) \end{aligned} \quad (2.e)$$

Las expresiones en (2.e) son válidas siempre y cuando se posea independencia condicional en X_2 o en X_3 , respectivamente.

Se puede escribir la función de densidad conjunta de un vector trivariado en términos de la densidad de su función cópula.

$$(X_1, X_2, X_3) \sim f_{123}(x_1, x_2, x_3) = c_{123}(F_1(x_1), F_2(x_2), F_3(x_3)) f_1(x_1) f_2(x_2) f_3(x_3) \quad (2.f)$$

Si las variables aleatorias X_2 y X_3 son condicionalmente independientes en X_1 , entonces es posible expresar la función de densidad conjunta utilizando funciones de densidad cópula bivariadas como se demostrará.

Observación 2.2.2. Sea:

$$\begin{aligned} f_{123}(x_1, x_2, x_3) &= f_{23|1}(x_2, x_3 | x_1) \cdot f_1(x_1) \\ &= f_{2|1}(x_2 | x_1) \cdot f_{3|1}(x_3 | x_1) \cdot f_1(x_1) \end{aligned}$$

Y:

$$\begin{aligned} f_{2|1}(x_2 | x_1) &= \frac{f_{12}(x_1, x_2)}{f_1(x_1)} \\ &= \frac{c_{12}(F_1(x_1), F_2(x_2))}{f_1(x_1)} \cdot f_1(x_1) \cdot f_2(x_2) \\ &= c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2) \end{aligned}$$

Así mismo:

$$\begin{aligned} f_{3|1}(x_3|x_1) &= \frac{f_{13}(x_1, x_3)}{f_1(x_1)} \\ &= \frac{c_{13}(F_1(x_1), F_3(x_3))}{f_1(x_1)} \cdot f_1(x_1) \cdot f_3(x_3) \\ &= c_{13}(F_1(x_1), F_3(x_3)) \cdot f_3(x_3) \end{aligned}$$

Por lo tanto:

$$f_{123}(x_1, x_2, x_3) = c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \quad (2.g)$$

Si las variables X_1 y X_3 son condicionalmente independientes en X_2 , la expresión (2.g) tendría la siguiente forma:

$$f_{123}(x_1, x_2, x_3) = c_{12}(F_1(x_1), F_2(x_2)) c_{23}(F_2(x_2), F_3(x_3)) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \quad (2.h)$$

Finalmente, si se da el caso en que X_1 y X_2 son condicionalmente independientes en X_3 :

$$f_{123}(x_1, x_2, x_3) = c_{13}(F_1(x_1), F_3(x_3)) c_{23}(F_2(x_2), F_3(x_3)) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \quad (2.i)$$

2.3. Un modelo teórico

El siguiente modelo busca ejemplificar lo expresado en (2.g) y también se probará si este modelo funciona bajo los supuestos de linealidad y normalidad comúnmente aplicados.

Sea $X_1 \sim \text{Unif}(0, 1)$ $X_2|X_1 = x_1 \sim \text{Unif}(0, x_1)$ $X_3|X_1 = x_1 \sim \text{Unif}(0, x_1)$

Primero se calcularán algunas funciones de densidad:

$$f(x_1) = \mathbf{1}_{\{0 < x_1 < 1\}}$$

$$f(x_2) = (-\log x_2) \mathbf{1}_{\{0 < x_2 < 1\}}$$

$$f(x_3) = \begin{cases} (-\log x_3) \mathbf{1}_{\{0 < x_3 < 1\}} & \text{si } \theta = 0 \\ \left[\log(x_3) + \frac{2}{\sqrt{x_3}} - 2 \right] \mathbf{1}_{\{0 < x_3 < 1\}} & \text{si } \theta = 1 \\ \left[\frac{1-x_3^{\frac{-\theta}{\theta+1}}}{\theta^2-\theta} + \frac{\theta x_3^{-1+\frac{1}{\theta}}}{(-\theta+1)} (1-x_3^{\frac{-1}{\theta^2+\theta}}) \right] \mathbf{1}_{\{0 < x_3 < 1\}} & \text{si } 0 < \theta \neq 1 \end{cases} \quad (2.j)$$

$$f(x_1, x_2) = \frac{1}{x_1} \mathbf{1}_{\{0 < x_2 < x_1 < 1\}} \quad (2.k)$$

$$f(x_1, x_3) = \frac{1}{x_1} \mathbf{1}_{\{0 < x_3 < x_1\}} \mathbf{1}_{\{0 < x_1 < 1\}} \quad (2.l)$$

$$f(x_2, x_3) = \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < x_2^\theta\}} \frac{1}{x_2^\theta} \left[\left(-1 + \frac{1}{x_2} \right) \mathbf{1}_{\{x_3 < x_2^{\theta+1}\}} + \left(-1 + \frac{x_2^\theta}{x_3} \right) \mathbf{1}_{\{x_3 \geq x_2^{\theta+1}\}} \right] \quad (2.m)$$

$$f_{1|23}(x_1|x_2x_3) = \begin{cases} \frac{x_3}{x_1^2(x_2^\theta-x_3)} \mathbf{1}_{\{x_3 < x_2^\theta < x_1 < 1\}} & \text{si } x_3 \geq x_2^{1+\theta} \\ \frac{x_2}{x_1^2(1-x_2)} \mathbf{1}_{\{x_2 < x_1 < 1\}} & \text{si } x_3 < x_2^{1+\theta} \end{cases} \quad (2.n)$$

Se realizará la simulación de 1000 observaciones del vector aleatorio (X_1, X_2, X_3) mediante el método condicional:

Para $k = 1, 2, \dots, 1000$ {

Simula $x_1[k]$ a partir de una $\text{Unif}(0, 1)$

Simula $x_2[k]$ a partir de una $\text{Unif}(0, x_1[k])$

Simula $x_3[k]$ a partir de una $\text{Unif}(0, x_1[k])$

Siguiente k }

Se muestra un resumen de estadística descriptiva acerca las observaciones simuladas.

Tabla 2.I: Resumen de estadística descriptiva

Estimador	X_1	X_2	X_3
Mín	0.0000234	0.0000085	0.000015
Primer cuartil	0.2250767	0.0698964	0.069565
Mediana	0.4981592	0.1978755	0.193401
Media	0.5013312	0.2575942	0.255022
Tercer cuartil	0.7494541	0.3935962	0.391099
Máx	0.9998597	0.9798222	0.995049

En la siguiente imagen se muestra una matriz de gráficos de 1×3 con el histograma y la función de densidad univariada de cada variable aleatoria simulada.

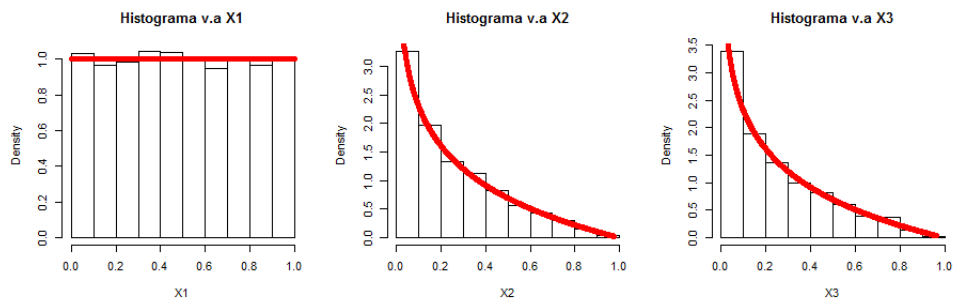


Ilustración 2.1: Histogramas contra funciones marginales de densidad

Se presentarán imágenes de las funciones de densidad conjunta, de las curvas de nivel y de los diagramas de dispersión de cada par de variables aleatorias:

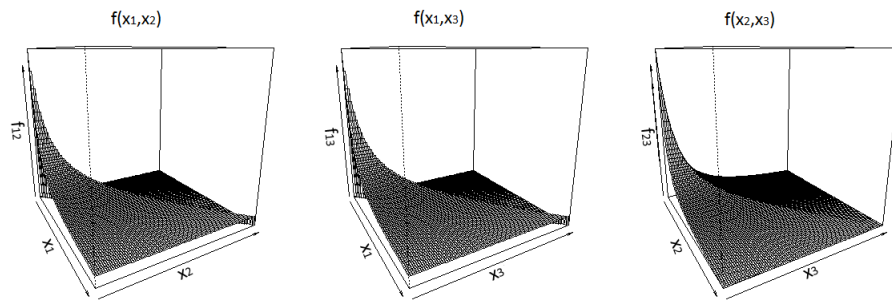


Ilustración 2.2: Gráficos 3D de función densidad bivariada

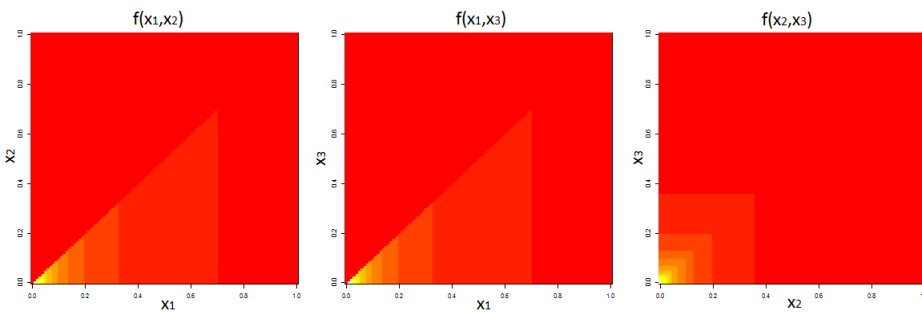


Ilustración 2.3: Curvas de nivel a color de función de densidad bivariada

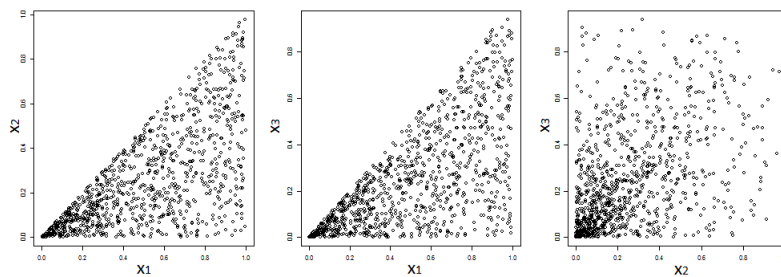


Ilustración 2.4: Diagramas de dispersión de observaciones

Para comprobar que es posible expresar la función de densidad trivariada de este modelo en términos de densidades de cópulas bivariadas como en la expresión (2.i), se muestran las observaciones simuladas empleando el método descrito en sección (1.4) del capítulo uno.

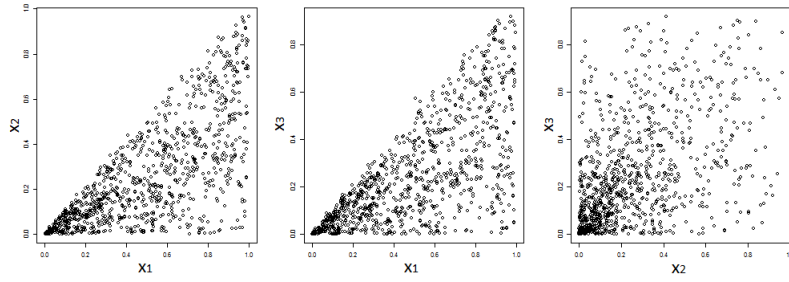


Ilustración 2.5: Diagramas de dispersión de observaciones simuladas

Se aprecia en las imágenes que las observaciones de la simulación original (2.4) coinciden con las nuevas observaciones (2.5).

La tabla a continuación contiene información que permite comparar el coeficiente de correlación teórico contra su estimación muestral.

Tabla 2.II: Coeficiente de correlación teórico v.s muestral

Par de variables aleatorias	Correlación Teórica	Correlación Muestral
(X_1, X_2)	0.6546537	0.6592560
(X_1, X_3)	0.6546537	0.6609024
(X_2, X_3)	0.4286000	0.4424104

Se calculará la esperanza de X_1 condicionada en X_2 y X_3 mediante la función $\varphi(x_2, x_3) = \mathbb{E}[X_1 | X_2 = x_2, X_3 = x_3]$:

$$\begin{aligned}
 \mathbb{E}[X_1 | X_2 = x_2, X_3 = x_3] &= \int_{\mathbb{R}} x_1 f_{1|2,3}(x_1 | x_2, x_3) dx_1 \\
 &= \frac{x_2 \vee x_3}{[1 - (x_2 \vee x_3)]} \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < 1\}} \int_{\mathbb{R}} \frac{\mathbf{1}_{\{x_2 \vee x_3 < x_1 < 1\}}}{x_1} dx_1 \\
 &= -\frac{x_2 \vee x_3}{[1 - (x_2 \vee x_3)]} \log(x_2 \vee x_3) \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < 1\}}
 \end{aligned}
 \tag{2.ñ}$$

Con los datos trivariados simulados se realizará el ajuste clásico de regresión lineal múltiple:

$$X_1 = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \text{ con } \varepsilon \sim \text{Normal}(0, \sigma^2)$$

Se presentará la comparación entre la esperanza condicional $\varphi(x_2, x_3)$ y el ajuste de regresión lineal $\ell_1(x_2, x_3) = \hat{\beta}_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3$.

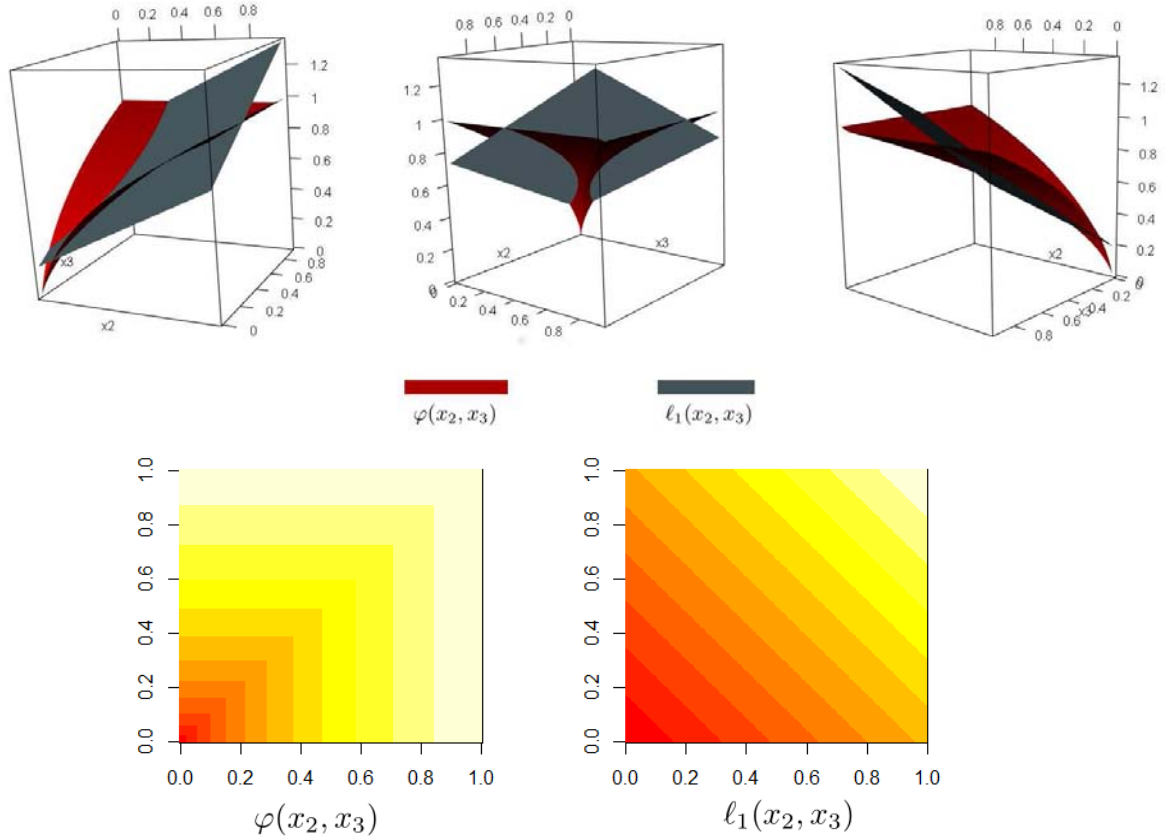


Ilustración 2.6: $\varphi(x_2, x_3)$ versus $\ell_1(x_2, x_3)$

Aquí se logra observar que el ajuste de regresión lineal múltiple falla a menudo en estimar X_1 . En algunas partes subestima el valor de X_1 y en otras partes lo sobrestima.

Se calculará varianza de X_1 condicionada en X_2 y X_3 mediante la función $\psi(x_2, x_3) := \mathbb{V}(X_1 | X_2 = x_2, X_3 = x_3)$ para compararla contra la varianza del ajuste de regresión lineal múltiple.

Sea:

$$\mathbb{V}(X_1 | X_2 = x_2, X_3 = x_3) = \mathbb{E}[X_1^2 | X_2 = x_2, X_3 = x_3] - \mathbb{E}[X_1 | X_2 = x_2, X_3 = x_3]^2$$

Primero calculamos $\mathbb{E}[X_1^2 | X_2 = x_2, X_3 = x_3]$:

$$\begin{aligned} \mathbb{E}[X_1^2 | X_2 = x_2, X_3 = x_3] &= \int_{\mathbb{R}} x_1^2 f_{1|2,3}(x_1 | x_2, x_3) dx_1 \\ &= \int_{\mathbb{R}} x_1^2 \frac{x_2 \vee x_3}{x_1^2 [1 - (x_2 \vee x_3)]} \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < 1\}} \\ &= (x_2 \vee x_3) \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < 1\}} \end{aligned}$$

Utilizando el resultado anterior:

$$\mathbb{V}(X_1 | X_2 = x_2, X_3 = x_3) = ((x_2 \vee x_3) - \frac{(x_2 \vee x_3)^2}{[1 - (x_2 \vee x_3)]^2} \text{Ln}|x_2 \vee x_3|^2) \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < 1\}}$$

Se mostrará gráficamente de la función $\psi(x_2, x_3) := \mathbb{V}(X_1 | X_2 = x_2, X_3 = x_3)$ comparada contra $\hat{\sigma}^2$:

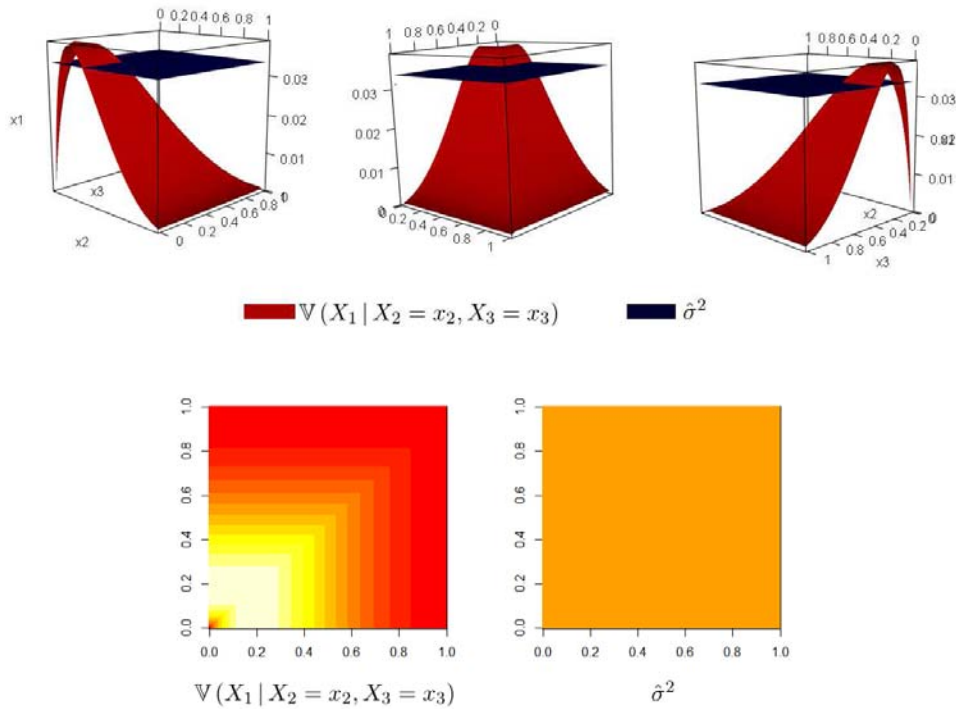


Ilustración 2.7: $\psi(x_2, x_3) := \mathbb{V}(X_1 | X_2 = x_2, X_3 = x_3)$ contra $\hat{\sigma}^2$ del ajuste de regresión lineal múltiple

Debido a que la regresión lineal múltiple tiene como supuesto que la varianza de las observaciones es constante (como se puede ver en el plano azul), se pierde mucha

información sobre el modelo. Para casi todos los casos la varianza del modelo lineal sobreestimó la varianza condicional y en otros pocos, la subestimó.

Aplicaremos la prueba de normalidad de Shapiro-Wilk a los residuales de la regresión lineal múltiple para comprobar que cumplen con los criterios de normalidad.

```

Console C:/Users/Propietario/Dropbox/Tesis Jorge Ingles/Codigo R y Maxima/Modelo trivariado
> shapiro.test(residuals(regresion.mult))

Shapiro-Wilk normality test

data: residuals(regresion.mult)
W = 0.905, p-value < 2.2e-16
> |

```

Ilustración 2.8: Prueba Shapiro-Wilk

Se aprecia que el p-value de la prueba Shapiro-Wilk es demasiado bajo por lo que se rechaza que los residuales estén distribuidos normalmente. Para corroborar esto, graficaremos los residuales y la función de distribución normal que más se acerque.

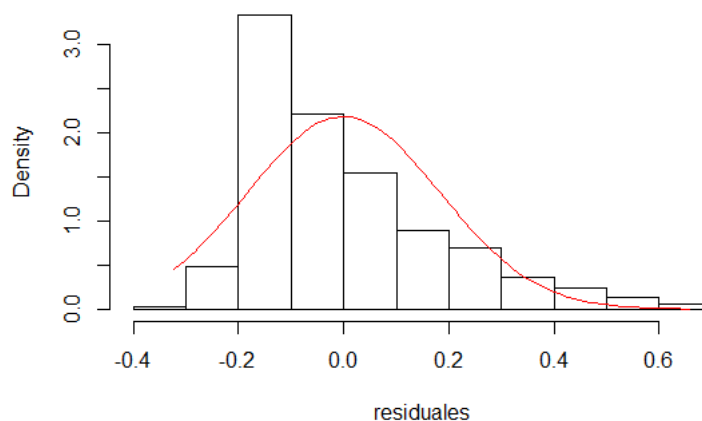


Ilustración 2.9: Histograma de residuales y Normal ajustada

Se puede notar en 2.9 que los residuales están muy lejos de estar distribuidos normalmente.

Mostraremos una gráfica de la función de densidad $f_{1|23}(x_1 | 0.9, 0.8)$ y una densidad $N(\hat{\beta}_1 + 0.9\hat{\beta}_2 + 0.8\hat{\beta}_3, \hat{\sigma}^2)$.

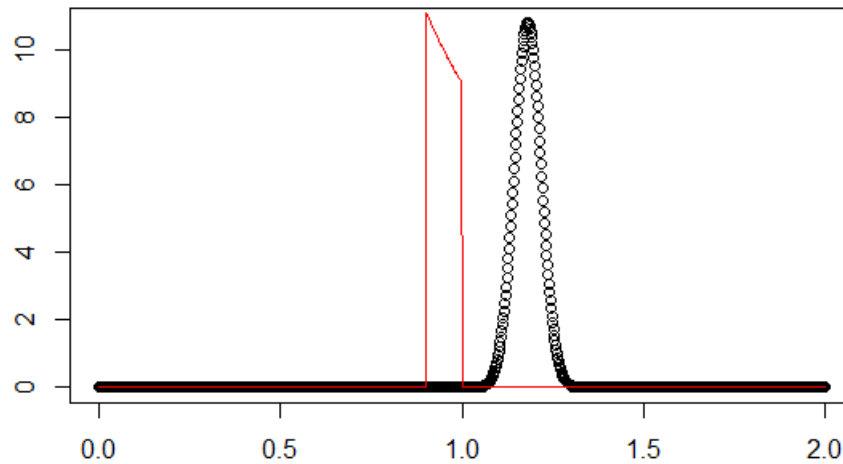


Ilustración 2.10: $f_{1|23}(x_1 | 0.9, 0.8)$ contra Normal $(\hat{\beta}_1 + 0.9\hat{\beta}_2 + 0.8\hat{\beta}_3, \hat{\sigma}^2)$

Se logra ver que la forma de ambas funciones de densidad es muy diferente y sus medidas de tendencia central están alejadas.

Como conclusión del análisis se observa que para este modelo en particular los supuestos de linealidad de las variables aleatorias y de normalidad en los residuales no aplican.

2.4. Cópula Trivariada en Racimo

2.4.1. Antecedentes

Lo expresado en esta subsección se basa en una presentación de Czado (2011).

En esta se comenta que Joe (1996) fue el primero en crear una construcción probabilística de funciones de distribución multivariadas basada en *cóputas por pares*.

Debido a que, para distribuciones de altas dimensiones existen múltiples *cóputas por pares*, Bedford y Cooke (2001) y Bedford y Cooke (2002) introdujeron modelos gráficos llamados racimos regulares y proporcionaron una expresión para la función de densidad conjunta involucrada.

La primera estimación para el caso Gaussiano fue considerado en Kurowicka y Cooke (2006).

Aas, Czado, Frigessi y Bakken (2009) usaron construcciones de *cóputas por pares* para construir cóputas multivariadas basadas en cóputas bivariadas como el caso Gaussiano, t-, Gumbel y Clayton y propusieron expresiones de verosimilitud.

Los primeros y los segundos talleres de cóputas en racimo tomaron lugar en Delft, Holanda en Noviembre de 2007 y Diciembre de 2008. Un tercer taller tomó lugar en Oslo, Noruega en 2009. Los resultados fueron publicados en Kurowicka y Joe (2012).

2.4.2. Justificación matemática

Se demostrará el fundamento de las Cópulas en Racimo caso trivariado.

Demostración 2.4.1. Sea $f_{123}(x_1, x_2, x_3) = f_{3|1,2}(x_3|x_1, x_2) \cdot f_{12}(x_1, x_2)$ tal que:

$$\begin{aligned} f_{3|1,2}(x_3|x_1, x_2) &= \frac{f_{123}(x_1, x_2, x_3)}{f_{12}(x_1, x_2)} \\ &= \frac{f_{23|1}(x_2, x_3|x_1) \cdot f_1(x_1)}{f_{2|1}(x_2|x_1) \cdot f_1(x_1)} \\ &= \frac{c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|1}(x_3|x_1)}{f_{2|1}(x_2|x_1)} \end{aligned} \quad (2.o)$$

Se cumple que:

$$f_{12}(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \quad (2.p)$$

Sustituyendo (2.o) y (2.p) en (2.4.1):

$$\begin{aligned} f_{123}(x_1, x_2, x_3) &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) \cdot \\ & f_{3|1}(x_3|x_1) \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2) \end{aligned} \quad (2.q)$$

Aplicando la igualdad:

$$\begin{aligned} f_{3|1}(x_3|x_1) &= \frac{f_{13}(x_1, x_3)}{f_1(x_1)} \\ &= \frac{c_{13}(F_1(x_1), F_3(x_3)) \cdot f_1(x_1) \cdot f_3(x_3)}{f_1(x_1)} \end{aligned} \quad (2.r)$$

Sustituyendo (2.r) en (2.q):

$$\begin{aligned} f_{123}(x_1, x_2, x_3) &= c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) \cdot \\ & c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \end{aligned} \quad (2.s)$$

Es relativamente fácil aproximar las densidades cópula c_{12} y c_{13} y las funciones de distribución y de densidad marginales condicionales y no condicionales, pero se complica obtener $c_{23|1}$.

En el caso que la cópula condicional $C_{2,3|1}$ es la cópula producto:

$$\begin{aligned} C_{23|1}(F_2(x_2), F_3(x_3) | F_1(x_1)) &= \Pi \\ c_{23|1}(F_2(x_2), F_3(x_3) | F_1(x_1)) &= 1 \end{aligned} \tag{2.t}$$

Por (2.t) se tiene que (2.s) queda así:

$$f_{123}(x_1, x_2, x_3) = c_{13}(F_1(x_1), F_3(x_3)) c_{23}(F_2(x_2), F_3(x_3)) \cdot f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \tag{2.u}$$

2.4.3. Generalización del modelo teórico

Supóngase que:

$$X_1 \sim \text{Unif}(0, 1) \quad X_2|X_1 = x_1 \sim \text{Unif}(0, x_1) \quad X_3|X_1 = x_1 \sim \text{Unif}(0, x_1 x_2^\theta)$$

Este modelo es la generalización del desarrollado en la subsección (2.3) y con él se buscará demostrar la factibilidad de la metodología de Cópulas en Racimo caso trivariado.

Se requiere obtener explícitamente las fórmulas de todas las densidades marginales univariadas y bivariadas, condicionales y no condicionales de las variables aleatorias involucradas.

$$f(x_1) = \mathbf{1}_{\{0 < x_1 < 1\}}$$

$$f(x_2) = (-\log x_2) \mathbf{1}_{\{0 < x_2 < 1\}}$$

$$f(x_3) = \begin{cases} (-\log x_3) \mathbf{1}_{\{0 < x_3 < 1\}} & \text{si } \theta = 0 \\ \left[\log(x_3) + \frac{2}{\sqrt{x_3}} - 2 \right] \mathbf{1}_{\{0 < x_3 < 1\}} & \text{si } \theta = 1 \\ \left[\frac{1-x_3^{-\frac{\theta}{\theta+1}}}{\theta^2-\theta} + \frac{\theta x_3^{-1+\frac{1}{\theta}}}{(-\theta+1)} \left(1 - x_3^{-\frac{1}{\theta^2+\theta}} \right) \right] \mathbf{1}_{\{0 < x_3 < 1\}} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$f(x_1, x_2) = \frac{1}{x_1} \mathbf{1}_{\{0 < x_2 < x_1 < 1\}}$$

$$f(x_1, x_3) = \begin{cases} \frac{1}{x_1} \mathbf{1}_{\{0 < x_3 < x_1 < 1\}} & \text{si } \theta = 0 \\ \left[\frac{2 \log x_1 - \log x_3}{x_1^2} \right] \mathbf{1}_{\{0 < x_3 < x_1^2\}} \mathbf{1}_{\{0 < x_1 < 1\}} & \text{si } \theta = 1 \\ \mathbf{1}_{\{0 < x_3 < x_1^{\theta+1}\}} \mathbf{1}_{\{0 < x_1 < 1\}} \frac{1}{-\theta+1} \left[x_1^{-\theta-1} - x_3^{-1+\frac{1}{\theta}} x_1^{-1-\frac{1}{\theta}} \right] & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$f(x_2, x_3) = \mathbf{1}_{\{0 < x_2 < 1\}} \mathbf{1}_{\{0 < x_3 < x_2^\theta\}} \frac{1}{x_2^\theta} \left[\left(-1 + \frac{1}{x_2} \right) \mathbf{1}_{\{x_3 < x_2^{\theta+1}\}} + \left(-1 + \frac{x_2^\theta}{x_3} \right) \mathbf{1}_{\{x_3 \geq x_2^{\theta+1}\}} \right]$$

$$f_{12|3}(x_1 x_2 | x_3) = \begin{cases} -\frac{\mathbf{1}_{\{x_3 \vee x_2 < x_1 < 1\}} \mathbf{1}_{\{0 < x_2 < 1\}}}{x_1^2 \log(x_3)} & \text{si } \theta = 0 \\ \frac{\mathbf{1}_{\{\frac{x_3}{x_1} < x_2 < 1\}} \mathbf{1}_{\{0 < x_1 < 1\}}}{x_1^2 x_2 \left(\log(x_3) + \frac{2}{\sqrt{x_3}} - 2 \right)} & \text{si } \theta = 1 \\ \frac{\mathbf{1}_{\{x_3^{\frac{1}{1+\theta}} < x_1 < 1\}} \mathbf{1}_{\{(\frac{x_3}{x_1})^{\frac{1}{\theta}} < x_2 < x_1\}}}{x_1^2 x_2^\theta \left(\frac{\theta x_3^{\frac{1}{\theta}-1} \left(1 - x_3^{\frac{-1}{\theta^2+\theta}} \right)}{1-\theta} + \frac{1-x_3^{\frac{-\theta}{\theta^2-\theta}}}{\theta^2-\theta} \right)} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$f_{13|2}(x_1 x_3 | x_2) = -\frac{1}{x_1^2 x_2^\theta \log x_2} \mathbf{1}_{\{0 < x_3 < x_1 x_2^\theta\}} \mathbf{1}_{\{x_2 < x_1 < 1\}}$$

$$f_{23|1}(x_2 x_3 | x_1) = \frac{1}{x_1^2 x_2^\theta} \mathbf{1}_{\{0 < x_3 < x_1 x_2^\theta\}} \mathbf{1}_{\{0 < x_2 < x_1\}}$$

$$f_{1|2}(x_1 | x_2) = -\frac{\mathbf{1}_{\{0 < x_2 < x_1 < 1\}}}{x_1 \log x_2}$$

$$f_{1|3}(x_1 | x_3) = \begin{cases} -\frac{\mathbf{1}_{\{0 < x_3 < x_1\}} \mathbf{1}_{\{0 < x_1 < 1\}}}{x_1 \log(x_3)} & \text{si } \theta = 0 \\ \frac{[2 \log(x_1) - \log(x_3)] \mathbf{1}_{\{\sqrt{x_3} < x_1 < 1\}}}{x_1^2 \left(\log(x_3) + \frac{2}{\sqrt{x_3}} - 2 \right)} & \text{si } \theta = 1 \\ \frac{\mathbf{1}_{\{x_3^{\frac{1}{\theta+1}} < x_1 < 1\}} \left(x_1^{-\theta-1} - x_1^{-\frac{1}{\theta}-1} x_3^{\frac{1}{\theta}-1} \right)}{(1-\theta) \left[\frac{\theta x_3^{\frac{1}{\theta}-1}}{1-\theta} \left(1 - x_3^{\frac{-1}{\theta^2+\theta}} \right) + \frac{1-x_3^{\frac{-\theta}{\theta^2-\theta}}}{\theta^2-\theta} \right]} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$f_{2|1}(x_2|x_1) = \frac{1}{x_1} \mathbf{1}_{\{0 < x_2 < x_1\}}$$

$$f_{2|3}(x_2|x_3) = \begin{cases} \frac{\mathbf{1}_{\{0 < x_2 < 1\}} \left[\left(-1 + \frac{1}{x_2} \right) \mathbf{1}_{\{x_3 < x_2\}} + \left(-1 + \frac{1}{x_3} \right) \mathbf{1}_{\{x_3 \geq x_2\}} \right]}{\log(x_3)} & \text{si } \theta = 0 \\ \frac{\mathbf{1}_{\{0 < x_3 < x_2 < 1\}} \left[\left(-1 + \frac{1}{x_2} \right) \mathbf{1}_{\{\sqrt{x_3} < x_2\}} + \left(-1 + \frac{x_2}{x_3} \right) \mathbf{1}_{\{\sqrt{x_3} \geq x_2\}} \right]}{x_2 \left(\log(x_3) + \frac{2}{\sqrt{x_3}} - 2 \right)} & \text{si } \theta = 1 \\ \frac{\mathbf{1}_{\{x_3^{\frac{1}{\theta}} < x_2 < 1\}} \left[\left(-1 + \frac{1}{x_2} \right) \mathbf{1}_{\{x_3 < x_2^{\theta+1}\}} + \left(-1 + \frac{x_2^\theta}{x_3} \right) \mathbf{1}_{\{x_3 \geq x_2^{\theta+1}\}} \right]}{x_2^\theta \left[\frac{\theta x_3^{\frac{1}{\theta}-1}}{1-\theta} \left(1 - x_3^{\frac{-1}{\theta^2+\theta}} \right) + \frac{1-x_3^{\frac{-\theta}{\theta+1}}}{\theta^2-\theta} \right]} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$f_{3|1}(x_3|x_1) = \begin{cases} \frac{1}{x_1} \mathbf{1}_{\{0 < x_3 < x_1\}} \sim Unif(0, x_1) & \text{si } \theta = 0 \\ \frac{\mathbf{1}_{\{0 < x_3 < x_1^2\}} (2 \log x_1 - \log x_3)}{x_1^2} & \text{si } \theta = 1 \\ \frac{\mathbf{1}_{\{0 < x_3 < x_1^{1+\theta}\}} \left[x_1^{-\theta-1} - x_1^{-\frac{1}{\theta}-1} x_3^{\frac{1}{\theta}-1} \right]}{1-\theta} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$f_{3|2}(x_3|x_2) = \frac{\mathbf{1}_{\{0 < x_3 < x_2^\theta\}} \left[\left(-1 + \frac{1}{x_2} \right) \mathbf{1}_{\{x_3 < x_2^{\theta+1}\}} + \left(-1 + \frac{x_2^\theta}{x_3} \right) \mathbf{1}_{\{x_3 \geq x_2^{\theta+1}\}} \right]}{x_2^\theta \log x_2}$$

$$f_{1|23}(x_1|x_2x_3) = \begin{cases} \frac{x_3}{x_1^2 (x_2^\theta - x_3)} \mathbf{1}_{\{x_3 < x_2^\theta < x_1 < 1\}} & \text{si } x_3 \geq x_2^{1+\theta} \\ \frac{x_2}{x_1^2 (1-x_2)} \mathbf{1}_{\{x_2 < x_1 < 1\}} & \text{si } x_3 < x_2^{1+\theta} \end{cases}$$

$$f_{2|13}(x_2|x_1, x_3) = \begin{cases} \frac{1}{x_1} \mathbf{1}_{\{0 < x_2 < x_1\}} \sim Unif(0, x_1) & \text{si } \theta = 0, \text{ sólo si } x_3 \leq x_1 \\ \frac{\mathbf{1}_{\{\frac{x_3}{x_1} < x_2 < x_1\}}}{x_2 (2 \log(x_1) - \log(x_3))} & \text{si } \theta = 1, \text{ sólo si } x_3 \leq x_1^2 \\ \frac{(1-\theta) \mathbf{1}_{\{0 < \theta \sqrt{\frac{x_3}{x_1}} < x_2 < x_1\}}}{x_2^\theta \left(x_1^{-\theta+1} - x_1^{-\frac{1}{\theta}+1} x_3^{\frac{1}{\theta}-1} \right)} & \text{si } 0 < \theta \neq 1, \text{ sólo si } x_3 \leq x_1^{1+\theta} \end{cases}$$

Deber ser $x_2 < x_1$:

$$f_{3|12}(x_3|x_1x_2) = \frac{1}{x_1x_2^\theta} \mathbf{1}_{\{0 < x_3 < x_1x_2^\theta\}} \sim Unif(0, x_1x_2^\theta), \text{ sólo si } x_2 \leq x_1$$

Se simularon 1000 observaciones utilizando el método condicional para cada uno de los valores $\theta := \{0.01, 0.49, 0.51, 0.99, 1.01, 1.98, 2.03\}$, aplicando el siguiente algoritmo:

Para $k = 1, 2, \dots, 1000$ {
 Simula $x_1[k]$ a partir de una $\text{Unif}(0, 1)$
 Simula $x_2[k]$ a partir de una $\text{Unif}(0, x_1[k])$
 Simula $x_3[k]$ a partir de una $\text{Unif}(0, x_1[k]x_2[k]^\theta)$
 Siguiente k }

Los datos simulados tienen los siguientes estimadores:

Tabla 2.III: Resumen estadística descriptiva (parte I)

Valor de θ :	0.01			0.49		
Estimador	X_1	X_2	X_3	X_1	X_2	X_3
Mín	0.0002778	0.0000189	0.000019	0.001117	0.0002016	0.0002016
Primer cuartil	0.2480216	0.0685859	0.066011	0.235966	0.0569116	0.0177851
Mediana	0.4962564	0.1834776	0.182900	0.489564	0.1674123	0.0676401
Media	0.4979165	0.2474927	0.244359	0.489526	0.2389155	0.1249695
Tercer cuartil	0.7474515	0.3796377	0.380847	0.740994	0.3836249	0.1791973
Máx	0.9993089	0.9674614	0.984238	0.999433	0.9843883	0.7921863

Tabla 2.IV: Resumen estadística descriptiva (parte II)

Valor de θ :	0.51			0.99		
Estimador	X_1	X_2	X_3	X_1	X_2	X_3
Mín	0.006948	0.0000395	0.0000114	0.0007908	0.000003	0.0000
Primer cuartil	0.257997	0.0703965	0.0196630	0.2560091	0.073335	0.005816
Mediana	0.501295	0.1927635	0.0694329	0.5100030	0.191072	0.031490
Media	0.502743	0.2480104	0.1300667	0.5045503	0.253906	0.085611
Tercer cuartil	0.739407	0.3784533	0.2012109	0.7499379	0.386899	0.113763
Máx	0.998623	0.9795470	0.9123191	0.9999650	0.978490	0.91556

Tabla 2.V: Resumen estadística descriptiva (parte III)

Valor de θ :	1.01			1.98		
Estimador	X_1	X_2	X_3	X_1	X_2	X_3
Mín	0.000001	0.0000004	0.000000	0.0007863	0.0000948	0.0000000
Primer cuartil	0.257338	0.0699703	0.005125	0.2390764	0.0700717	0.0004798
Mediana	0.503960	0.1912412	0.031066	0.4868146	0.1821216	0.0053027
Media	0.504263	0.2555920	0.086188	0.4922640	0.2415809	0.0406740
Tercer cuartil	0.754926	0.3945663	0.113532	0.7425003	0.3489544	0.0320442
Máx	0.999838	0.9788690	0.894279	0.9991382	0.9559116	0.8696321

Tabla 2.VI: Resumen estadística descriptiva (parte IV)

Valor de θ :	2.03		
Estimador	X_1	X_2	X_3
Mín	0.001182	0.0002351	0.0000000
Primer cuartil	0.256781	0.0620581	0.0003029
Mediana	0.501598	0.1875705	0.0052339
Media	0.502812	0.2524870	0.0433318
Tercer cuartil	0.754666	0.3931779	0.0426139
Máx	0.999809	0.9580420	0.6808802

Nótese que se emplearon valores de θ cercanos a 0.5, 1 y 2 respectivamente. Esto se debe a que si el modelo es evaluado en dichos valores se incurre en casos especiales, por lo que se prefirió utilizar números que se aproximarán.

Se mostrarán los histogramas de las observaciones simuladas, acompañados de la función de distribución marginal teórica correspondiente.

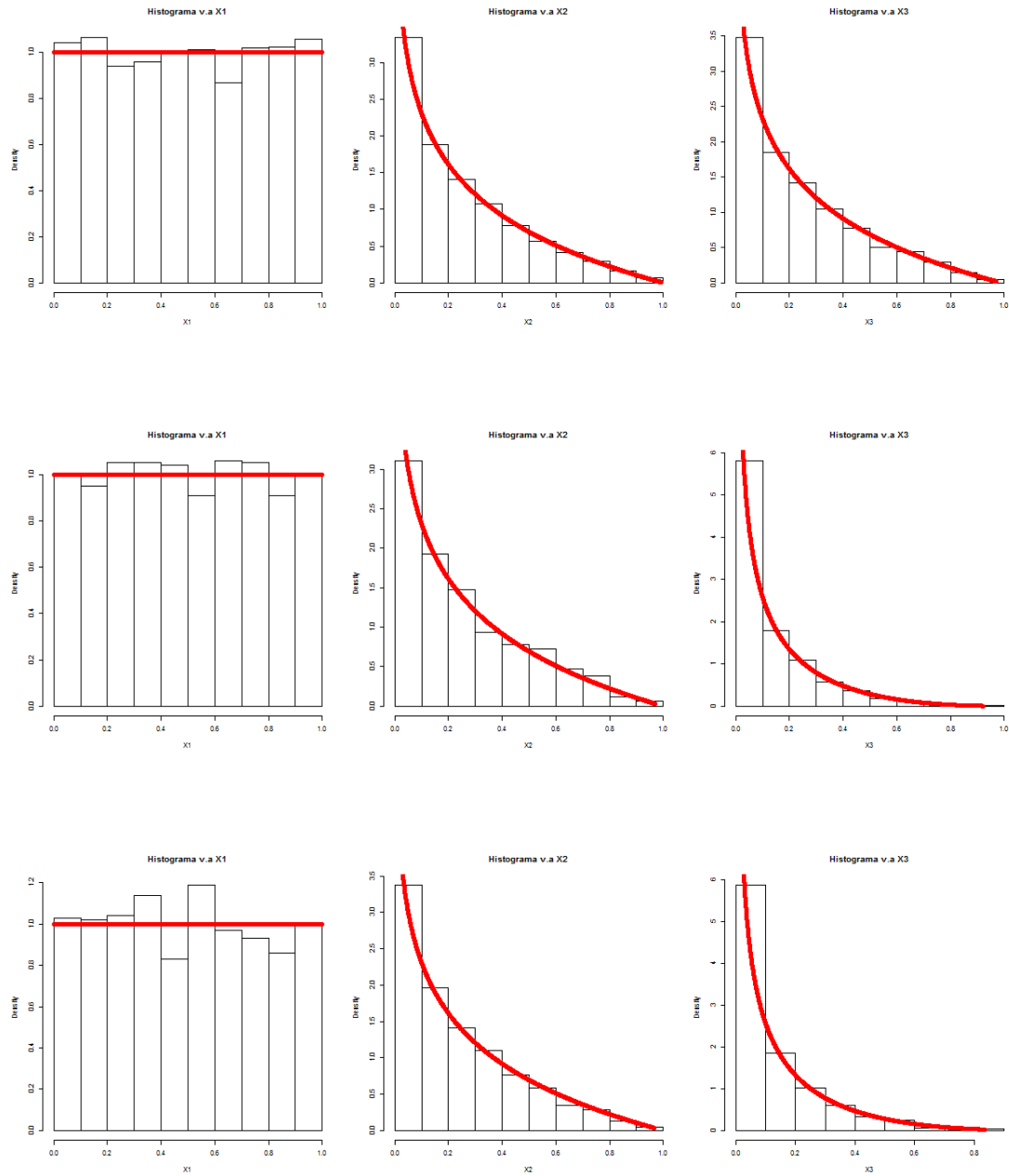


Ilustración 2.11: Histogramas de X_1, X_2 y X_3 con $\theta = \{0.01, 0.49, 0.51\}$

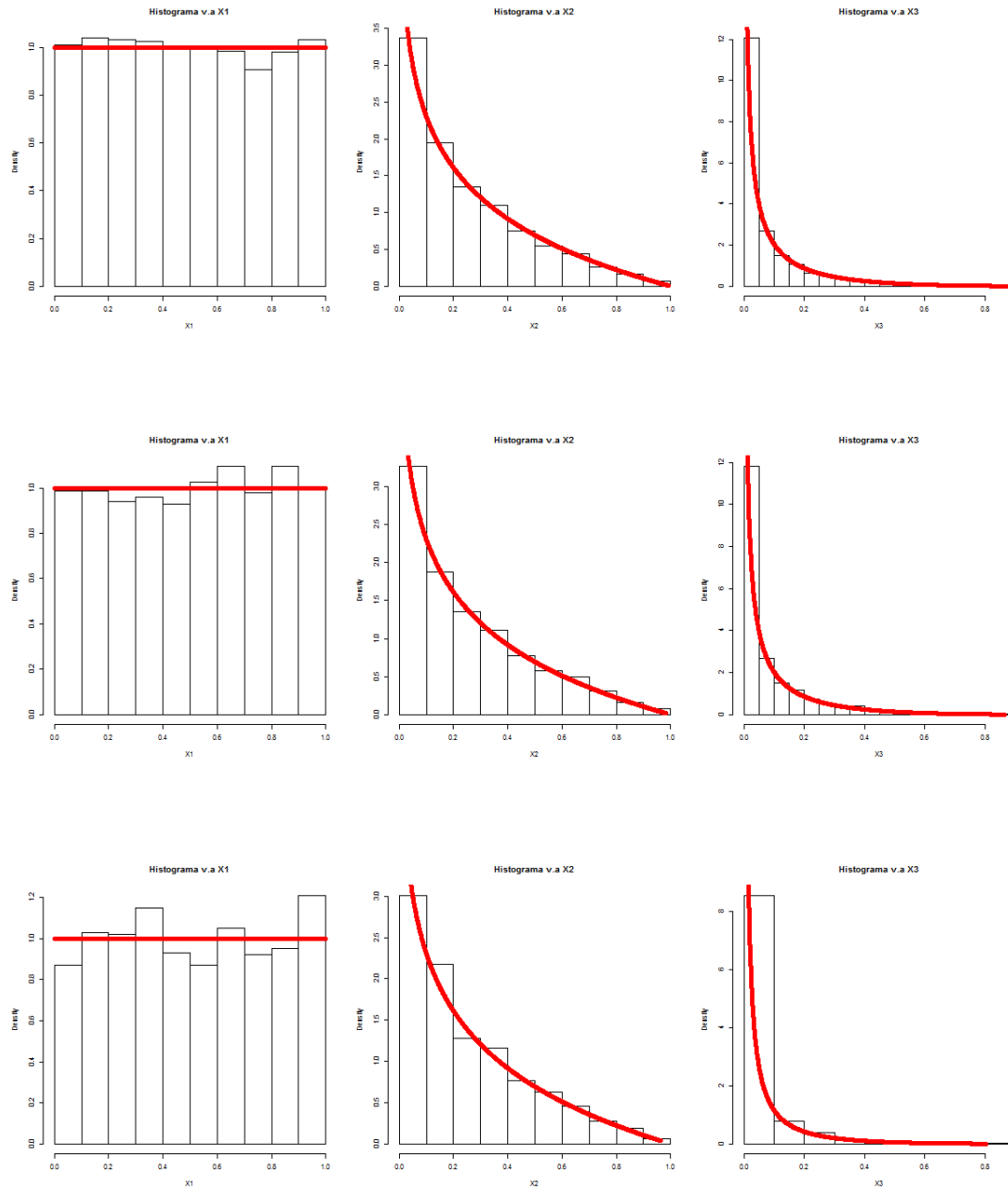


Ilustración 2.12: Histogramas de X_1, X_2 y X_3 con $\theta = \{0.99, 1.01, 1.98\}$

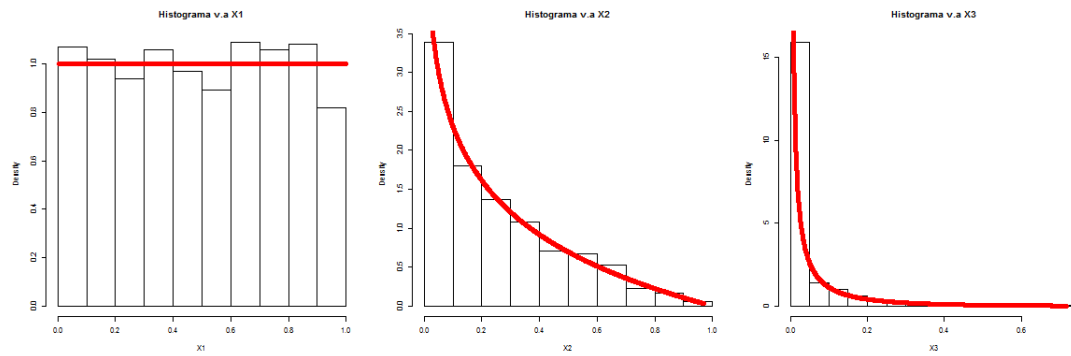
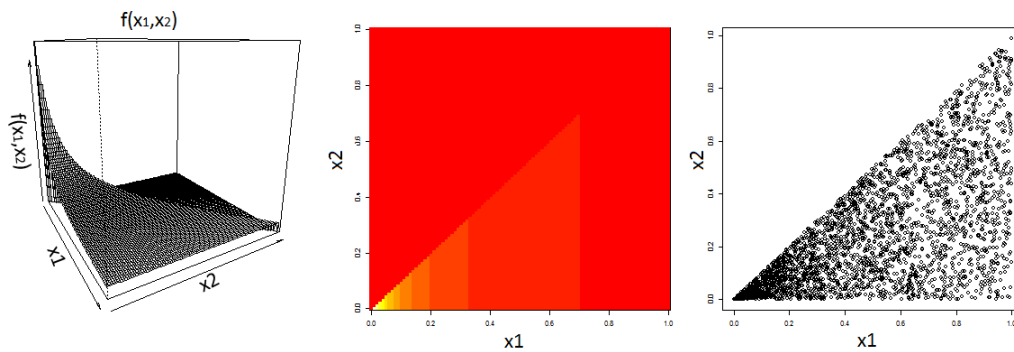


Ilustración 2.13: Histogramas de X_1, X_2 y X_3 con $\theta = 2.03$

Ahora se mostrarán gráficas de las funciones de densidad conjunta y de las observaciones simuladas de cada par de variables aleatorias para todos los valores de θ propuestos en el modelo.



Nótese aquí que no aparece en esta imagen referencia alguna a θ porque la relación entre X_1 y X_2 no se ve alterada por dicho parámetro.

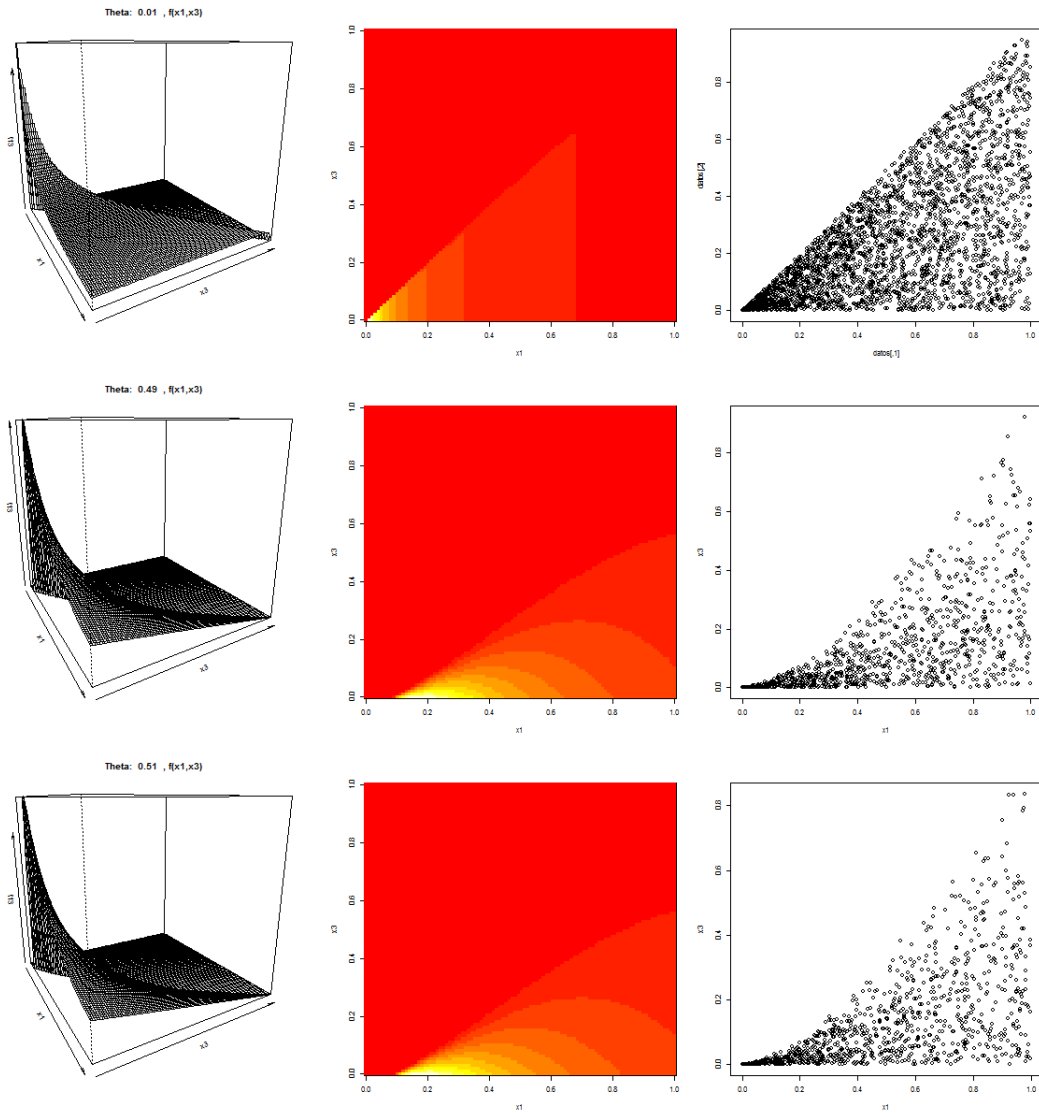


Ilustración 2.14: Densidad de función bivariada (X_1, X_3) con $\theta = \{0.01, 0.49, 0.51\}$

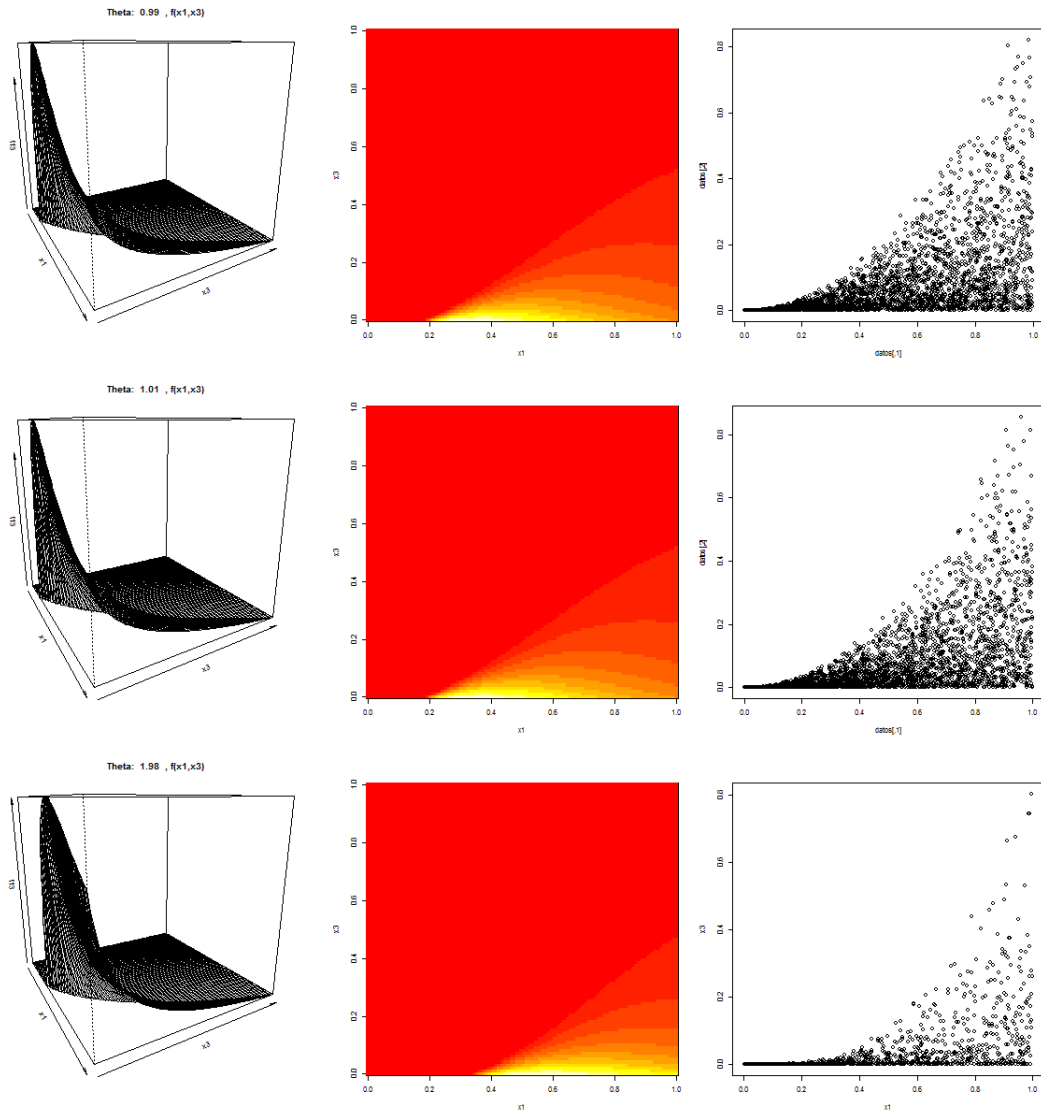


Ilustración 2.15: Densidad de función biviada (X_1, X_3) con $\theta = \{0.99, 1.01, 1.98\}$

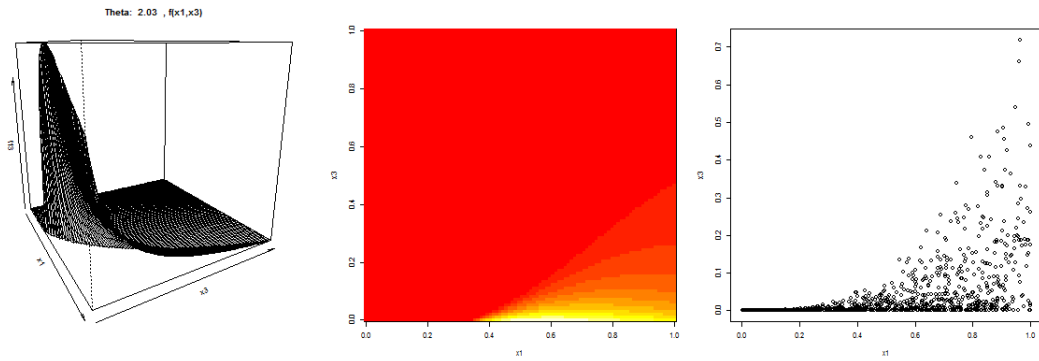


Ilustración 2.16: Densidad de función bivariada (X_1, X_3) con $\theta = 2.03$

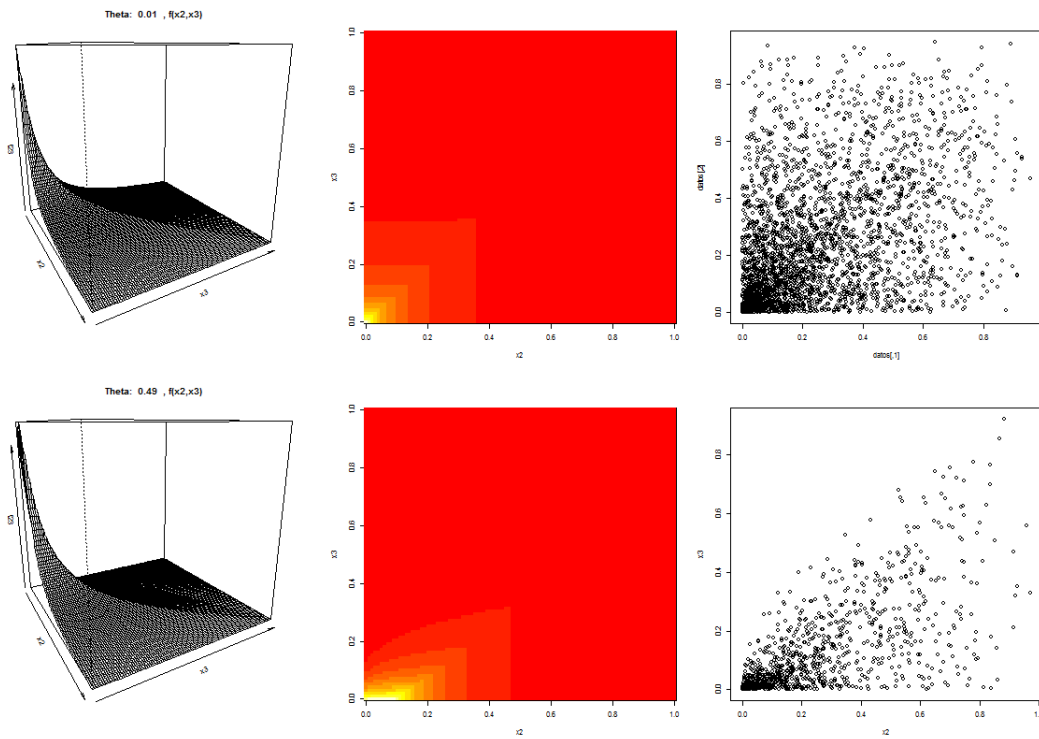


Ilustración 2.17: Densidad de función bivariada (X_2, X_3) con $\theta = \{0.01, 0.49\}$

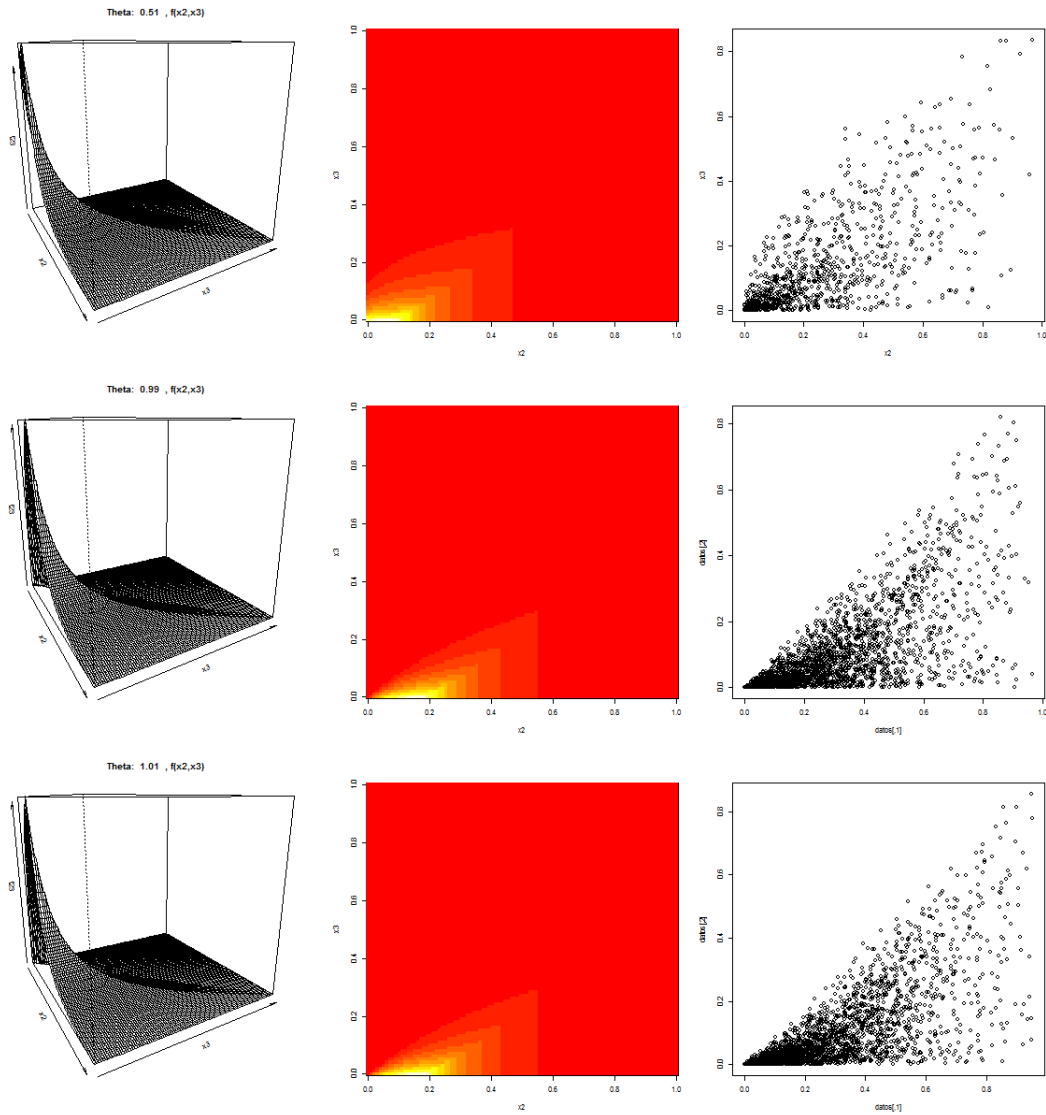


Ilustración 2.18: Densidad de función bivariada (X_2, X_3) con $\theta = \{0.51, 0.99, 1.01\}$

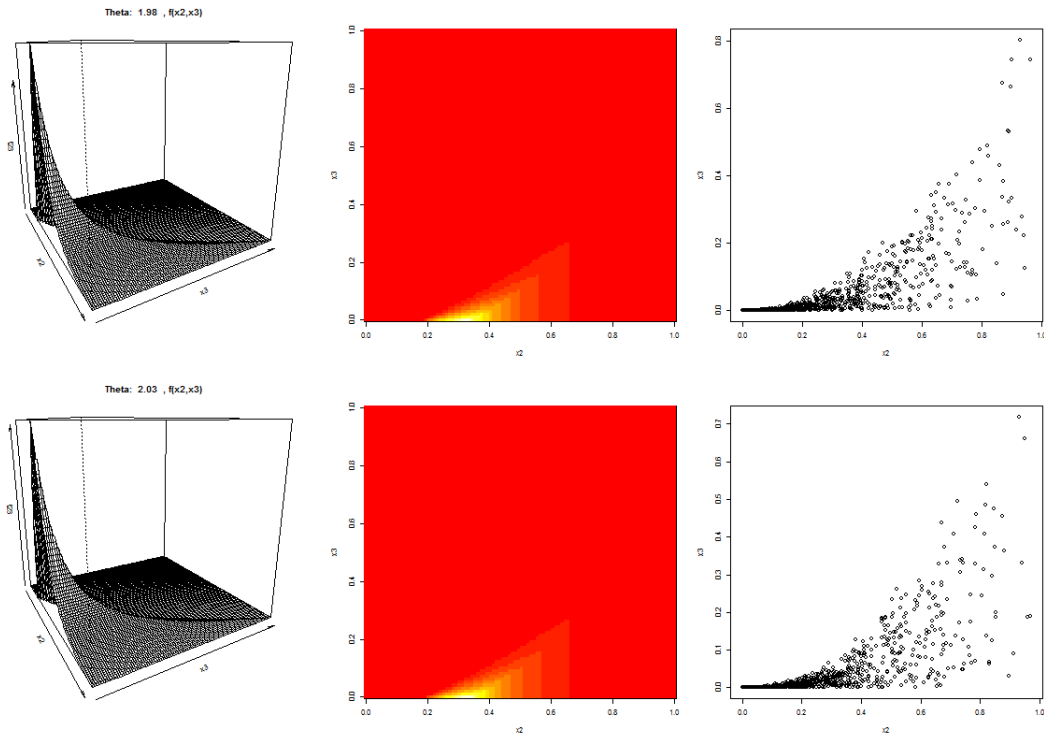


Ilustración 2.19: Densidad de función biviada (X_2, X_3) con $\theta = \{1.98, 2.03\}$

Una de las metas del presente trabajo es demostrar que las cópulas capturan mejor la estructura de dependencia de variables aleatorias frente al enfoque tradicional del coeficiente de correlación.

Se presentará un análisis de las observaciones simuladas empleando el coeficiente de correlación.

$$\rho_{1,2}(x_1, x_2) = 0.6546537$$

$$\rho_{1,3}(x_1, x_3) = \begin{cases} 0.6546537 & \text{si } \theta = 0 \\ 0.583874 & \text{si } \theta = 1 \\ \frac{2\sqrt{3}\left(\frac{1}{2\theta^2+8\theta+6} - \frac{1}{2(2\theta^2+6\theta+4)}\right)}{\sqrt{\frac{1}{12\theta^2+24\theta+9} - \frac{1}{(2\theta^2+6\theta+4)^2}}} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$\rho(X_2, X_3) = \begin{cases} 0.4286 & \text{si } \theta = 0 \\ 0.7645 & \text{si } \theta = 1 \\ \frac{\frac{-1}{4(2\theta^2+6\theta+4)} - \frac{1}{2(\theta+3)} + \frac{1}{2(\theta+2)}}{\frac{\sqrt{7}}{12} \sqrt{\frac{1}{12\theta^2+24\theta+9} - \frac{1}{(2\theta^2+6\theta+4)^2}}} & \text{si } 0 < \theta \neq 1 \end{cases}$$

$$\rho(X_1, X_2|X_3 = x_3, \theta = 0) = (x_3^2 - 1) \log(x_3) - 2x_3^2 + 4x_3 - 2$$

$$\cdot \left[\frac{1}{4 \log(x_3)^2 \sqrt{\frac{-1}{\log(x_3)} \left(\frac{1}{2} - \frac{x_3^2}{2}\right) - \frac{1}{\log(x_3)^2} (1 - x_3)^2}} \right]$$

$$\cdot \left[\frac{1}{\sqrt{\frac{1}{\log(x_3)} \left(\frac{-1}{6} + \frac{x_3^2}{6}\right) - \frac{1}{\log(x_3)^2} \left(\frac{x_3}{2} - \frac{1}{2}\right)^2}} \right]$$

$$\rho(X_1, X_2 | X_3 = x_3, \theta = 1) =$$

$$\left[\left[\frac{1 + x_3 - \left(\sqrt{x_3} + \frac{x_3}{\sqrt{x_3}} \right)}{f_3(x_3)} \right] - \left[\frac{-\log \sqrt{x_3}^2 + (\log x_3 \cdot \log \sqrt{x_3})}{f_3(x_3)} \cdot \frac{x_3 - \log \sqrt{x_3} - \frac{1}{2}}{f_3(x_3)} \right] \right]$$

$$\cdot \frac{1}{\sqrt{\frac{2 \cdot (-1 - \sqrt{x_3} \cdot \log \sqrt{x_3} + \sqrt{x_3})}{f_3(x_3)} - \frac{\log x_3 \cdot (1 - \sqrt{x_3})}{f_3(x_3)} - \left(\frac{\log x_3 \cdot \log \sqrt{x_3} - \log \sqrt{x_3}^2}{f_3(x_3)} \right)^2}}$$

$$\cdot \frac{1}{\sqrt{\frac{\frac{1}{2} - 2 \cdot \frac{\sqrt{x_3}}{3} + \frac{x_3^2}{6}}{f_3(x_3)} - \left(\frac{\frac{x_3}{2} - \log \sqrt{x_3} - \frac{1}{2}}{f_3(x_3)} \right)^2}}$$

Para $0 < \theta \neq 1$:

$$\rho(X_1, X_2 | X_3 = x_3) = \frac{-\frac{x_3^{\frac{2-\theta}{\theta+1}}}{2-\theta} - \theta x_3^{\frac{1}{\theta+1}} + \theta x_3^{\frac{1}{\theta}} + \frac{1}{2-\theta}}{(2-\theta) f_3(x_3 | \theta) \sqrt{\frac{(\theta+1) \left(1 - x_3^{\frac{2-\theta}{\theta+1}} \right)}{(2\theta^2 - 5\theta + 2) f_3(x_3 | \theta)} - \left(\frac{(-\theta-1) x_3^{\frac{1-\theta}{\theta+1}} + \theta x_3^{\frac{1}{\theta} - 1} + 1}{(1-\theta)^2 f_3(x_3 | \theta)} \right)^2}}$$

$$\cdot \frac{1}{\sqrt{\frac{\frac{\frac{3}{3} - 1}{3} - \frac{x_3^{\frac{2-\theta}{\theta+1}}}{2-\theta} + \frac{1}{\theta^2 - 5\theta + 6}}{(1-\theta) f_3(x_3 | \theta)} - \left(\frac{\frac{x_3^{\frac{2}{\theta+1} - 1}}{2} - \frac{x_3^{\frac{1-\theta}{\theta+1}}}{1-\theta} + \frac{1}{\theta^2 - 3\theta + 2}}{(1-\theta)^2 f_3(x_3 | \theta)^2} \right)^2}}$$

$$\rho(X_1, X_3 | X_2 = x_2) =$$

$$\frac{\sqrt{6} \left((x_2 - 1) \log(x_2) + (-x_2^2 + 2x_2 - 1) \right)}{\sqrt{(2x_2^2 - 2) \log(x_2) + (-3x_2^2 + 6x_2 - 3) \sqrt{(x_2^2 - 1) \log(x_2) - 2x_2^2 + 4x_2 - 2}}}$$

$$\rho(X_2, X_3 | X_1 = x_1) = \begin{cases} 0.4286 & \text{si } \theta = 0 \\ \frac{\sqrt{3}}{\sqrt{7}} & \text{si } \theta = 1 \\ \frac{\sqrt{3}\theta}{(2\theta^2 + 6\theta + 4) \sqrt{\frac{4\theta^2 + 2\theta + 1}{24\theta^3 + 60\theta^2 + 48\theta + 12}}} & \text{si } 0 < \theta \neq 1 \end{cases}$$

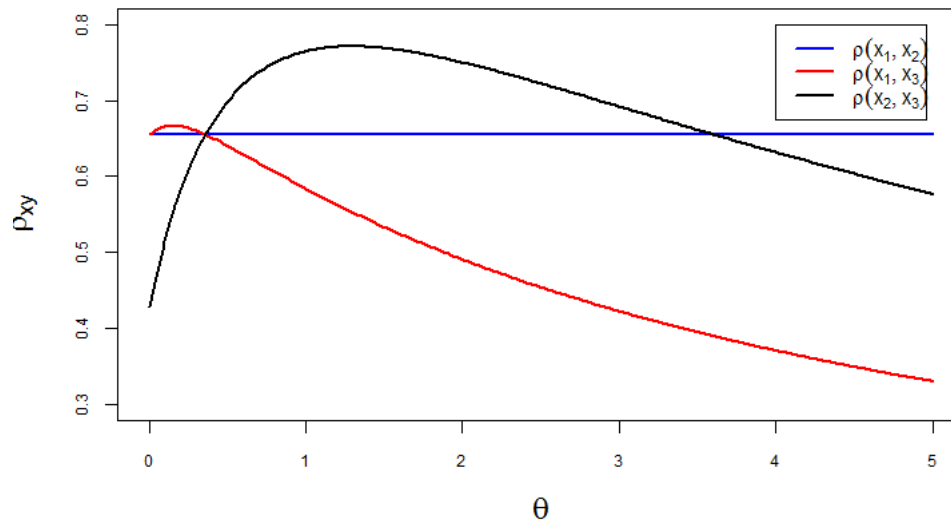
Se presentará una comparación entre las correlaciones no condicionales calculadas a partir de las observaciones simuladas y las correlaciones obtenidas teóricamente.

Tabla 2.VII: Coeficiente de correlación teórico v.s muestral

	$\rho(X_1, X_2)$		$\rho(X_1, X_3)$		$\rho(X_2, X_3)$	
Valor de θ :	Teórico	Muestral	Teórico	Muestral	Teórico	Muestral
0.01	0.6546537	0.6557658	0.656432	0.6502129	0.4382452	0.4406902
0.49	0.6550596	0.6195391	0.6414544	0.61425	0.6961264	0.7035372
0.51	0.6550596	0.6205	0.6392598	0.6170814	0.7011849	0.684851
0.99	0.6550596	0.6507174	0.5849638	0.5798865	0.7639731	0.7728218
1.01	0.6546537	0.6444483	0.5827876	0.5690081	0.7649462	0.7675104
1.98	0.6550596	0.6457003	0.49226	0.4818018	0.7504977	0.7406679
2.03	0.6550596	0.6596439	0.4883223	0.5048251	0.7480347	0.757471

Es de apreciar que las correlaciones muestrales difieren muy poco de los resultados teóricos, lo cual sugiere que el ejercicio se ha calculado de forma correcta.

Se muestra una gráfica de las tres correlaciones no condicionales en donde se toman los valores de θ para formar el eje de las abscisas.

**Ilustración 2.20:** Comparación entre correlaciones

Es preciso observar que el valor de $\rho(X_1, X_3)$ crece hasta que $\theta = 0.15$, posteriormente

su valor desciende para todos los valores restantes de θ . Se tiene que $\rho(X_1 X_3)$ toma el valor de $\rho(X_1 X_2)$ cuando $\theta = 0$ y $\theta = 0.35$.

En el caso de $\rho(X_2 X_3)$, su valor crece para $0 \leq \theta < 1.35$, para descender cuando $\theta \geq 1.35$. Toma el valor de $\rho(X_1 X_2)$ cuando $\theta = 0.35$ y $\theta = 3.6$.

El valor de $\rho(X_1 X_2)$ se mantiene constante para todos los valores de θ .

$\rho(X_1 X_3)$ y $\rho(X_2 X_3)$ ascienden en algunas partes de la gráfica debido a que X_2^θ tiende a X_2 cuando $\theta \approx 1$, y el valor de estas correlaciones llegan a 0 debido a que X_2^θ tiende a 0 cuando $\theta \rightarrow \infty$.

2.4.4. Supuesto de simplificación

Es necesario saber si la cópula condicional del modelo explicado en la subsección (2.4.3) es la misma o muy parecida aún cuando cambian los valores de la variable en que está condicionada. De ser así es posible asumir que la cópula es independiente de dicha variable, a esto se le denomina supuesto de simplificación.

Para obtener una aproximación a la cópula condicional se deben obtener primero las pseudo-observaciones condicionales de la muestra. En el siguiente teorema se explican estos conceptos.

Teorema 2.4.2. *Sea (Y_1, Y_2, X) un vector trivariado donde Y_1 y Y_2 están condicionadas en X tal que $(Y_1, Y_2, X) \sim H(y_1, y_2, x)$. Las marginales bivariadas de H son las siguientes:*

$$\begin{aligned} (Y_1, X) &\sim H_1(y_1, x) \rightarrow F_{1x}(y_1) := \mathbb{P}(Y_1 \leq y_1 | X = x) \\ (Y_2, X) &\sim H_2(y_2, x) \rightarrow F_{2x}(y_2) := \mathbb{P}(Y_2 \leq y_2 | X = x) \end{aligned}$$

De aquí se desprende que:

$$\begin{aligned} H_x(y_1, y_2) &= \mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2 | X = x) \\ &= C_X(F_{1x}(y_1), F_{2x}(y_2)) \end{aligned}$$

Entonces C_X es la cópula bivariada condicional en X tal que:

$$C_X(u_1, u_2) = H_X(F_{1x}^{-1}(u_1), F_{2x}^{-1}(u_2))$$

Por lo tanto se concluye que:

$$(y_{1i}, y_{2i}, x_i) \rightarrow (F_{1x}(y_{1i}), F_{2x}(y_{2i})) = (u_{1i}, u_{2i}) \sim C_{X_i}(u_1, u_2) \quad (2.v)$$

Cuando la cópula condicional puede expresarse de forma independiente a la variable en que está condicionada, o sea $C_{X_i}(u_1, u_2) \sim C^*(u_1, u_2)$, se dice que se cumple el supuesto de simplificación.

A u_{1i} y u_{2i} se les denomina pseudo-observaciones condicionales en X_i .

Se mostrarán las curvas de nivel a color de las densidades de las cópulas condicionales del modelo explicado en (2.4.3). En cada caso las correlaciones condicionales se encuentra arriba de cada conjunto de curvas de nivel.

Correlaciones condicionales vs No condicionales

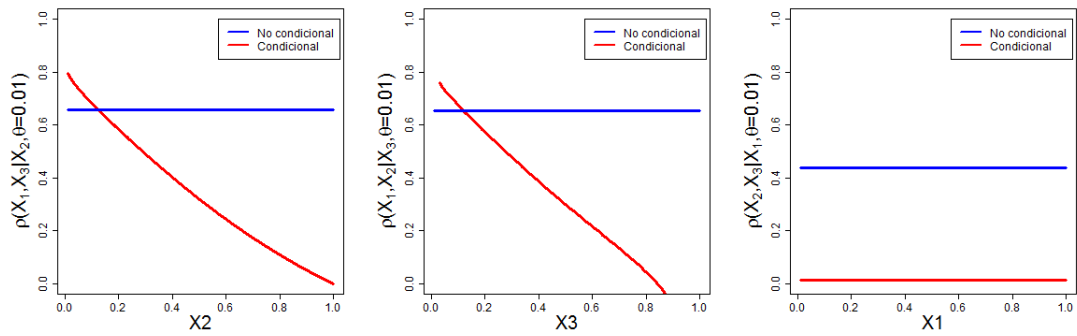


Ilustración 2.21: Comparación entre correlaciones condicionales con $\theta = 0.01$

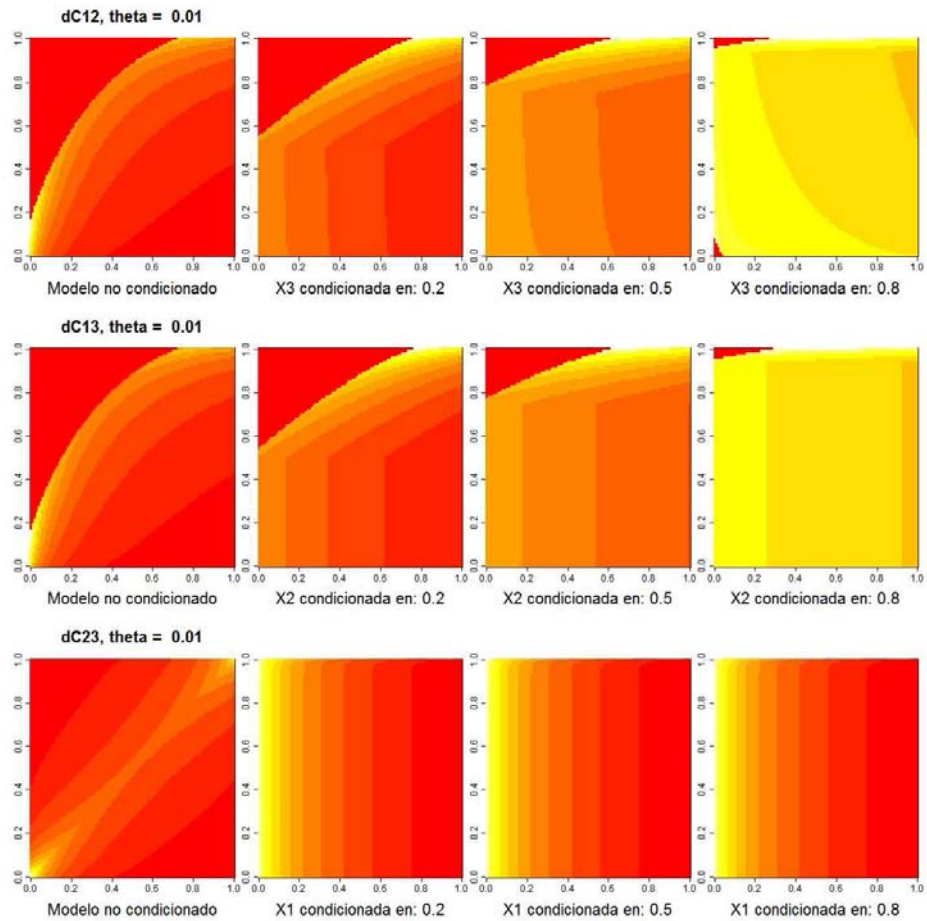


Ilustración 2.22: Comparación entre densidades de cópula con $\theta = 0.01$

Correlaciones condicionales vs No condicionales

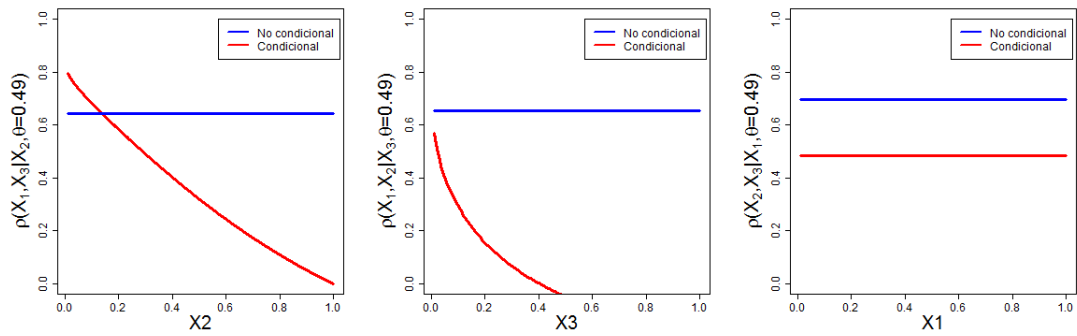


Ilustración 2.23: Comparación entre correlaciones condicionales con $\theta = 0.49$

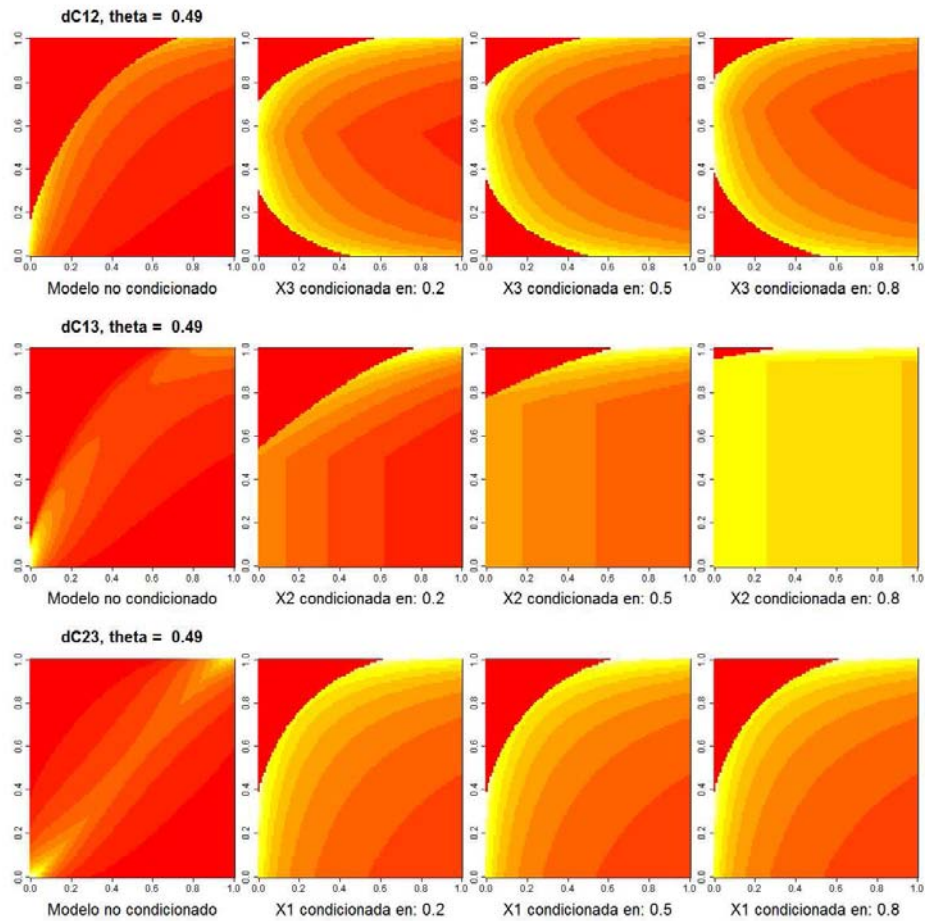


Ilustración 2.24: Comparación entre densidades de cópula con $\theta = 0.49$

Correlaciones condicionales vs No condicionales

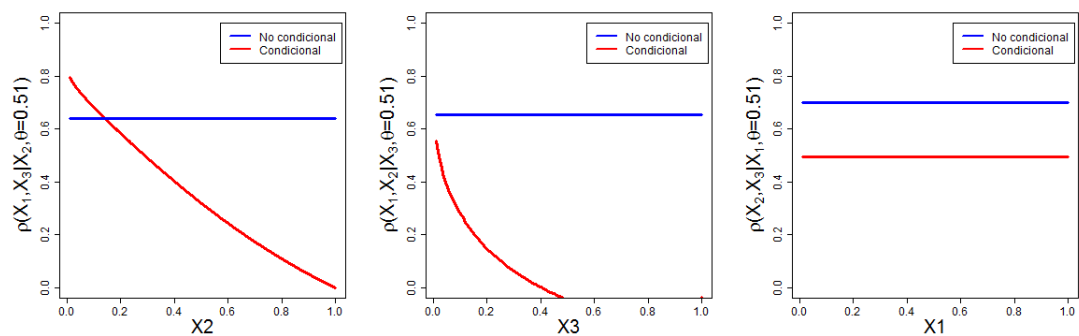


Ilustración 2.25: Comparación entre correlaciones condicionales con $\theta = 0.51$

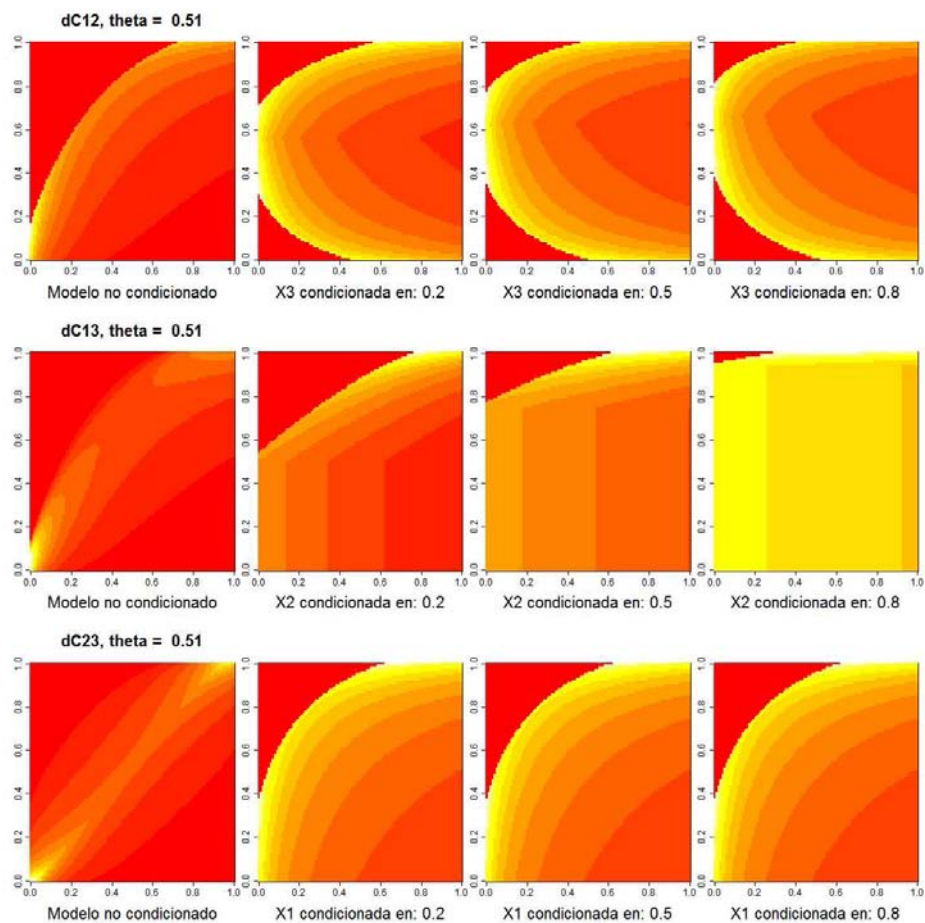


Ilustración 2.26: Comparación entre densidades de cópula con $\theta = 0.51$

Correlaciones condicionales vs No condicionales

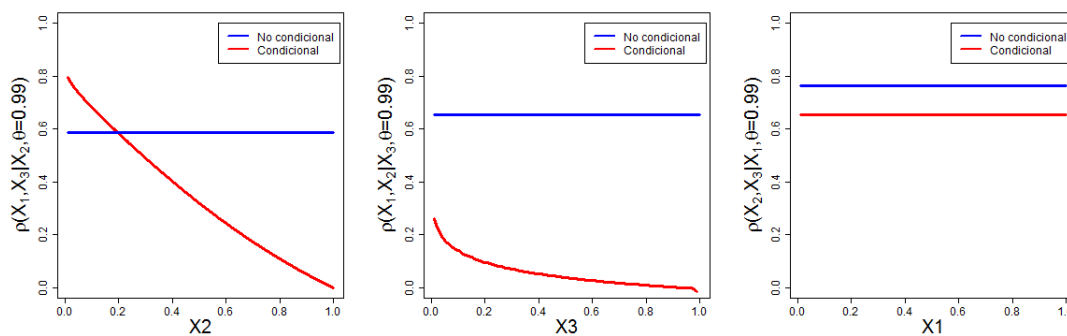


Ilustración 2.27: Comparación entre correlaciones condicionales con $\theta = 0.99$

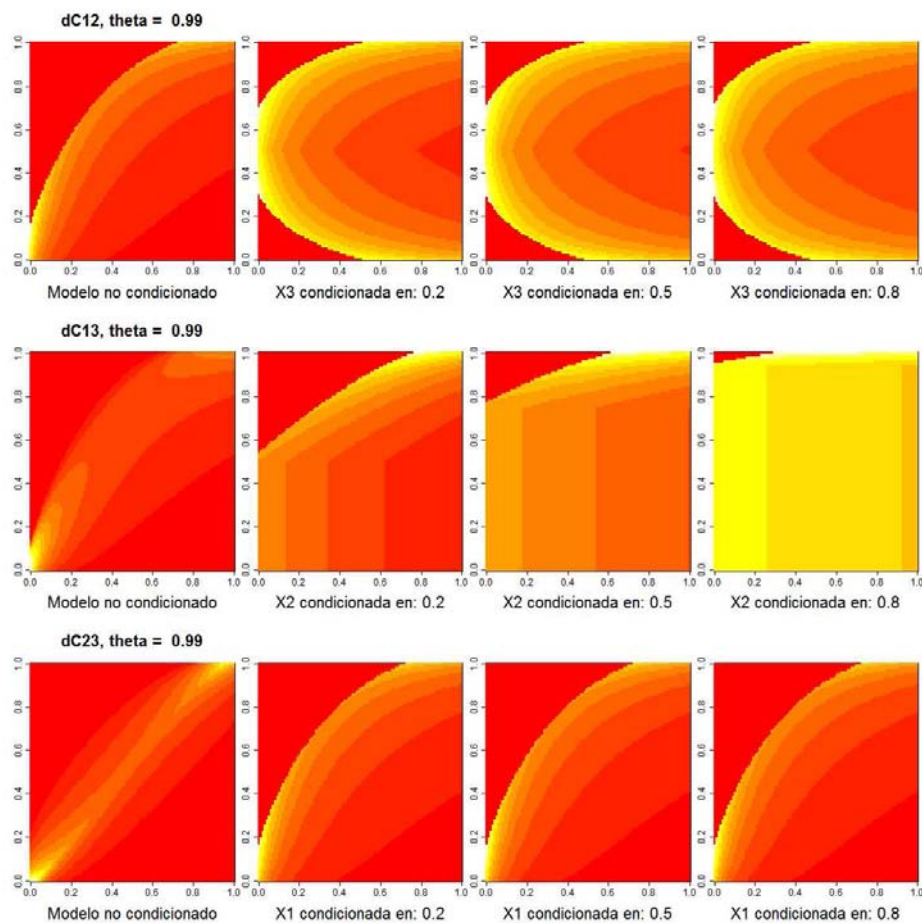


Ilustración 2.28: Comparación entre densidades de cópula con $\theta = 0.99$

Correlaciones condicionales vs No condicionales

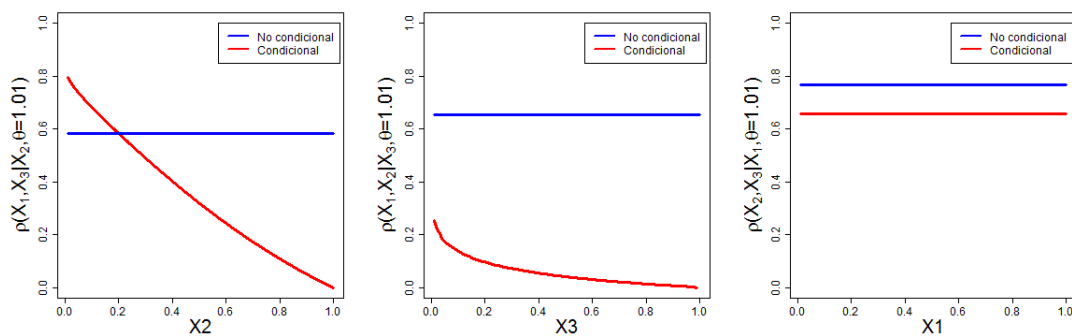


Ilustración 2.29: Comparación entre correlaciones condicionales con $\theta = 1.01$

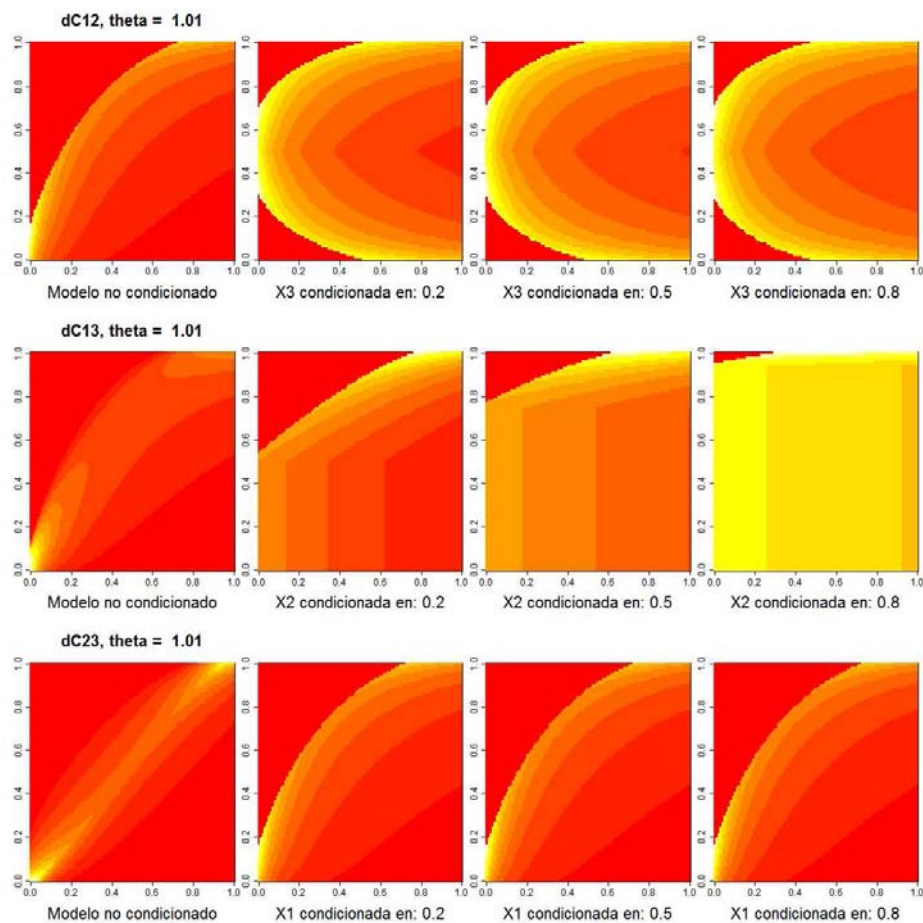


Ilustración 2.30: Comparación entre densidades de cópula con $\theta = 1.01$

Correlaciones condicionales vs No condicionales

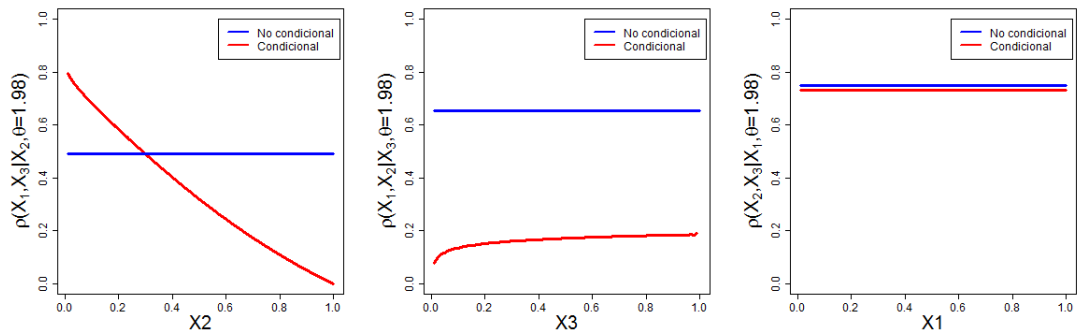


Ilustración 2.31: Comparación entre correlaciones condicionales con con $\theta = 1.98$

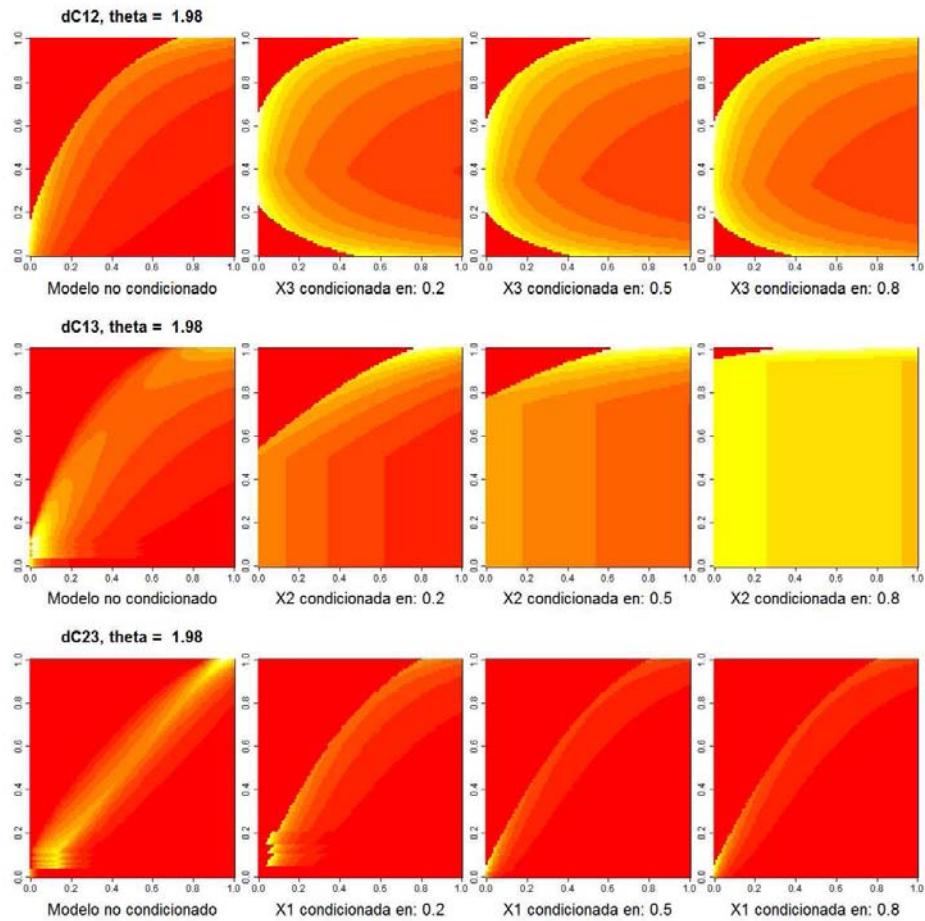


Ilustración 2.32: Comparación entre densidades de cópula con $\theta = 1.98$

Correlaciones condicionales vs No condicionales

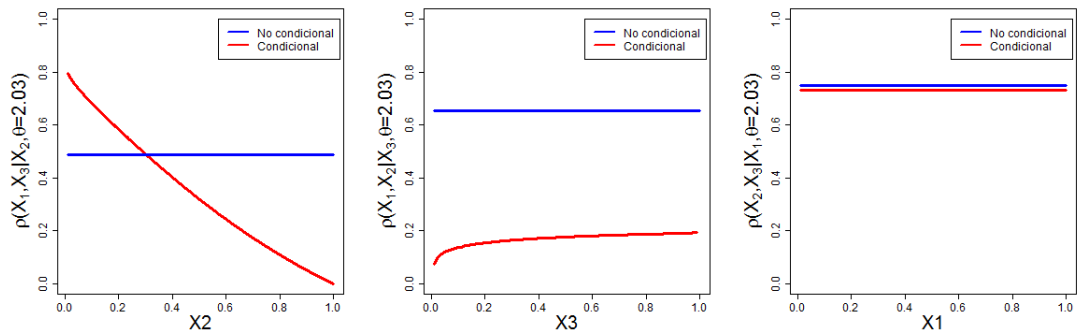


Ilustración 2.33: Comparación entre correlaciones condicionales con $\theta = 2.03$

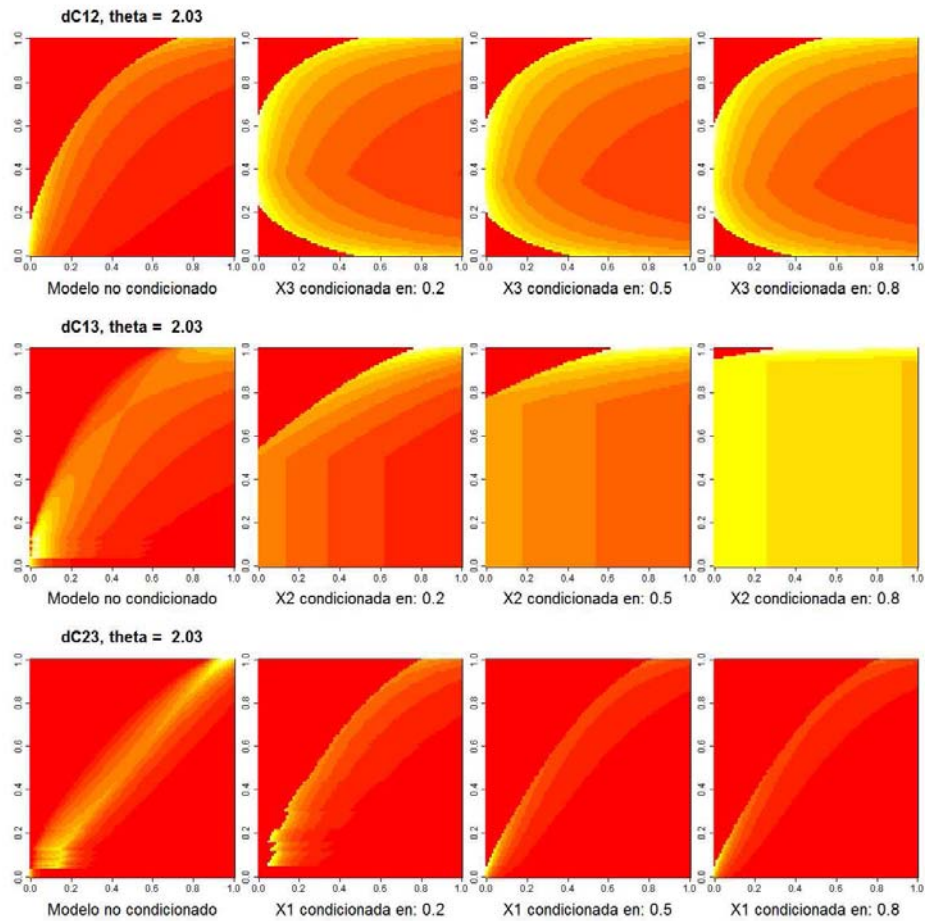


Ilustración 2.34: Comparación entre densidades de cópula con $\theta = 2.03$

En resumen, se aprecia visualmente que para $\theta = 0.01$ la densidad de cópula $\partial C_{23|1}$ es

similar a pesar de que cambie el valor de X_1 . Y para $\theta = \{0.45, 0.51, 0.99, 1.01, 1.98, 2.03\}$ las densidades de cópula $\partial C_{12|3}$ y $\partial C_{23|1}$ permanecerán similares, a pesar de que los valores que toman X_3 y X_1 cambien.

En todos los casos mencionados es posible aplicar el supuesto simplificador si la prueba de Rémillard y Scaillet (2009) valida nuestra apreciación visual.

2.4.5. Prueba de Rémillard y Scaillet

La prueba propuesta por Rémillard y Scaillet (2009) compara conjuntos de datos para determinar que tanto se parecen sus cópulas subyacentes y arroja como resultado un *p-value* que indica la probabilidad de equivocarse si rechazamos que los conjuntos de datos analizados tienen la misma cópula subyacente.

Se aplica la prueba de Rémillard y Scaillet (2009) a las pseudo-observaciones condicionales de la muestra con el objetivo de comprobar que la cópula condicional C_{x_i} no cambia al tomar diferentes valores de x_i .

Para aplicar la prueba de Rémillard y Scaillet (2009) a la muestra, esta se divide en dos partes: la primera correspondiente a las observaciones en que la variable condicional tenga valores menores a su mediana y la segunda con las observaciones en que la variable condicional tenga valores mayores o iguales a su mediana.

Se presentarán las pseudo-observaciones condicionales para cada conjunto de datos simulados:

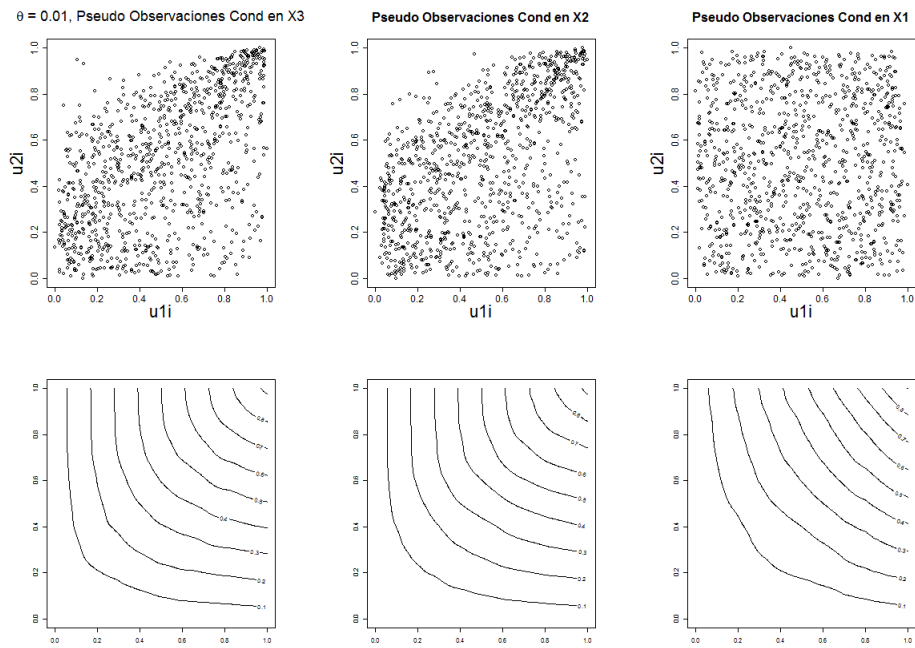


Ilustración 2.35: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.01$

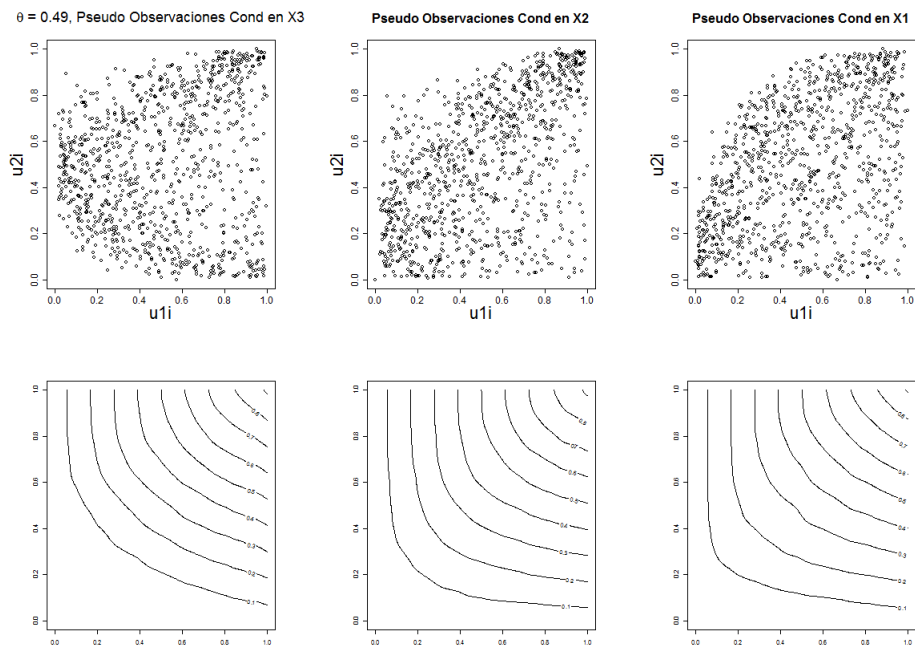


Ilustración 2.36: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.49$

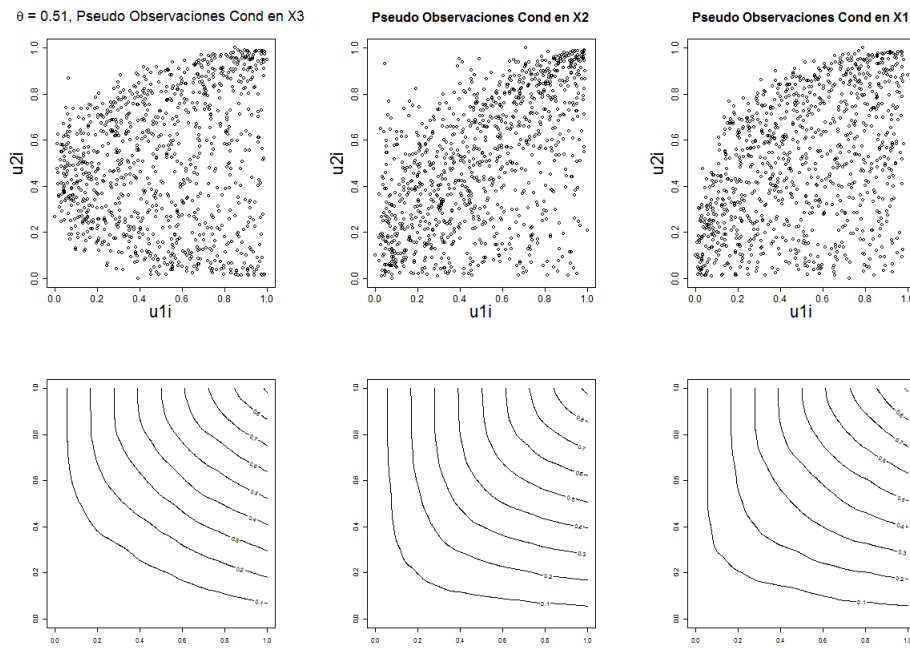


Ilustración 2.37: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.51$

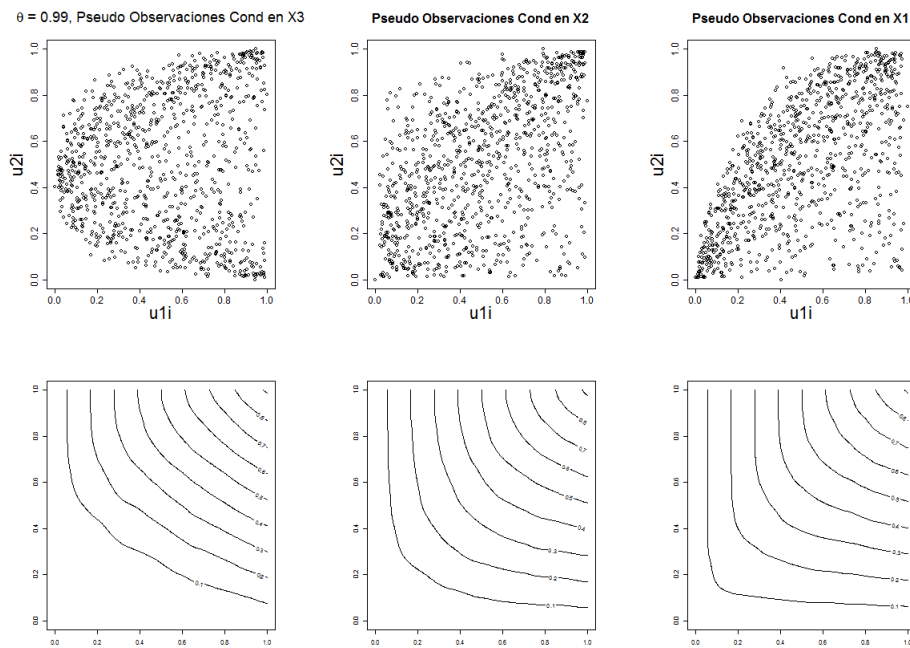


Ilustración 2.38: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 0.99$

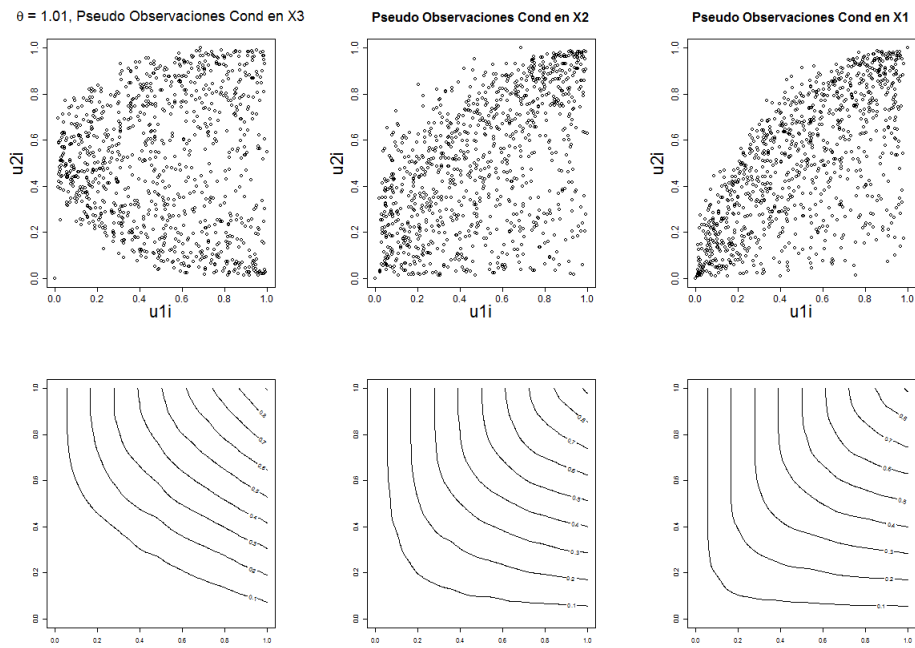


Ilustración 2.39: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 1.01$

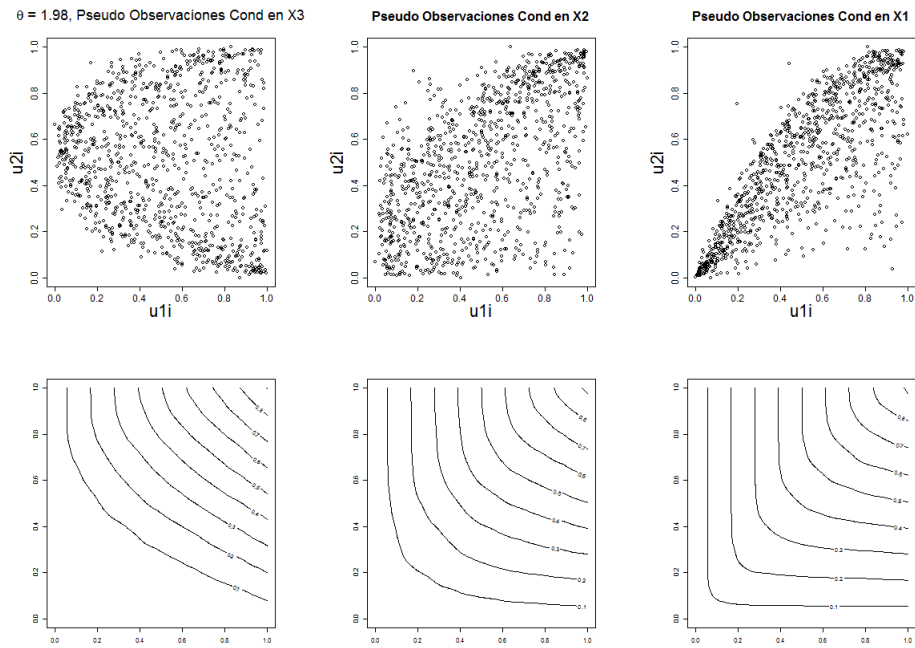


Ilustración 2.40: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 1.98$

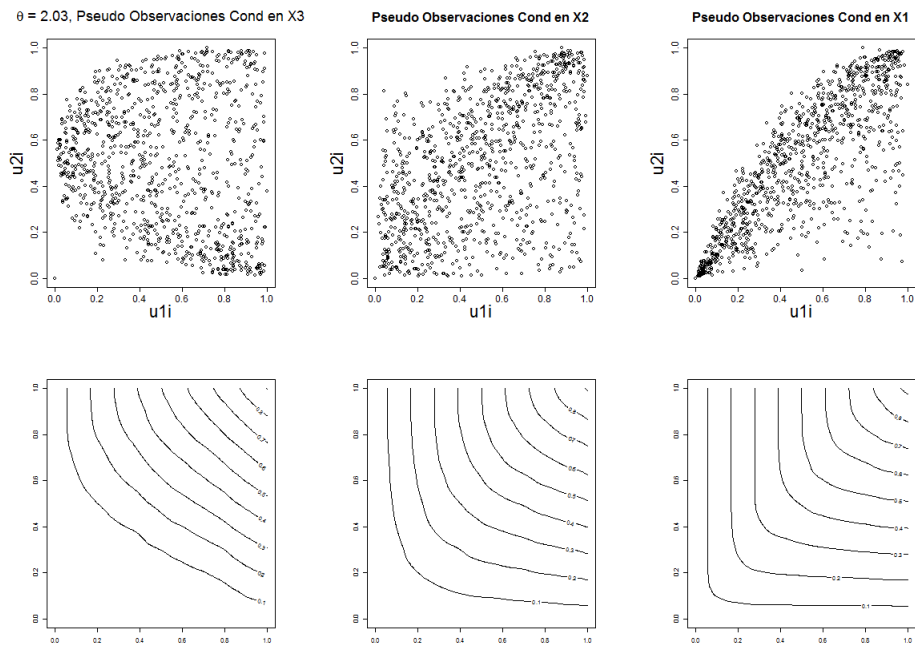


Ilustración 2.41: Pseudo Observaciones Condicionales y Cópulas Bernstein para $\theta = 2.03$

Se mostrarán las pseudo-observaciones condicionales divididas en dos partes de acuerdo a la mediana de la variable en que están condicionadas. El objetivo de esto es visualizar que tan similares son estas partes.

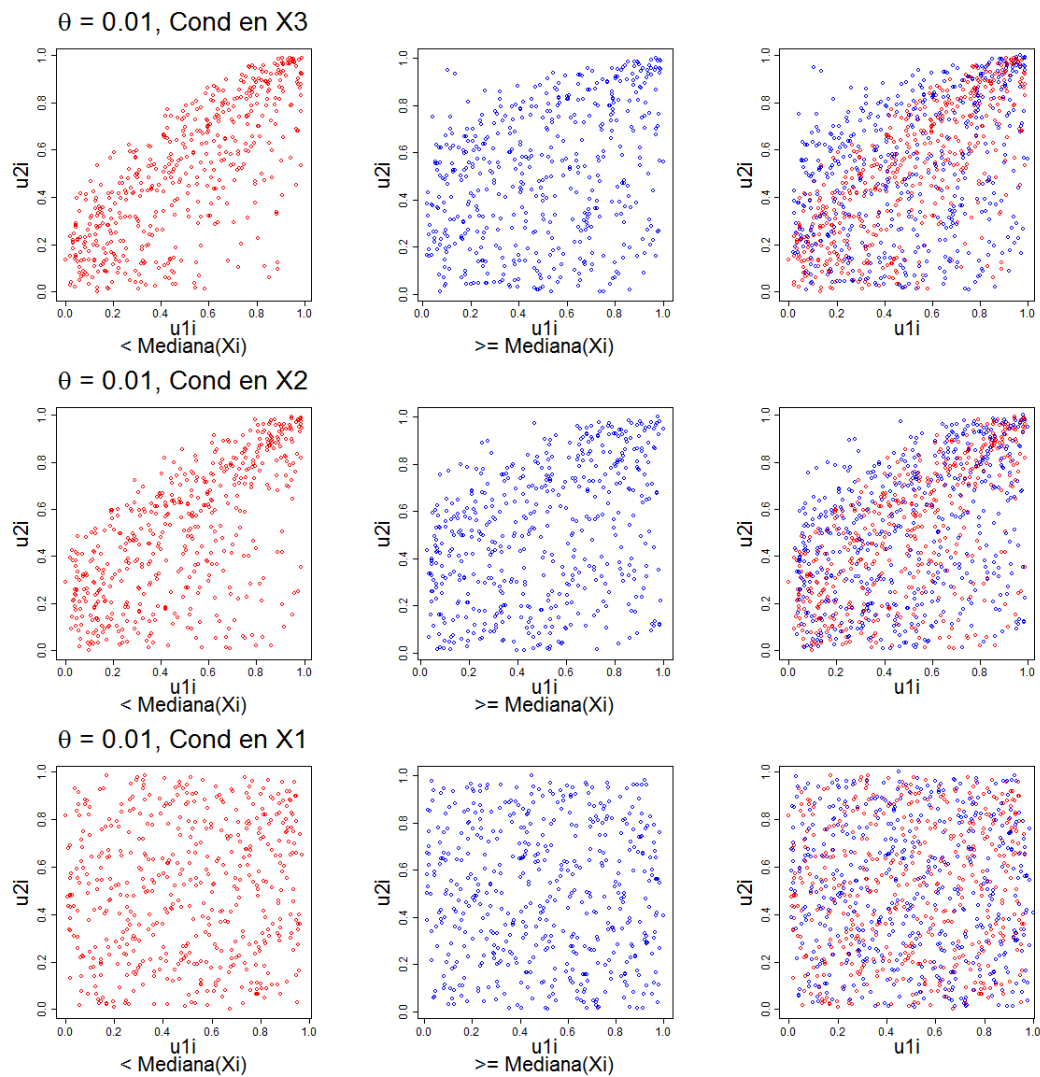


Ilustración 2.42: Pseudo Observaciones Condicionales divididas para $\theta = 0.01$

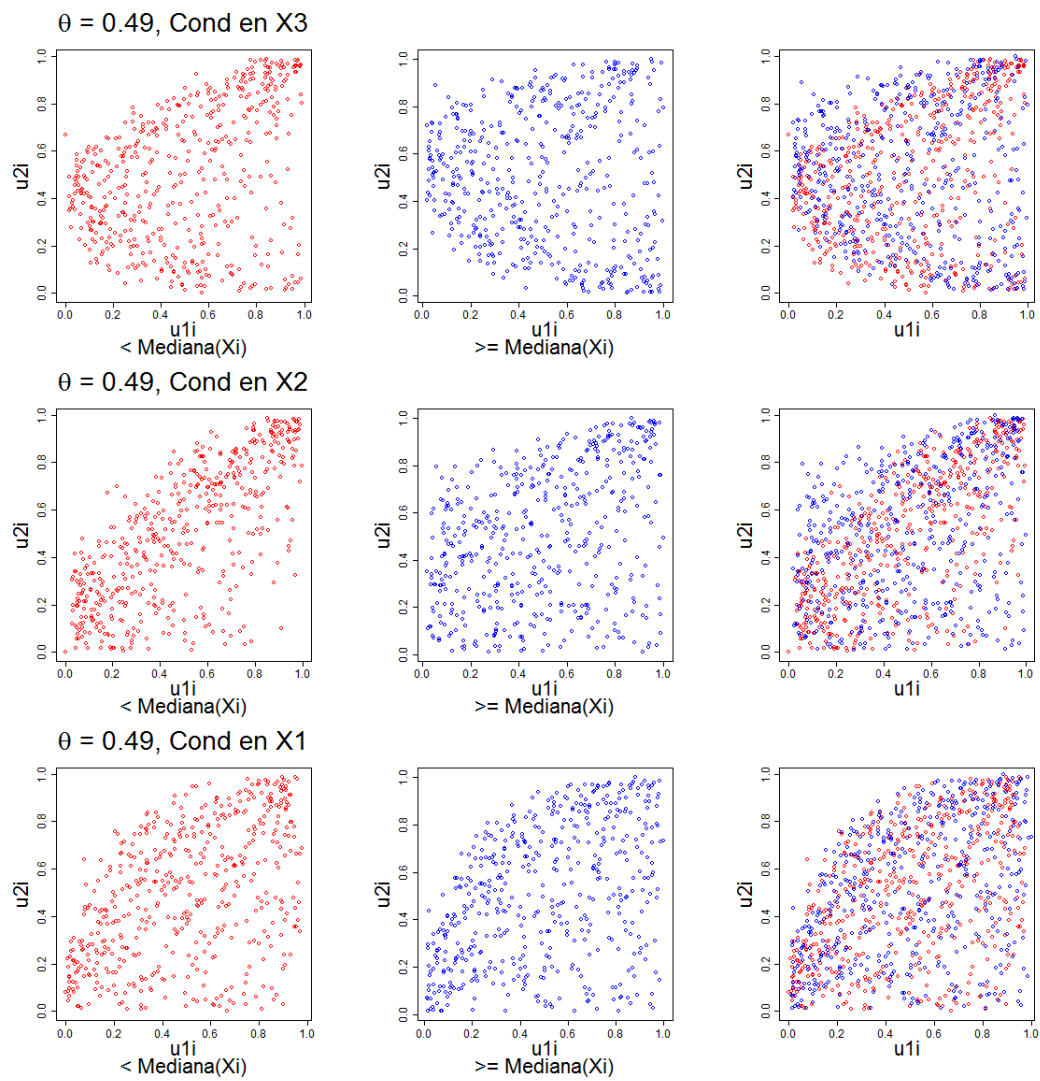


Ilustración 2.43: Pseudo Observaciones Condicionales divididas para $\theta = 0.49$

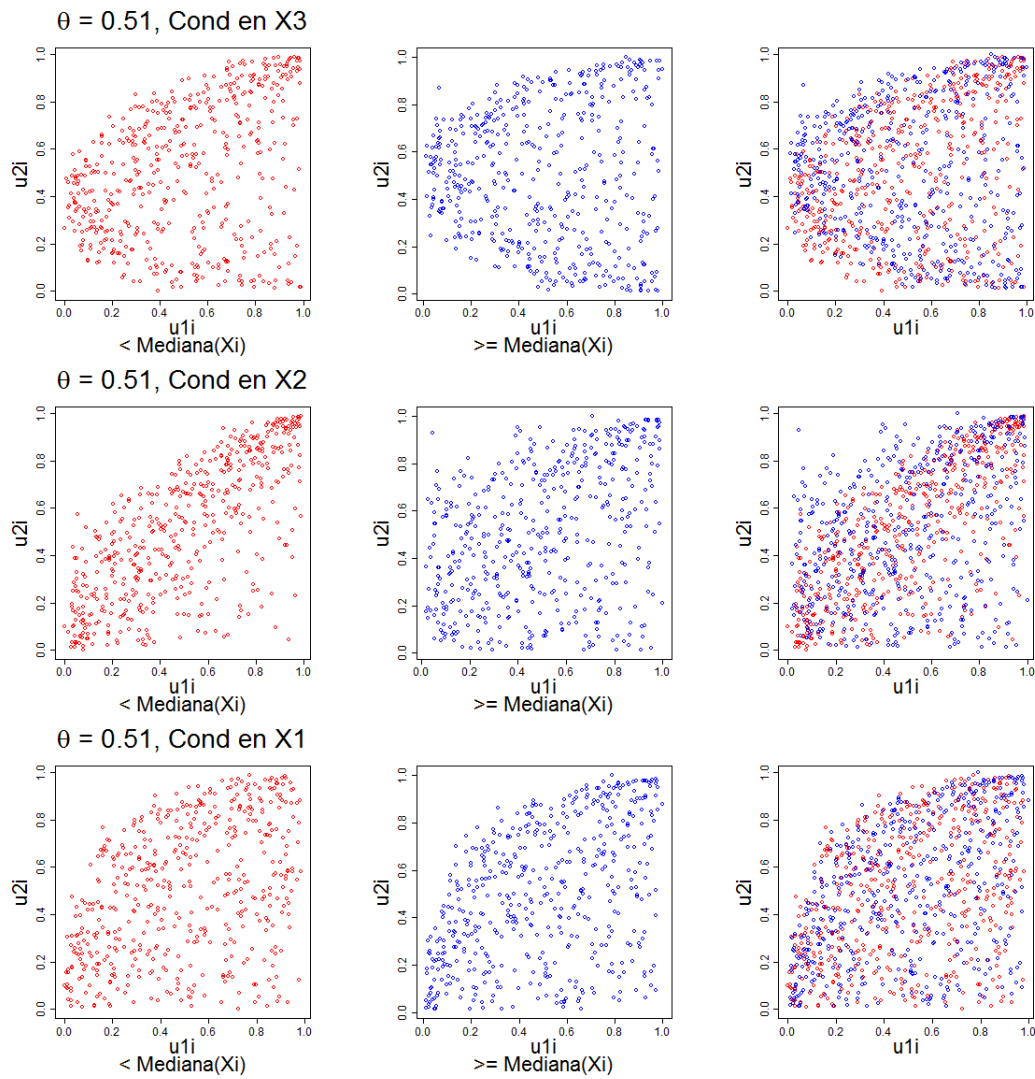


Ilustración 2.44: Pseudo Observaciones Condicionales divididas para $\theta = 0.51$

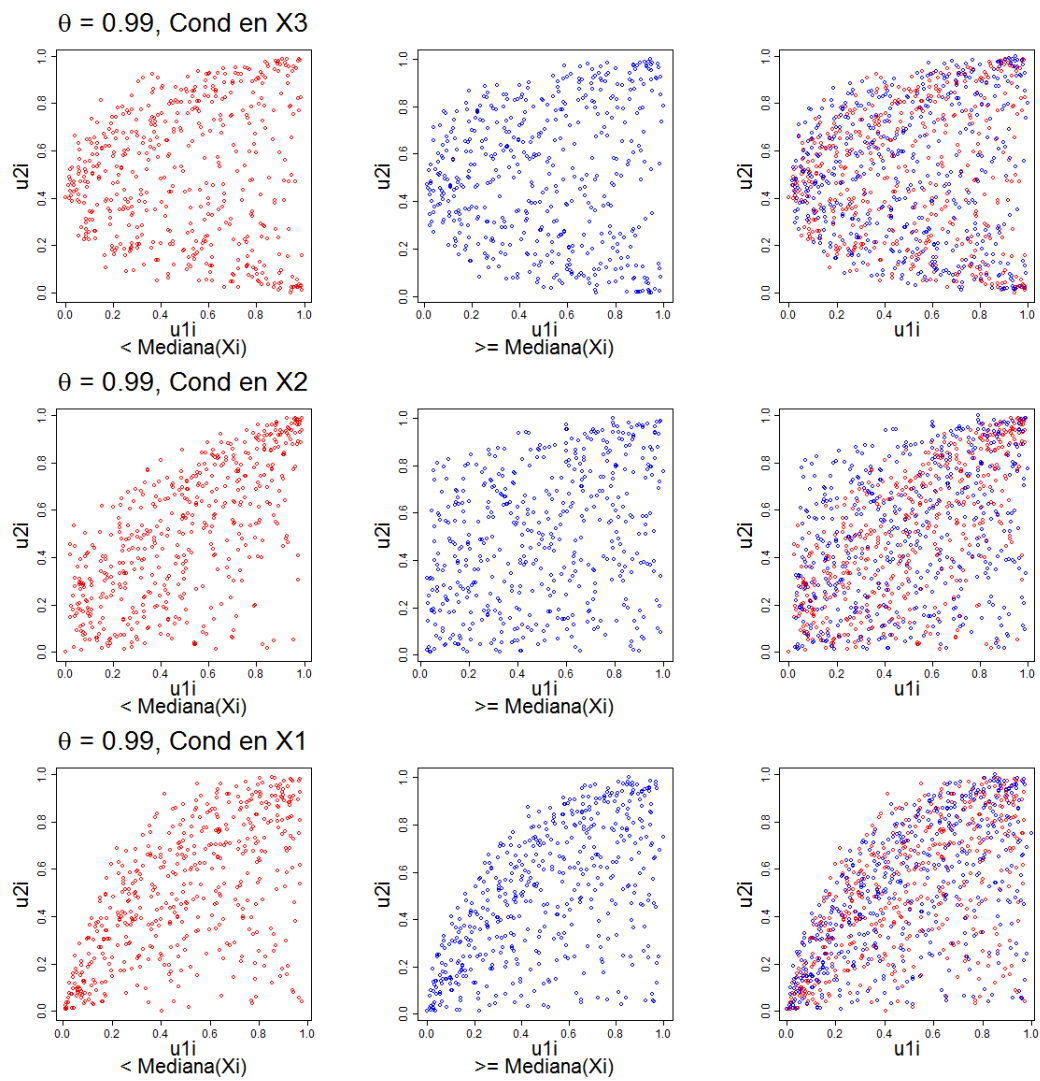


Ilustración 2.45: Pseudo Observaciones Condicionales divididas para $\theta = 0.99$

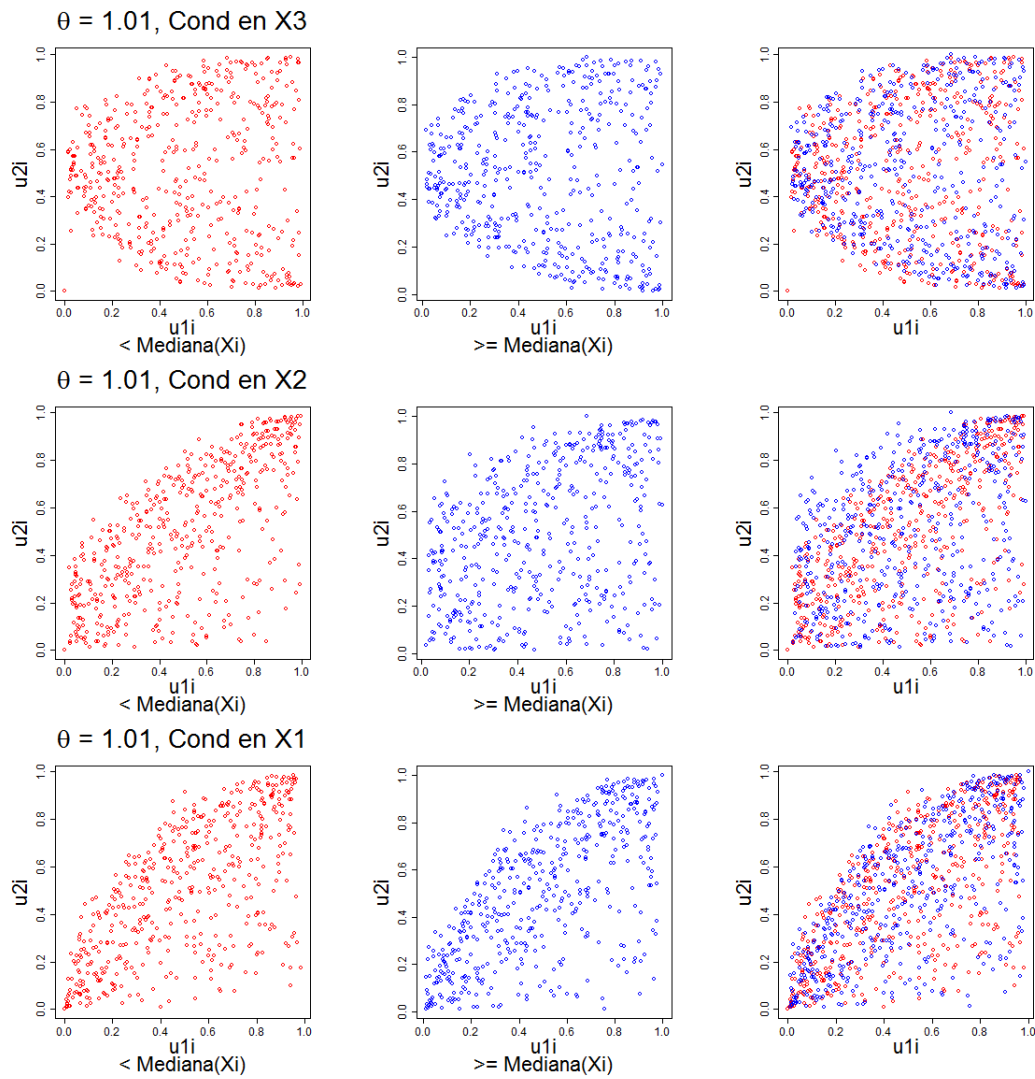


Ilustración 2.46: Pseudo Observaciones Condicionales divididas para $\theta = 1.01$

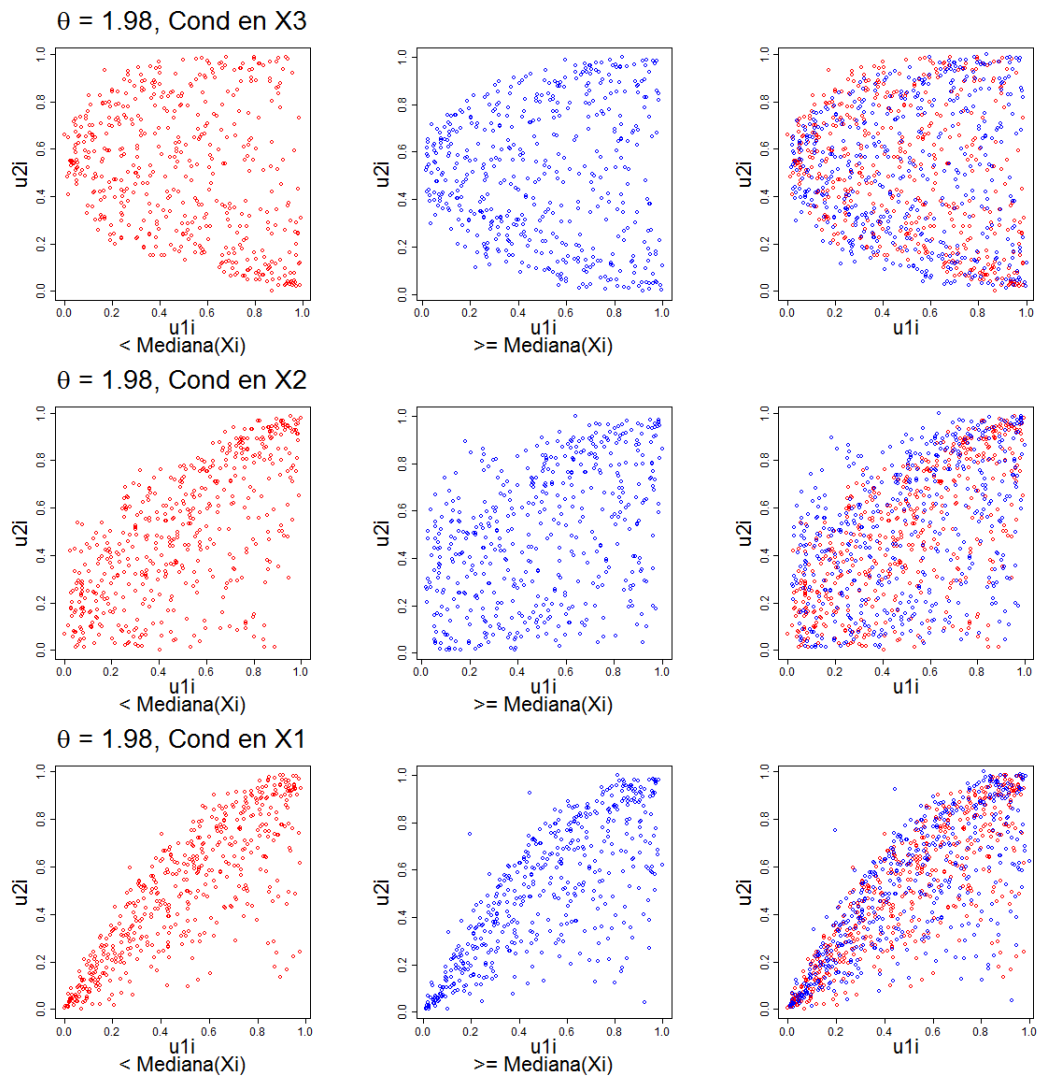


Ilustración 2.47: Pseudo Observaciones Condicionales divididas para $\theta = 1.98$

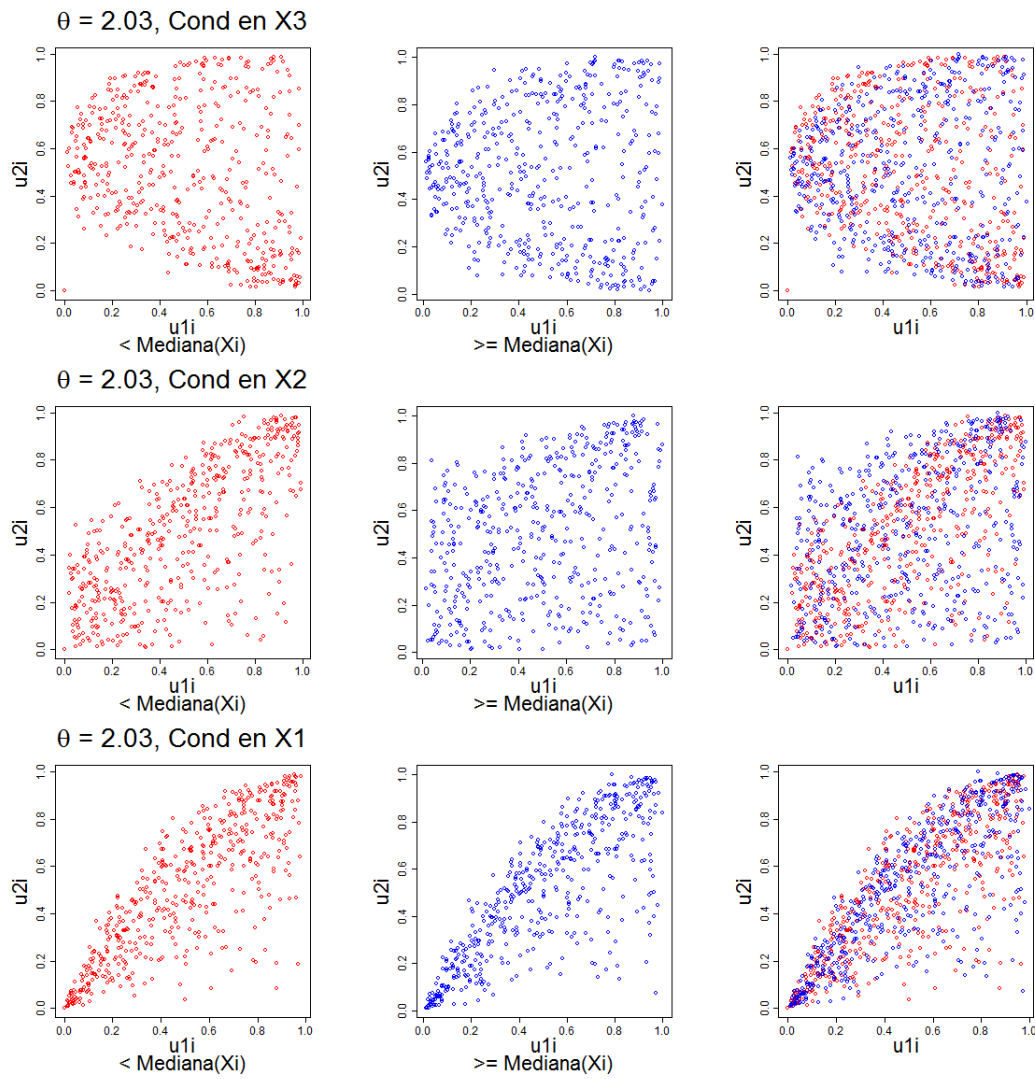


Ilustración 2.48: Pseudo Observaciones Condicionales divididas para $\theta = 2.03$

Los resultados de aplicar la prueba de Rémillard y Scaillet (2009) fueron los siguientes:

Tabla 2.VIII: Prueba de Rémillard y Scaillet, p-values

θ	$C_{12 3}$	$C_{13 2}$	$C_{23 1}$
0.01	0.00	0.00	0.61
0.49	0.00	0.00	0.40
0.51	0.00	0.00	0.44
0.99	0.53	0.00	0.85
1.01	0.93	0.00	0.86
1.98	0.00	0.00	0.18
2.03	0.00	0.00	0.16

En (2.VIII), se puede ver que la prueba de Rémillard y Scaillet (2009) concuerda con la apreciación visual de las densidades de cópula condicional y con las imágenes de las pseudo-observaciones condicionales divididas para el caso $C_{23|1}$. En cambio, para $C_{12|3}$ el modelo solo pasa la prueba cuando $\theta = 0.99, 1.01$.

Aplicación a un caso de estudio

3.1. Análisis descriptivo

Erdely y Díaz-Viera (2010) mencionan que la estimación de la formación de permeabilidad rocosa es un problema complejo y desafiante que juega un rol clave en la modelación de reservas de crudo, pronósticos de producción y manejo óptimo en la explotación de reservas petroleras.

También comentan que las relaciones de dependencia entre variables aleatorias petrofísicas, son usualmente no lineales y complejas y que muchas de las herramientas estadísticas disponibles para tratar el tema asumen supuestos de linealidad y/o de normalidad, y que por dicho motivo no son de mucha ayuda.

Para el presente caso de estudio se emplearon datos petrofísicos analizados previamente en Erdely y Díaz-Viera (2010), los cuales presentan con tres atributos: Porosidad (PHIV), Permeabilidad (K) y Velocidad de propagación de onda (VP).

Mediante el presente caso de estudio se buscó simular valores que conservaran la estructura de dependencia de las observaciones de la muestra original.

Se mostrarán los principales estimadores del conjunto de datos proporcionados:

Estimador	PHIV	VP	K
Mín	0.001937	1.647000	0.000012
Primer cuartil	0.225981	2.643000	0.060986
Mediana	0.630967	2.794000	0.222346
Media	0.649177	2.836000	0.969534
Tercer cuartil	1.016979	2.975000	1.714010
Máx	1.679685	4.471000	4.054170
Máx	1.679685	4.471000	4.054170

Tabla 3.I: Resumen de Estadística descriptiva

A continuación se mostrará un análisis de caja y brazos para cada variable aleatoria:

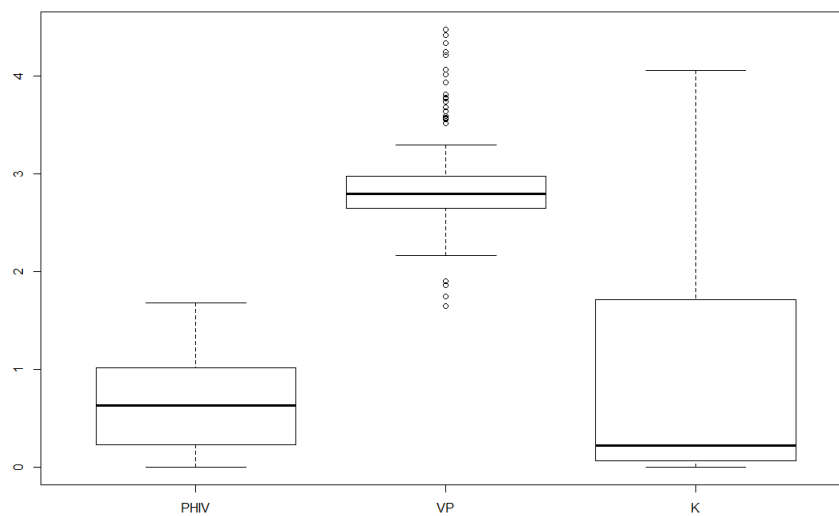


Ilustración 3.1: Diagramas de caja y brazos

En el caso de PHIV, el primer cuartil guarda simetría con el tercero. Es similar el comportamiento de VP, solo que los datos están más cerca de su media. En cuanto a K, la distribución de los datos está sesgada hacia el tercer cuartil.

Se mostrarán algunos diagramas de los datos muestrales y de las pseudo-observaciones en $2D$ y $3D$.

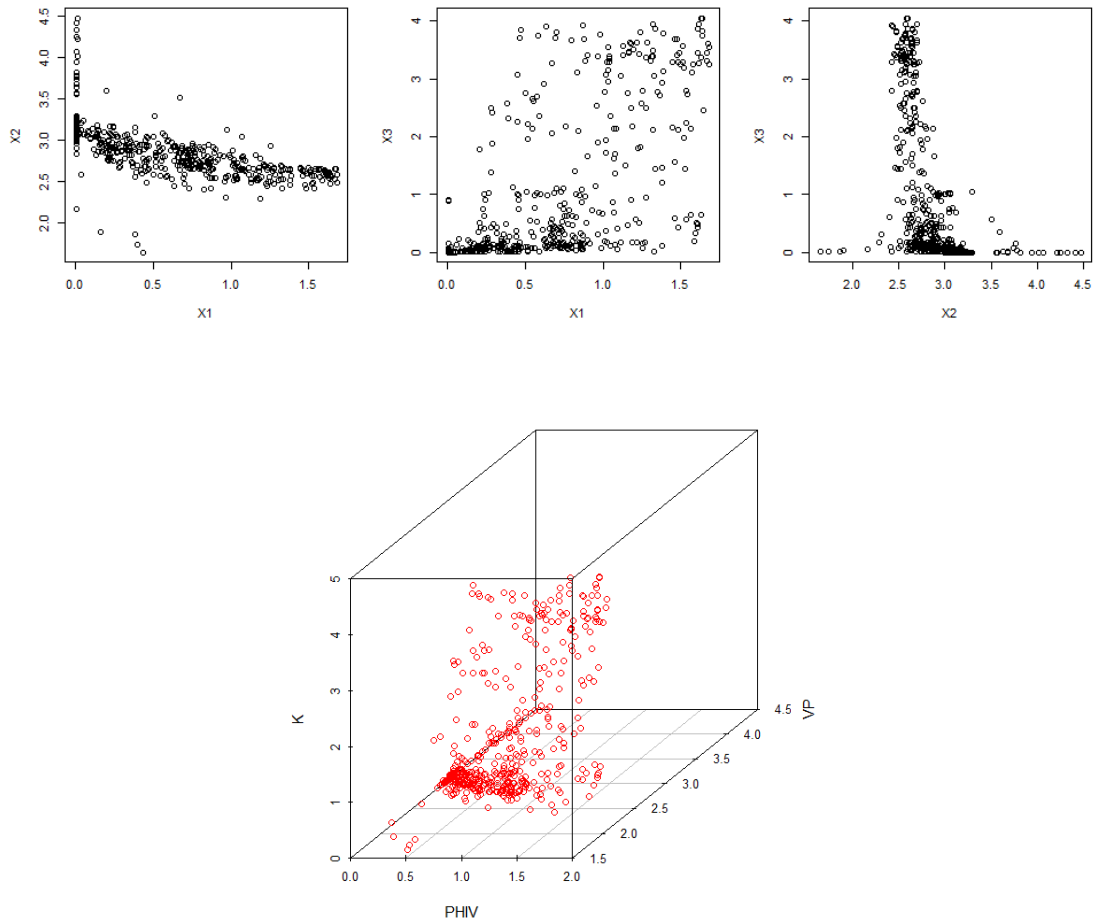


Ilustración 3.2: Diagramas de dispersión de la muestra original en $2D$ y $3D$

Se presentará un histograma por cada variable aleatoria.

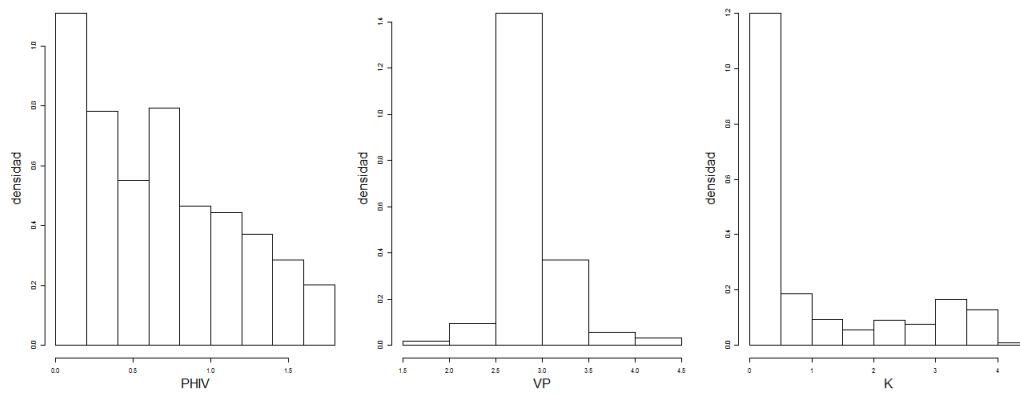


Ilustración 3.3: Histograma de cada variable aleatoria de la muestra original

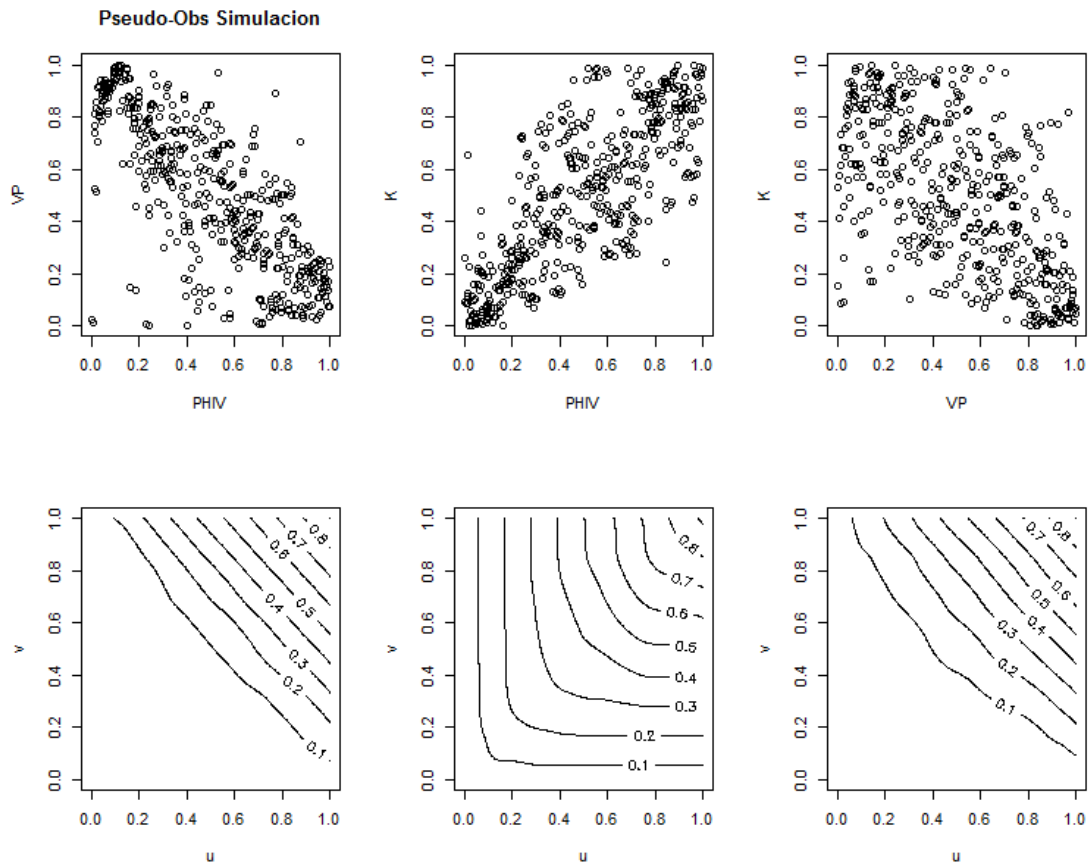


Ilustración 3.4: Pseudo observaciones de la muestra original

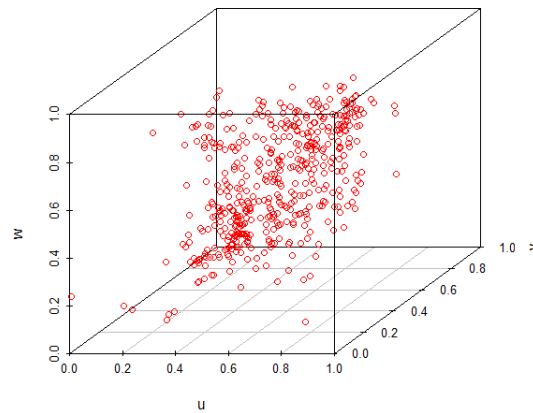


Ilustración 3.5: Pseudo observaciones de la muestra original

En la siguiente tabla se encuentran las medidas de dependencia ρ_{XY} y σ_{XY} aplicadas a los tres pares de pseudo-observaciones:

	ρ_{XY}	σ_{XY}
(X_1, X_2)	-0.7707	0.7686
(X_1, X_3)	0.7455	0.7424
(X_2, X_3)	-0.6885	0.6876

Tabla 3.II: Comparación entre medidas de ρ de Spearman vs σ de Wolff-Schweizer de las pseudo-observaciones

En los tres casos $\sigma_{XY} \approx |\rho_{XY}|$. Para los pares (X_1, X_2) y (X_2, X_3) la relación de dependencia es negativa y para (X_1, X_3) la relación de dependencia es positiva.

Es posible aplicar el supuesto simplificador al conjunto de datos con el que estamos trabajando y así modelar su estructura de dependencia de forma más fácil.

3.2. Supuesto simplificador

Se presentará un diagrama de dispersión de las pseudo-observaciones condicionales y las curvas de nivel de sus respectivas cópulas Bernstein.

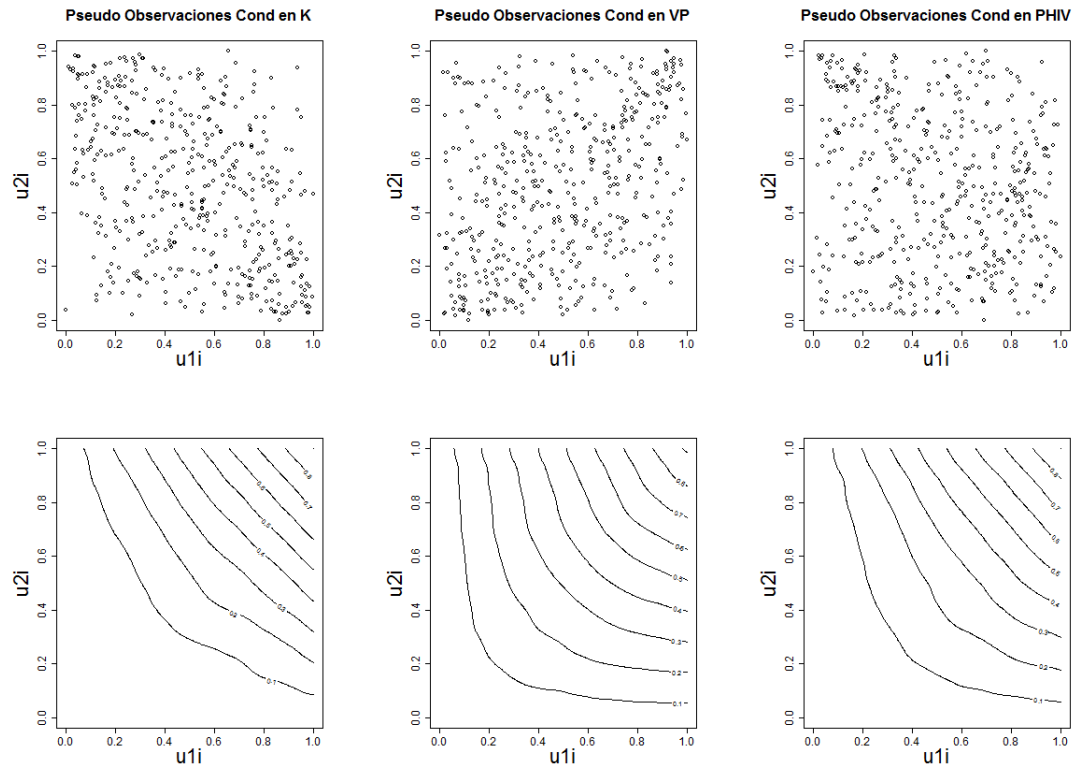


Ilustración 3.6: Pseudo observaciones condicionales

Ahora dividiremos las pseudo-observaciones condicionales en dos partes: la primera que contiene los valores menores a la media de la variable aleatoria condicional y la segunda que contiene los valores mayores o iguales a la media de la variable aleatoria condicional.

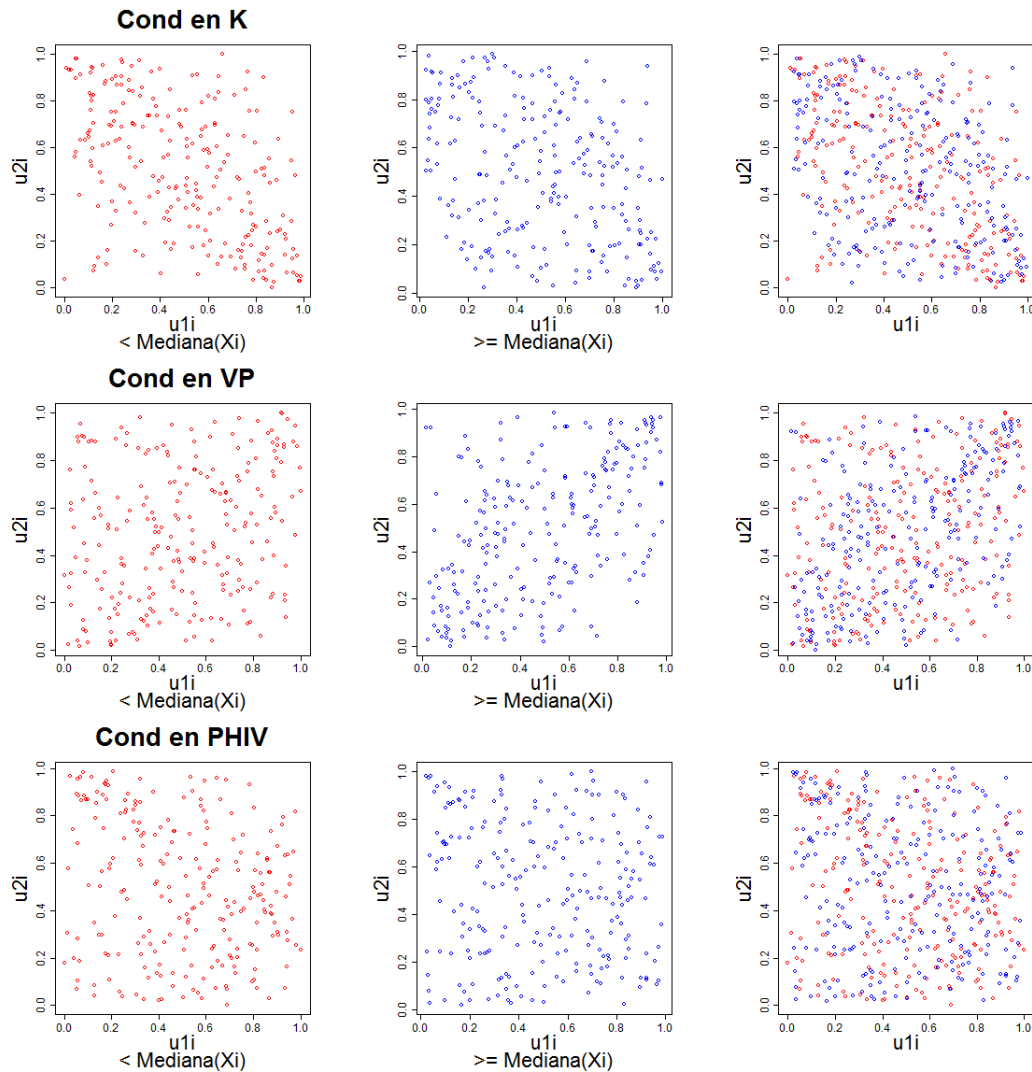


Ilustración 3.7: Pseudo observaciones condicionales clasificadas en dos grupos

Se aplicó la prueba de Rémillard y Scaillet (2009), comparando los datos divididos de las pseudo-observaciones condicionales con el objetivo de conocer el condicionamiento más apto para aplicar el supuesto de simplificación.

Par de variables aleatorias	p-value
$C_{12 3}$	0.06
$C_{13 2}$	0.00
$C_{23 1}$	0.37

Tabla 3.III: *p-values* de pruebas de Rémillard-Scaillet

El *p-value* más alto lo obtuvo la cópula $C_{23|1}$ y por lo tanto se debe aplicar el condicionamiento sobre X_1 .

En la siguiente tabla se encuentran las medidas de dependencia ρ_{XY} y σ_{XY} aplicadas a las pseudo-observaciones condicionadas en X_1 :

	ρ_{XY}	σ_{XY}
$(X_2, X_3 X_1)$	-0.1522	0.1733

Tabla 3.IV: Comparación entre medidas de ρ de Spearman vs σ de Wolff-Schweizer de las pseudo-observaciones condicionales

Se aprecia que σ_{XY} es similar a $|\rho_{XY}|$ y por lo tanto existe un relación negativa entre las pseudo-observaciones condicionales involucradas en el caso $(X_2, X_3 | X_1)$.

3.3. Construcción del modelo trivariado

Se realizó la simulación de las observaciones empleando el algoritmo de la sección (1.4) del capítulo uno y las cópulas Bernstein C_{12} y C_{13} . Se muestran imágenes de las pseudo-observaciones que se simularon, así como de sus respectivas cópulas Bernstein:

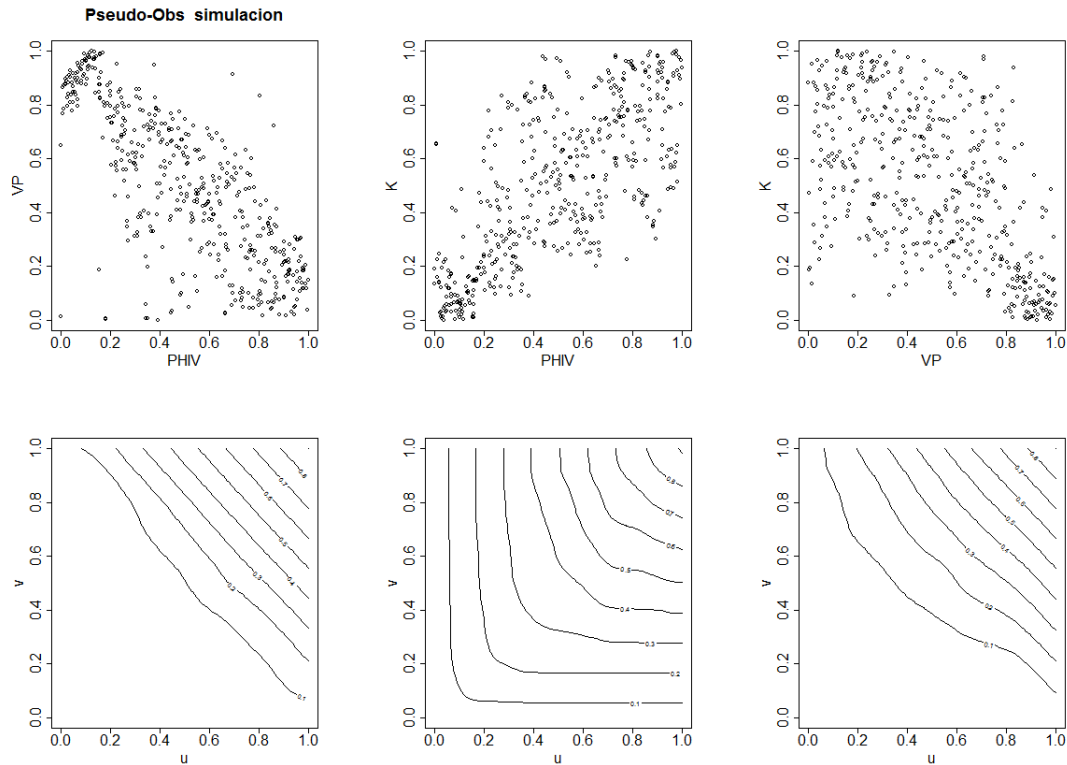


Ilustración 3.8: Pseudo-observaciones de simulación

Se aplicarán las cuasi-inversas de cada variable aleatoria a las pseudo-observaciones obtenidas. Se muestran imágenes de los datos petrofísicos originales y de la simulación realizada:

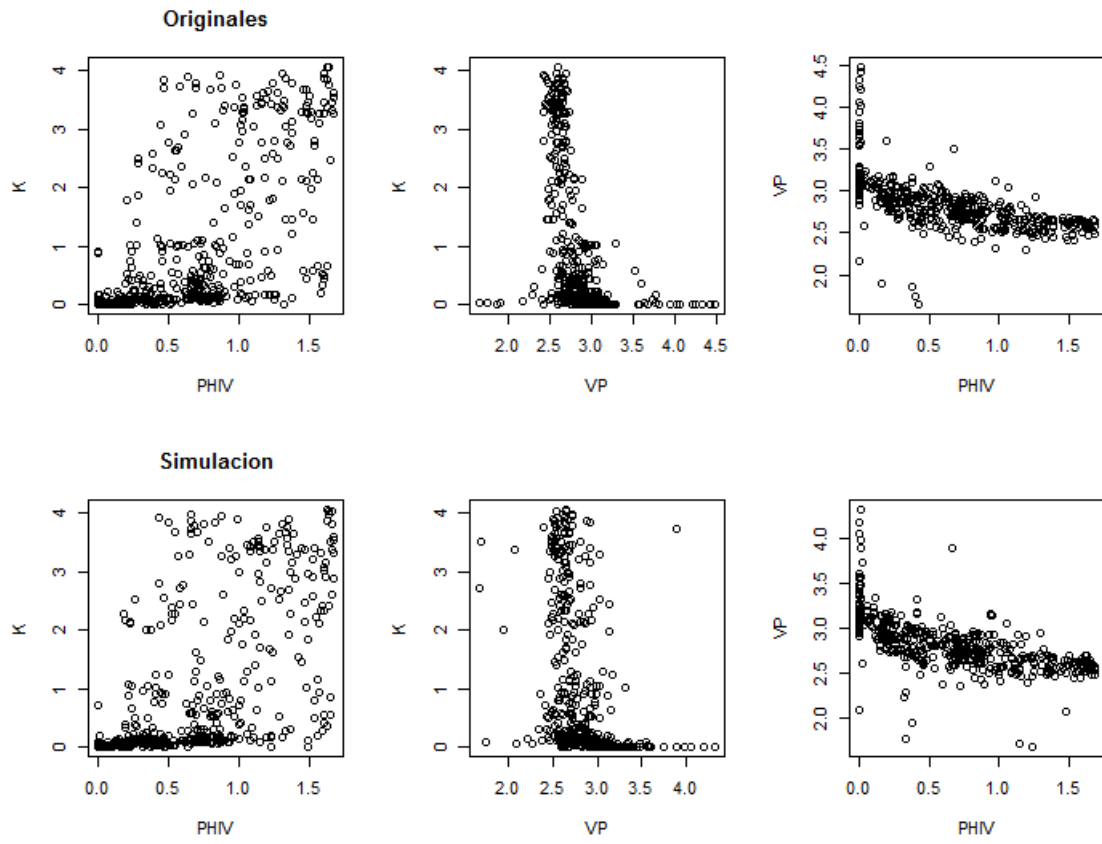


Ilustración 3.9: Datos de muestra originales vs Simulación

Se muestran las pseudo-observaciones condicionales de las observaciones simuladas:

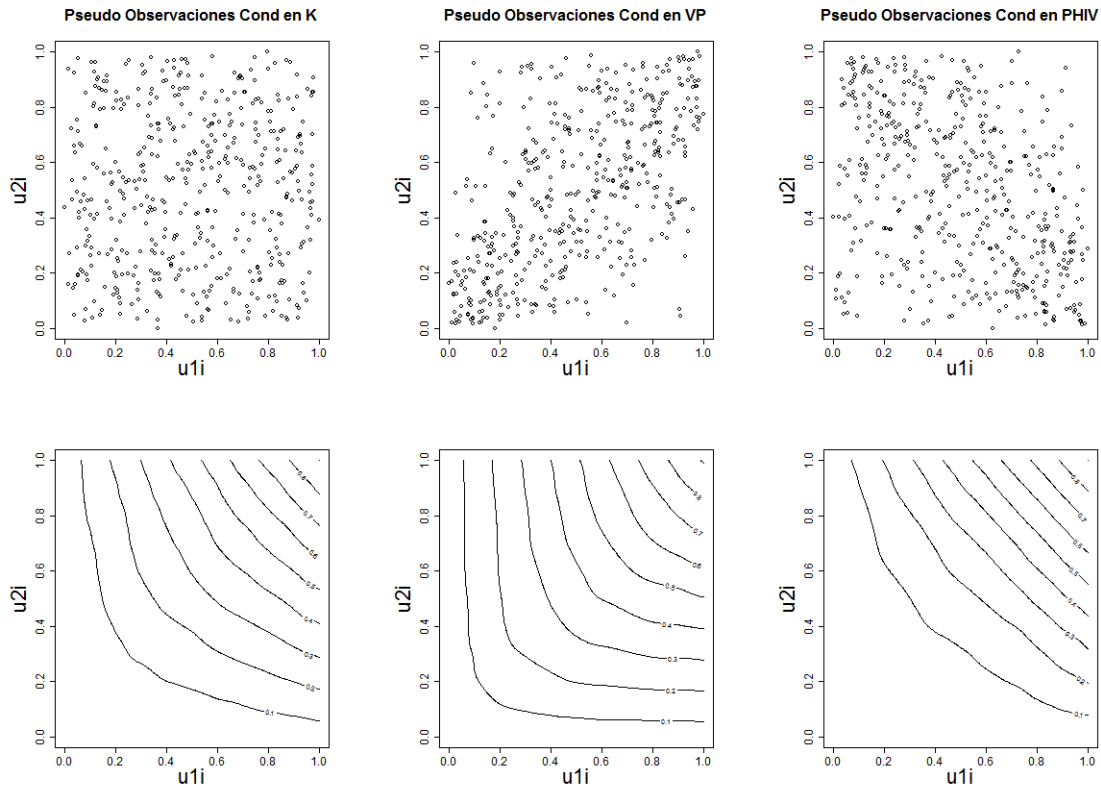


Ilustración 3.10: Pseudo-observaciones condicionales de las observaciones simuladas

Se aplicó la prueba de Rémillard y Scaillet (2009) a las observaciones originales y a la simulación. Los resultados fueron los siguientes:

Cópula	p -value
C_{12}	0.76
C_{23}	0.20
C_{13}	0.97
C_{123}	0.47

Tabla 3.V: p -values de pruebas de Rémillard-Scaillet de comparación datos originales vs simulación

No se rechaza que las observaciones simuladas posean la misma estructura de dependencia que las observaciones de la muestra y por lo tanto se considera la simulación de los

datos petrofísicos un éxito.

Se muestra a continuación imágenes con los diagramas de dispersión de los datos originales, 500 observaciones simuladas mediante el método desarrollado y la simulación de 1000 datos mediante el método de regresión lineal tradicional y con interacción, esto último debido a que al simular las observaciones empleando supuestos de linealidad se generan observaciones negativas.

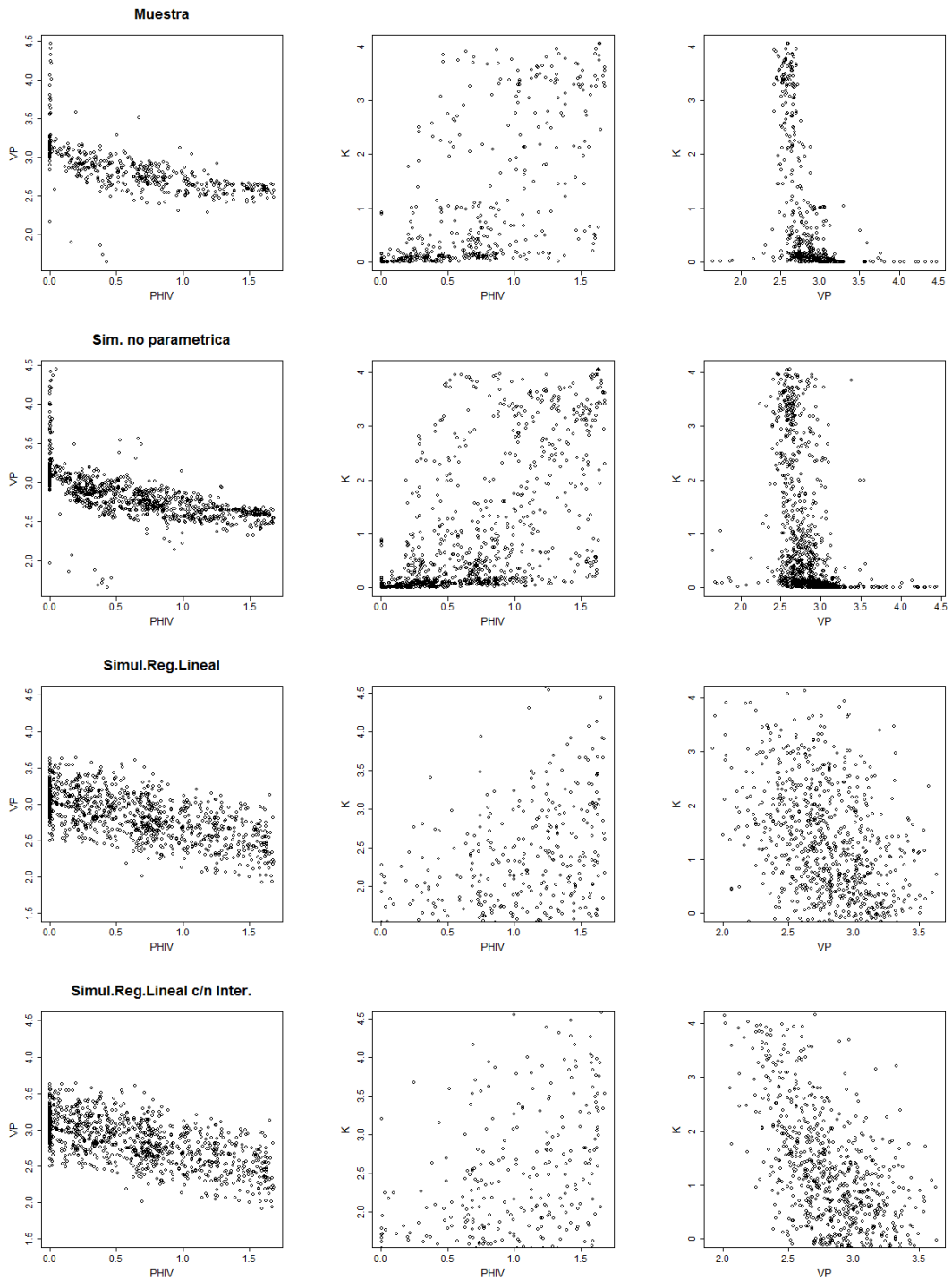


Ilustración 3.11: Datos de muestra originales vs Simulación

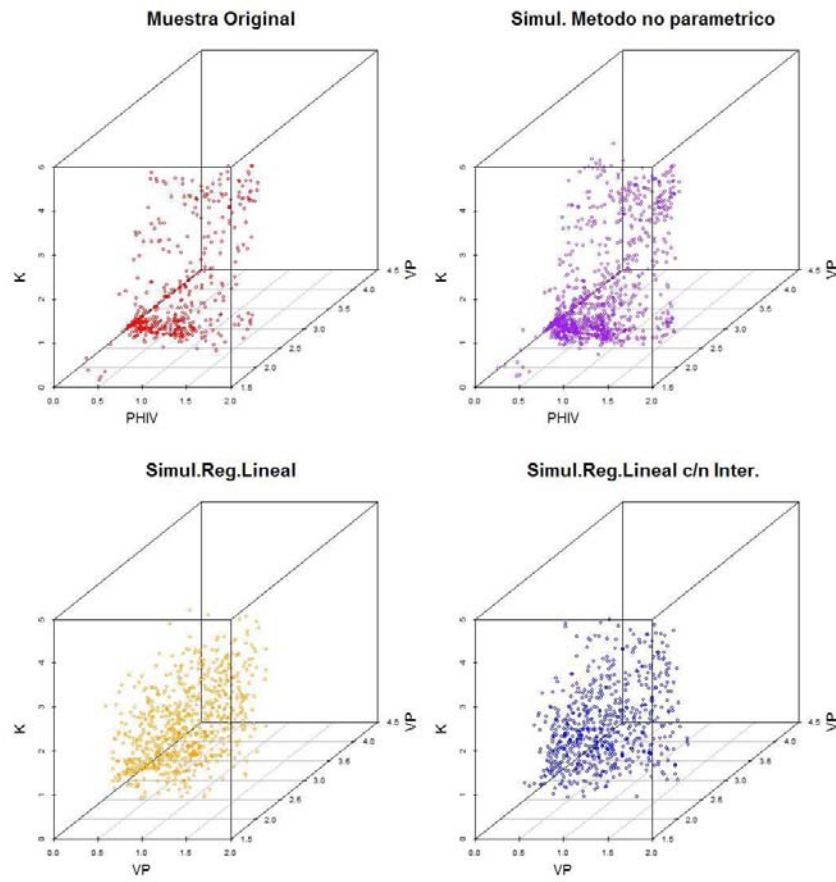


Ilustración 3.12: Datos de muestra originales vs Simulación en 3D

Conclusiones

En el capítulo uno se explicó por qué el coeficiente de correlación lineal no es una medida de dependencia. También se mostró que la sigma de Schweizer y Wolff (1981) es una medida de dependencia y es posible determinar con la correlación de Spearman si un conjunto de datos tiene dependencia positiva o negativa. Adicionalmente se introdujo la función cópula Berstein que permite obtener una aproximación a la cópula subyacente de forma *suavizada*.

En el capítulo dos se demostró que el ajuste de regresión lineal y su varianza describen pobremente el comportamiento de X_1 al ser condicionada en X_2 y X_3 . Además se hizo notar que las variables aleatorias del modelo teórico cuando $\theta = 0$ no cumplen con los supuestos de linealidad y normalidad requeridos, y se constató que en contraste a las densidades de cópula condicional, los coeficientes de correlación condicionales capturan deficientemente la estructura de dependencia. Así mismo, se observa que para una gran cantidad de casos del modelo teórico, las densidades de cópula condicional son muy parecidas, a pesar de que los valores de las variables en las que se condicionaron cambien. Debido a esto y a que pasaron la prueba de Rémillard y Scaillet (2009) se concluyó que es factible aplicarles el supuesto simplificador. Sin embargo, no es posible emplear este supuesto a cualquier conjunto de datos, por lo que no se considera que el enfoque de Cópulas en Racimo sea una metodología de utilización universal.

En el capítulo tres se demostró que los datos petrofísicos eran aptos para ser sometidos a la metodología de Cópulas en Racimo debido a que se logró emplear el supuesto simplificador a uno de sus condicionamientos. Se comprobó que los datos muestrales y los simulados via Cópulas en Racimo poseen una estructura de dependencia muy similar, de

acuerdo a la prueba de Rémillard y Scaillet (2009). Así mismo, al comparar visualmente las observaciones simuladas empleado Cópulas en Racimo contra las obtenidas vía ajuste de regresión lineal con interacción y sin interacción, se apreció que las simulaciones de Cópulas en Racimo poseen observaciones más parecidas a las existentes en la muestra original.

A.1. Código en R

Puede consultar el código en el siguiente link:

<https://www.dropbox.com/s/d6vv00y23970znd/apendice.R?dl=0>

Para el correcto funcionamiento del código se requieren cargar paquetes externos, por lo que se recomienda tener conexión a Internet. Algunas funciones crean imágenes en una carpeta llamada directorio de trabajo, para conocer dónde se encuentra se debe escribir:

```
getwd()
```

Y deberá regresar una ruta como la siguiente:

```
## [1] "C:/Users/USER_NAME/Documents"
```

Donde USER_NAME es el nombre del usuario

A.1.1. Instalación y carga de paquetes

```
install.packages("rgl")
install.packages("fitdistrplus")
install.packages("TwoCop")
install.packages("scatterplot3d")

library(rgl)
library(fitdistrplus)
```

```
library(TwoCop)
library(scatterplot3d)
```

A.1.2. Funciones base creadas por el Dr. Arturo Erdely Ruiz

```
#Función de distribución Bernstein
Fn.inv.Bernshtein<-
function (u, valores.emp)
{
  x <- sort(valores.emp)
  n <- length(x)
  xm <- rep(0, n + 1)
  for (j in 2:n) {
    xm[j] <- (x[j - 1] + x[j])/2
  }
  xm[1] <- x[1]
  xm[n + 1] <- x[n]
  return(sum(xm * dbinom(0:n, n, u)))
}

#datos_capa_2 corresponde a los datos enviados a
#través de la función Fn.Bernshtein
Fn.Bernshtein.aux<-function (u, xdada,datos_capa_2)
  Fn.inv.Bernshtein(u, datos_capa_2) - xdada

#datos_capa_1 corresponde a los datos enviados
#por el usuario
Fn.Bernshtein<-function (x,datos_capa_1){
  contador<<-contador+1
  if(x>min(datos_capa_1) && x<max(datos_capa_1)){
    return(uniroot(Fn.Bernshtein.aux, interval = c(0.00001,0.999999),
      datos_capa_2 = datos_capa_1, xdada = x)$root)
  }
}
```

```

else{
  if (x<=min(datos_capa_1)) return(0)
  else{
    if (x>=max(datos_capa_1)) return(1)
  }
}
}

#Cópula Empírica
genmat.copem<-
function (mat.xy)
{
  n <- dim(mat.xy)[1]
  mat.copem <- matrix(0, ncol = (n + 1), nrow = (n + 1))
  mat.xyord <- mat.xy
  orden <- order(mat.xy[, 1])
  for (i in 1:n) {
    mat.xyord[i, ] <- mat.xy[orden[i], ]
  }
  mat.copem[n + 1, ] <- (0:n)/n
  y.ord <- sort(mat.xyord[, 2])
  for (i in 1:(n - 1)) {
    columna <- (((mat.xyord[, 2][i] <= y.ord)) * 1)/n
    mat.copem[i + 1, ] <- mat.copem[i, ] + c(0, columna)
  }
  return(mat.copem)
}

#Cópula Bernstein
copula.Bernshtein.emp<-
function (u, v)
  sum(matriz.copem *
      (dbinom(0:(dim(matriz.copem)[1] - 1), dim(matriz.copem)[1] -

```

```

1, u) %*% t(dbinom(0:(dim(matriz.copem)[1] - 1),
dim(matriz.copem)[1] - 1, v))))

genmat.copem.Bernshtein<-
function (u.vec, v.vec)
{
  copula.B <- matrix(0, nrow = length(u.vec), ncol = length(v.vec))
  for (i in 1:(length(u.vec))) {
    for (j in 1:(length(v.vec))) {
      copula.B[i, j] <- copula.Bernshtein.emp(u.vec[i],
                                              v.vec[j])
    }
  }
  return(list(u = u.vec, v = v.vec, copemB = copula.B))
}

#Primera derivada parcial de cópula Bernstein
cv.du<-
function (u, v,matriz.copem)
  (dim(matriz.copem)[1] - 1) * sum(matriz.copem *
  ((dbinom(-1:(dim(matriz.copem)[1] - 2),
dim(matriz.copem)[1] - 2, u) - dbinom(0:
(dim(matriz.copem)[1] - 1),dim(matriz.copem)[1] - 2, u) *
c(-1, rep(1, dim(matriz.copem)[1] - 1)))) %*%
t(dbinom(0:(dim(matriz.copem)[1] - 1),
dim(matriz.copem)[1] - 1, v))))

```

A.1.3. Funciones empleadas en el desarrollo de la tesis

```

# Modelo teórico
# con esta sección
# se define lo necesario Para
# instrumentar el modelo teórico

#Funciones que permiten simular las observaciones del modelo teórico
uniforme<-function(max) runif(n=1,min=0,max=max)
gen.vec.unif<-function(tamanyo,theta){
  x1<-x2<-x3<-c()
  x1<-runif(n=tamanyo,min=0,max=1)
  x2<-sapply(x1,uniforme)
  x3<-sapply(x1*(x2^theta),uniforme)
  return(cbind(x1,x2,x3))
}

#Función de densidad trivariada
f.x1.x2.x3<-function(x1,x2,x3,theta)
  (1/x1^2)*(1/x2^theta)*(0<x1)*(x1<1)*
  (0<x2)*(x2<x1)*(0<x3)*(x3<x1*x2^theta)

#Funciones de densidad condicionadas y no condicionadas
#f(x2)
f.x2<-function(x2) -log(x2)*(0<x2)*(x2<1)

#f(x2/x3)
f2d3<-
  function(x2,x3,theta)
    ((x3^(1/theta)<x2)*(x2<1))/(x2^(theta)*f.x3(x3,theta))*
    ((1/x2-1)*(x2>x3^(1/(1+theta)))+(x2^theta/x3-1)*
    (x2<x3^(1/(1+theta))))

```

```

#f(x3/x2)
f3d2<-function(x3,x2,theta) (0<x3)*(x3<x2^theta)/(x2^theta*-log(x2))*
  ((1/x2-1)*(x2^(1+theta)>x3)+(x2^theta/x3-1)*(x2^(1+theta)<=x3))

#f(x3)
f.x3.0<-function(x3) -log(x3)*(x3>0)*(x3<1)
f.x3.1<-function(x3) (log(x3)+2/sqrt(x3)-2)*(x3>0)*(x3<1)
f.x3.t<-function(x3,theta)
  (1-x3^(-theta/(theta+1)))/(theta^2-theta) +
  theta*(x3^(-1+1/theta))/(-theta+1)*
  (1-x3^(-1/(theta^2+theta)))*(0<x3)*(x3<1)
f.x3<-function(x3,theta){
  if(theta<0) return("El valor de theta dado es incorrecto")
  if(theta==0) return(f.x3.0(x3))
  if(theta==1) return(f.x3.1(x3))
  else{return(f.x3.t(x3,theta))}
}

#f(x2 /x1)
f2d1<-function(x2,x1) 1/x1*(0<x2)*(x2<x1)

#f(x3 /x1)
#Para theta = 0
f3d1.theta.0<-function(x3,x1) 1/x1*(0<x3)*(x3<x1)

#Para theta = 1
f3d1.theta.1<-
  function(x3,x1) (2*log(x1) - log(x3))/x1^2 *(0<x3)*(x3<x1^2)

#Para 0<theta != 1
f3d1.theta<-function(x3,x1,theta)

```

```

(x1^(1-theta) - (x3/x1)^(1/theta-1))/
((1-theta)*x1^2)*(0<x3)*(x3<x1^(1+theta))

#Uniendo todos los casos
f3d1<-function(x3,x1,theta) {
  if(theta == 0) return(f3d1.theta.0(x3,x1))
  if(theta == 1) return(f3d1.theta.1(x3,x1))
  if(0<theta & theta!=1) return(f3d1.theta(x3,x1,theta))
}

#f(x1/x3)
#Para theta = 0
f1d3.t.0<-function(x1,x3) (x3<x1)*(x1<1)/(x1*-log(x3))

#Para theta = 1
f1d3.t.1<-function(x1,x3)
  (2*log(x1)-log(x3))/(x1^2*f.x3(x3,theta=1))*(sqrt(x3)<x1)*(x1<1)

#Para 0<theta != 1
f1d3.t<-function(x1,x3,theta)
  (x1^(1-theta)-(x3/x1)^(1/theta-1))/((1-theta)*x1^2*f.x3(x3,theta))*
  (x3^(1/(1+theta))<x1)*(x1<1)

#Uniendo todos los casos
f1d3<-function(x1,x3,theta){
  if(theta==0) return(f1d3.t.0(x1,x3))
  if(theta==1) return(f1d3.t.1(x1,x3))
  if(0<theta & theta!=1) return(f1d3.t(x1,x3,theta))
}

#f(x1/x2)
f1d2<-function(x1,x2) (x2<x1)*(x1<1)/(x1*-log(x2))

```

```

#funciones de densidad bivariada condicionadas y no condicionadas
#f(x1,x2)
f.1.2<-function(x1,x2)
  return(1/x1*(0<x2)*(x2<x1))

#f(x1,x3)
#Para theta = 0
f.1.3.0<-function(x1,x3) 1/x1*(0<x3)*(x3<x1)

#Para theta = 1
f.1.3.1<-function(x1,x3) (2*log(x1)-log(x3))/x1^2*(sqrt(x3)<x1)*(x1<1)

#Para 0<theta != 1
f.1.3.tn<-function(x1,x3,theta){
  return((1/(x1^2))*((x1^(-theta+1)))/
  (-theta+1)-(x3/x1)^((-theta+1)/theta)/(-theta+1))*
  (x3^(1/(theta+1))<x1)*(x1<1)*(x3>0)*(x3<1))
}

#Uniendo todos los casos
f.1.3<-function(x1,x3,theta){
  resultado<-0
  if (theta<0) return("Error, no se admiten valores menores a 0")
  if(theta==0) return(f.1.3.0(x1,x3))
  if (theta==1) return(f.1.3.1(x1,x3))
  else {return(f.1.3.tn(x1,x3,theta))}
}

#f(x2,x3)
f.2.3<-function(x2,x3,theta){
  return(1/x2^theta*(-1+(1/max((x3/x2^theta),x2)))*
  (0<x2)*(x2<1)*(0<x3)*(x3<x2^theta))
}

```



```

#f(x1 x2/x3)
#Para theta = 0
f.1.2d3.0<-function(x1,x2,x3) -((max(x2,x3)<x1)*(x1<1))/(x1^2*log(x3))

#Para theta = 1
f.1.2d3.1<-
  function(x1,x2,x3)
  ((0<x3)*(x3<(x1*x2))*(0<x2)*(x2<x1)*(0<x1)*(x1<1))/
  (x1^2*x2*(log(x3)+2/sqrt(x3)-2))

#Para 0<theta != 1
f.1.2d3.t<-
  function(x1,x2,x3,theta)((x3/x1)^(1/theta)<x2)*(x2<x1)*
  (x3^(1/(1+theta))<x1)*(x1<1))/(x1^2*x2^theta*
  ((theta*x3^(1/theta-1)*(1-x3^(-1/(theta^2+theta))))/(1-theta)+
  (1-x3^(-theta/(theta+1)))/(theta^2-theta)))

#Uniendo todos los casos
f.1.2d3<-function(x1,x2,x3,theta){
  if(theta<0) return("Error, no se admiten valores menores a 0")
  if(theta==0) return(f.1.2d3.0(x1,x2,x3))
  if(theta==1) return(f.1.2d3.1(x1,x2,x3))
  else {return(f.1.2d3.t(x1,x2,x3,theta))}
}

#f(x2, x3 |x1)
f.2.3d1<-function(x2,x3,x1,theta)
  1/(x1^2*x2^theta)*(0<x3)*(x3<x1*x2^theta)*(0<x2)*(x2<x1)

#f(x1, x3 |x2)
f.1.3d2<-function(x1,x3,x2,theta)
  1/(-log(x2)*x1^2*x2^theta)*(x2<x1)*(x1<1)*(0<x3)*(x3<x2^theta*x1)

```

```

#Funciones de distribución F2,F3
F2<-function(t) t-t*log(t)*(0<t)*(t<1)

#Para theta = 0
F3.theta.0<-function(t) t-t*log(t)*(0<t)*(t<1)

#Para theta = 1
F3.theta.1<-function(t) t* log(t) + 4*t^(1/2) - 3*t

#Para 0<theta != 1
F3.theta<-
  function(t,theta) ((t^(theta/(theta+1))+1/theta)-t)*
  theta^3-t*theta^2+t*theta-t^(theta/(theta+1)+1+t)/
  (t^(theta/(theta+1))*(1-theta)*theta)

#Uniendo todos los casos
F3<-function(t,theta){

  if (theta==0) return(F3.theta.0(t))
  if (theta==1) return(F3.theta.1(t))
  if (theta>0 & theta!=1) return(F3.theta(t,theta))
}

#F1d3
#Para theta=0
F1d3.0<-function(t,x3) -1/log(x3) *(log(t)-log(x3))*(x3<t)*(t<1)

#Para theta=1
F1d3.1<-function(t,x3) 1/f.x3(x3)*(2*((-log(t)-1)/t+
(log(sqrt(x3))+1)/sqrt(x3))-log(x3)*(-1/t+1/sqrt(x3)))*
(sqrt(x3)<t)*(t<1)

```

```

#Para 0<theta!=1
F1d3.theta<-function(t,x3,theta) 1/((1-theta)*f.x3(x3,theta))*
  (-t^(-theta)/theta+theta*x3^(1/theta-1)*t^
  (-1/theta)+x3^(-theta/(1+theta))/theta-theta*x3^(1/theta-1)*
  x3^(-1/(theta*(1+theta))))*(x3^(1/(1+theta))<t)*(t<1)

#Uniendo todos los casos
F1d3<-function(t,x3,theta){
  if (theta==0) return(F1d3.0(t,x3))
  else if(theta==1) return(F1d3.1(t,x3))
  else if(theta>0 & theta!=1) return(F1d3.theta(t,x3,theta))
}

#Función de distribución para (x2/x1)
F2d1<-function(t,x1) t/x1 * (0<t) * (t<x1) * (0<x1) * (x1<1)

#Función de distribución para (x3/x1)
F3d1.theta.0<-function(t,x1) t/x1*(0<t)*(t<x1)*(0<x1)*(x1<1)
F3d1.theta.1<-function(t,x1) (2*log(x1)/x1^2*t -
  t *(log(t)-1)/x1^2)*(0<t)*(t<x1^2)
F3d1.theta<-
  function(t,x1,theta) (x1^(1-theta)*t-theta*x1*(t/x1)^(1/theta))/
  ((1-theta)*x1^2)*(0<t)*(t<x1^(1+theta))
F3d1<-function(t,x1,theta) {
  if(theta==0) return(F3d1.theta.0(t,x1))
  if(theta==1) return(F3d1.theta.1(t,x1))
  if(0<theta & theta!=1) return(F3d1.theta(t,x1,theta))
}

#Función de distribución para (x2/x3)
F2d3<-function(t,x3,theta) {
  return(
    ((x3^(1/theta)<= t)*(t<=x3^(1/(1+theta))))*

```

```

( t/x3 + (t^(1-theta))/(theta-1) - theta*(x3^(1/theta-1))/
(theta-1) ) + (x3^(1/(1+theta))<t)*(t<=1)*( t^(1-theta))/
(theta-1) - (t^(-theta))/
theta + (1+1/theta)*(x3^(-theta/(1+theta))) -
theta*(x3^(1/theta-1))/(theta-1) ) )/f.x3(x3, theta)
}

#Función de distribución Para (x3/x2)
F3d2<-function(t,x2,theta)
(x2^(1+theta) <= t)*(t <= x2^theta)*
( 1+theta + t/((x2^theta)*log(x2)) -
(1+log(t+(t<x2^(1+theta))))/log(x2) ) -
(0 <= t)*(t < x2^(1+theta))*t*(1-x2)/((x2^(1+theta))*log(x2))

#Inversa de la funcion de distribución F3
F3.inv.aux<-function(t,theta,v.input) F3(t,theta) - v.input
F3.inv.v<-function(v,theta)
uniroot(F3.inv.aux,interval=c(0.0000001,.999999),
theta=theta,v.input=v)$root

#Inversa de la funcion de disitribucion F1d3
F1d3.inv.aux<-function(t,x3,theta,u.input)
F1d3(t,x3,theta)-u.input
F1d3.inv.v<-function(u,w,theta)
uniroot(F1d3.inv.aux,interval=c(.001,0.999999),
u.input=u,x3=w,theta=theta)$root

#Inversa de la funcion de disitribucion F2d3
F2d3.inv.aux<-function(t,x3,theta,v.input)
F2d3(t,x3,theta)- v.input
F2d3.inv.v<-function(v,w,theta)
uniroot(F2d3.inv.aux,interval=c(0.0000001,0.9999),
v.input=v,x3=w,theta=theta,

```

```

        extendInt = "yes")$root

#Inversa de la funcion de disitribucion F3d2
F3d2.inv.aux<-function(t,x2,theta,w.input)
  F3d2(t,x2,theta) - w.input
F3d2.inv.v<-function(w,v,theta)
  uniroot(F3d2.inv.aux,interval=c(.0000001,v^theta),
    w.input=w,x2=v,theta=theta,
    extendInt = "yes")$root

#Inversa de la funcion de disitribucion F3d1
F3d1.inv.aux<-function(t,x1,theta,v.input)
  F3d1(t,x1,theta) - v.input
F3d1.inv.v<-function(v,u,theta)
  uniroot(F3d1.inv.aux,interval=c(0.00000001,u^(1+theta)-.0000001),
    v.input=v,x1=u,theta=theta)$root

#Esperanza de x2
E.2<-1/4

#Esperanza de x3 dato theta = 0
E.3.t0<-1/4
E.3.t1<-1/12

#Esperanza de X3 Para 0<theta<1
E.3.t<-function(theta) 1/(2*theta^2+6*theta+4)

#Esperanza de X3^2
E.2.f.3<-function(theta) 1/(12*theta^2+24*theta+9)

#Desv. Est. de X2
S.d.f.2<-sqrt(1/9-1/16)

```

```

#Desv. Est. de X3 Para theta = 0
S.d.f.3.t0<-sqrt(1/9-1/16)

#Desv. Est. de x3 Para theta = 1
S.d.f.3.t1<-sqrt(1/45-1/144)

#Desv. Est. de x3 Para 0<theta<>1
S.d.f.3.t<-function(theta)
  sqrt(1/(12*theta^2+24*theta+9)-1/(2*theta^2+6*theta+4)^2)

#Rho(x1, x2)
rho.x1.x2<-(1/6-1/8)/(.2204*.2886)

#Rho(x1, x3)
#Para 0<theta!=1

#Esperanza de f(x1, x3)
E.f.1.3<- function(theta) 1/(2*theta^2+8*theta+6)

#Varianza de x_3
V.f.3<-function(theta) E.2.f.3(theta)-E.3.t(theta)^2

#Desviación estandar de x_3
S.d.f.1<-sqrt(1/3-1/4)
S.d.f.3<-function(theta) sqrt(V.f.3(theta))

#Calculando rho para todos los casos de theta
Cov.x1.x3<-function(theta)
  E.f.1.3(theta)-.5*E.3.t(theta)
rho.x1.x3<-function(x2,theta)
  Cov.x1.x3(theta)/(S.d.f.1*S.d.f.3(theta))

```

```

#Rho(x2,x3)
#Para theta = 0
#E(x2,x3)
Cov.x2.x3.t0<-(1/6+1/4-1/3-E.2*E.3.t0)
rho.x2.x3.t0<-Cov.x2.x3.t0/(S.d.f.2*S.d.f.3.t0)

#Para theta = 1
#E(x2,x3)
Cov.x2.x3.t1<-(1/8+1/6-1/4-E.2*E.3.t1)
rho.x2.x3.t1<-Cov.x2.x3.t1/ (S.d.f.2* S.d.f.3.t1)

#Para 0<theta <> 1
#E(x2,x3)
Cov.x2.x3.t<-
  function(theta)(-1)/(4*(2*theta^2+6*theta+4))-1/
  (2*(theta+3))+1/(2*(theta+2))
rho.x2.x3.t<-
  function(theta)Cov.x2.x3.t(theta)/(S.d.f.2*S.d.f.3.t(theta))

#Cov(x2,x3) para todos los casos de theta
Cov.x2.x3<-function(theta) {
  if (theta < 0 ) return("Valor de theta no permitido")
  if (theta ==0) return(Cov.x2.x3.t0)
  if (theta == 1) return(Cov.x2.x3.t1)
  else { return(Cov.x2.x3.t(theta))}
}

#Rho(x2,x3) para todos los casos de theta
rho.x2.x3<-function(x1,theta) {
  if (theta < 0 ) return("Valor de theta no permitido")
  if (theta ==0) return(rho.x2.x3.t0)
  if (theta == 1) return(rho.x2.x3.t1)
}

```

```

else { return(rho.x2.x3.t(theta))}
}

#Rho(x1, x3|x2)
E13d2<-function(x2,theta) -x2^theta/(4*log(x2))*(1-x2^2)
E1d2<-function(x2) -1/log(x2)*(1-x2)
E2_1d2<-function(x2) -1/(2*log(x2))*(1-x2^2)
E3d2<-function(x2,theta) x2^(theta+1)/(2*log(x2))-x2^theta/(2*log(x2))
E2_3d2<-function(x2,theta)
  +x2^(2*(theta+1))/(6*log(x2))-x2^(2*theta)/(6*log(x2))
cov<-function(x2,theta) E13d2(x2,theta)-E1d2(x2)*E3d2(x2,theta)
sigma1d2<-function(x2) sqrt(E2_1d2(x2)-E1d2(x2)^2)
sigma3d2<-function(x2,theta) sqrt(E2_3d2(x2,theta)-E3d2(x2,theta)^2)
rho13d2<-function(x2,theta) cov(x2,theta)/
  ((sigma1d2(x2))*(sigma3d2(x2,theta)))

#Rho(x2, x3|x1)

#no aplica la correlación condicional para theta = 0
rho23d1.1<-sqrt(3/7)
rho23d1.t<-function(theta) return((sqrt(3)*theta)/
  (((2*theta^2+6*theta+4)*sqrt(((4*theta^2+2*theta+1)))/
  (24*theta^3+60*theta^2+48*theta+12))))

#función que agrupa a todos los casos de esta correlación condicional
rho23d1<-function(x1,theta){
  if(theta == 0) return(0)
  if (theta == 1) return (rho23d1.1)
  if(theta>0 & theta !=1) return (rho23d1.t(theta))
}

#Rho(x1, x2|x3) Para theta = 0
rho12d3.0<-function(x3) ((x3^2-1)*log(x3)-2*x3^2+4*x3-2)*

```



```

(1/(4*log(x3)^2*sqrt(-1/log(x3)*(1/2-x3^2/2)-1/log(x3)^2*(1-x3)^2))*
(1/sqrt(1/log(x3)*(-1/6+x3^2/6)-1/log(x3)^2*(x3/2-1/2)^2)))

#Rho(x1,x2/x3) dado theta = 1
#E(x1,x2/x3)
E12d3.1<-function(x3) ((1+x3)-(sqrt(x3)+x3/sqrt(x3)))/f.x3(x3,1)

#E(x1/x3)
E1d3.1<-
function(x3) (-log(sqrt(x3)))^2+(log(x3)*log(sqrt(x3)))/f.x3(x3,1)

#E(x1^2/x3)
E2_1d3.1<-
function(x3)
(2*(-1-sqrt(x3)*log(sqrt(x3))+sqrt(x3)))/
f3(x3)-log(x3)*(1-sqrt(x3))/f.x3(x3,1)

#E(x2/x3)
E2d3.1<-function(x3) (-1/2-log(sqrt(x3))+x3/2)/f.x3(x3,1)

#E(x2^2/x3)
E2_2d3.1<-function(x3) (1/2-2*sqrt(x3)/3+x3^2/6)/f.x3(x3,1)

#Cov(x1,x2/x3)
Cov12d3.1<-function(x3) E12d3.1(x3)-E1d3.1(x3)*E2d3.1(x3)

#Sigma(x1/x3)
sigma1d3.1<-function(x3) sqrt(E2_1d3.1(x3)-E1d3.1(x3)^2)

#Sigma(x2/x3)
sigma2d3.1<-function(x3) sqrt(E2_2d3.1(x3)-E2d3.1(x3)^2)

```

```

#Rho(x1, x2/x3)
rho12d3.1<-function(x3)
  Cov12d3.1(x3)/(sigma1d3.1(x3)*sigma2d3.1(x3))

#Rho(x1, x2/x3) dado 0<theta != 1

#E(x1/x3)
E1d3.t<-function(x3,theta) 1/((1-theta)*f.x3(x3,theta))*
((1-x3^((1-theta)/(1+theta)))/(1-theta)-x3^(1/theta-1)/
(1-1/theta)*(1-x3^((theta-1)/(theta*(theta+1))))))

#E(x1^2/x3)
E2_1d3.t<-function(x3,theta)1/(f.x3(x3,theta)*(1-theta))*
(-(-theta^2+1)/(2*theta^2-5*theta+2)*x3^((-theta+2)/(theta+1))+1/
(-theta+2)-x3^((1/theta)-1)/(-1/theta+2))

#Sigma(x1/x3)
Sigma1d3.t<-
  function(x3,theta) sqrt(E2_1d3.t(x3,theta)-E1d3.t(x3,theta)^2)

#E(x2/x3)
E2d3.t<-
  function(x3,theta)
    ((1/(1-theta)-1/(2-theta))-x3^((1-theta)/(theta+1))/(1-theta)+
x3^((2/(theta+1)-1))/2-x3^(2/theta-1)/2+x3^((2-theta)/theta)/(2-theta))/
  f.x3(x3,theta)

#E(x2^2/x3)
E2_2d3.t<-
  function(x3,theta)
    (1/(2-theta)-1/(3-theta)-x3^((2-theta)/(1+theta))/(2-theta)+
x3^(3/(1+theta)-1)/3-x3^(3/theta-1)/3+x3^((3-theta)/theta)/
(3-theta))/f.x3(x3,theta)

```

```

#E(x1, x2/x3)
E12d3.t<-
  function(x3,theta) 1/f.x3(x3,theta)*((1-x3^((2-theta)/(1+theta)))/
(2-theta)^2-x3^((2-theta)/theta)/((2-theta)*
(-2/theta+1))*((1-x3^((-2/theta+1)/(1+theta))))

#Sigma(x2/x3)
Sigma2d3.t<-function(x3,theta)
  sqrt(E2_2d3.t(x3,theta)-E2d3.t(x3,theta)^2)

#Cov(x1, x2/x3)
Cov12d3.t<-function(x3,theta)
  E12d3.t(x3,theta)-E1d3.t(x3,theta)*E2d3.t(x3,theta)

#Rho(x1, x2/x3)
rho12d3.t<-
  function(x3,theta)
    Cov12d3.t(x3,theta)/(Sigma1d3.t(x3,theta)*Sigma2d3.t(x3,theta))

#Rho(x1, x2/x3) unificado
rho12d3<-function(x3,theta) {
  if(theta == 0 ) return(rho12d3.0(x3))
  if(theta == 1 ) return(rho12d3.1(x3))
  if(0<theta & theta != 1) return(rho12d3.t(x3,theta))
  else return("Fuera de rango")}

#Densidades de las cópulas del modelo teórico
d.C.12<-function(u,v){
  x1<-u
  x2<-F3.inv.v(v,0)
  return((1/x1)/(((0<x1)*(x1<1))*(-log(x2)))*(0<x2)*(x2<x1))}

```

```
d.C.13<-function(u,v,theta){
  x1<-u
  x3<-F3.inv.v(v,theta)
  return(f.1.3(x1,x3,theta)/((0<x1)*(x1<1)*f.x3(x3,theta)))
}

d.C.23<-function(v,w,theta){
  x2<-F3.inv.v(v,0)
  x3<-F3.inv.v(w,theta)
  return( f.2.3(x2,x3,theta)/(f.x2(x2)*f.x3(x3,theta)))}

d.C.23d1<-function(v,w,x1,theta){
  x2<-F3d1.inv.v(v,x1,0)
  x3<-F3d1.inv.v(w,x1,theta)
  return( f.2.3d1(x2,x3,x1,theta)/((f2d1(x2,x1))*(f3d1(x3,x1,theta))))
}

d.C.12d3<-function(u,v,x3,theta){
  x1<-F1d3.inv.v(u,x3,theta)
  x2<-F2d3.inv.v(v,x3,theta)
  return(f.1.2d3(x1,x2,x3,theta)/(f1d3(x1,x3,theta)*f2d3(x2,x3,theta)))
}

d.C.13d2<-function(u,w,x2,theta){
  x1<-F1d3.inv.v(u,x2,0)
  x3<-F3d2.inv.v(w,x2,theta)
  return(f.1.3d2(x1,x3,x2,theta)/(f1d2(x1,x2)*f3d2(x3,x2,theta)))
}
```

```
#Densidad de la copula C12
dC.12.analitica<-
  function(u,v)
    outer(u,v,function(uu,vv) mapply(d.C.12,uu,vv))

#Densidad de la copula C12d3
dC.12d3.analitica<-
  function(u,v,w,theta) outer(u,v,
  function(uu,vv,ww=w,tt=theta)
    mapply(d.C.12d3,uu,vv,x3=ww,theta=tt))

#Densidad de la copula C13
dC.13.analitica<-
  function(u,w,theta) outer(u,w,
  function(uu,ww) mapply(d.C.13,uu,ww,theta=theta))

#Densidad de la copula C13d2
dC.13d2.analitica <-
  function(u,w,v,theta) outer(u,w,
  function(uu,ww,vv=v,tt=theta)
    mapply(d.C.13d2,uu,ww,x2=vv,theta=tt))

#Densidad de la copula C23
dC.23.analitica<-
  function(v,w,theta) outer(v,w,
  function(vv,ww) mapply(d.C.23,vv,ww,theta=theta))

#Densidad de la copula C23/1
dC.23d1.analitica <-
  function(v,w,u,theta) outer(v,w,
  function(vv,ww) mapply(d.C.23d1,vv,ww,x1=u,theta=theta))
```

```
# Metodología de Cópulas en Racimo:
# en esta parte se definen las
# funciones que permiten desa-
# llar la metodología de có-
# pulas en racimo

#Obtiene las pseudo-observaciones no condicionales de un vector
#trivariado
#El indice indica cual de las pseudo-observaciones regresara la funcion
#si es 0, las devuelve todas
#datos_muestra: datos de los que se obtendra las pseudo-observaciones
#datos_pob: todos los datos de la poblacion original

get_pseudo_observaciones_no_cond<-
  function(indice,datos_muestra,datos_pob)
{
  if(indice==0){
    contador<-0
    u1<-sapply(datos_muestra[,1],Fn.Bernshtein,datos_pob[,1])
    contador<-0
    u2<-sapply(datos_muestra[,2],Fn.Bernshtein,datos_pob[,2])
    contador<-0
    u3<-sapply(datos_muestra[,3],Fn.Bernshtein,datos_pob[,3])
    return(cbind(u1,u2,u3))
  }
  if(indice==1){
    u1<-sapply(datos_muestra[,1],Fn.Bernshtein,datos_pob[,1])
    return(u1)
  }
  if(indice==2){
    u2<-sapply(datos_muestra[,2],Fn.Bernshtein,datos_pob[,2])
    return(u2)
  }
}
```

```
if(indice==3){
  u3<-sapply(datos_muestra[,3],Fn.Bernshtein,datos_pob[,3])
  return(u3)
}
}

#Parte en 2 la muestra original:
# ->En dos partes en base a la mediana de X1, si se condiciono en X1
# ->En dos partes en base a la mediana de X2, si se condiciono en X2
# ->En dos partes en base a la mediana de X3, si se condiciono en X3
#Y obtiene las pseudo observaciones
#Además, obtiene las pseudo-observaciones condicionales de los datos
#ingresados
#Parametros:
# ->indice: variable aleatoria condicionada
# ->datos : datos a analizar

get_pseudo_observaciones_cond<-function(indice,datos)
{
  if(indice==1)
  {
    indice_1<-2
    indice_2<-3
  }
  if(indice==2)
  {
    indice_1<-1
    indice_2<-3
  }
  if(indice==3)
  {
    indice_1<-1
```

```

    indice_2<-2
  }
  cat("Diviando muestra, de acuerdo a la media de la
      variable que condiciona...\n")
  muestra.parte1_dividida_en_X_indice<-
    datos[which(datos[,indice]<=median(datos[,indice])),]
  rownames(muestra.parte1_dividida_en_X_indice)<-NULL
  muestra.parte2_dividida_en_X_indice<-
    datos[which(datos[,indice]>median(datos[,indice])),]
  rownames(muestra.parte2_dividida_en_X_indice)<-NULL
  cat("Listo ...\n")
  cat("Obteniendo pseudo-observaciones no
      condicionales de primera particion...\n")
  u1_u2_u_3_P1<-
    get_pseudo_observaciones_no_cond(0,
    muestra.parte1_dividida_en_X_indice,datos)
  cat("Listo ...\n")
  cat("Obteniendo pseudo-observaciones no
      condicionales de segunda particion...\n")
  u1_u2_u_3_P2<-
    get_pseudo_observaciones_no_cond(0,
    muestra.parte2_dividida_en_X_indice,datos)
  cat("Listo ...\n")
  pseudo_obs_cond_p1_dividida_en_indice<-
    matrix(rep(0,length(muestra.parte1_dividida_en_X_indice[,1])*2),
           ncol=2)
  pseudo_obs_cond_p2_dividida_en_indice<-
    matrix(rep(0,length(muestra.parte2_dividida_en_X_indice[,1])*2),
           ncol=2)
#obteniendo copulas empiricas
  matriz.copem_1<-genmat.copem(cbind(datos[,indice],datos[,indice_1]))
  matriz.copem_2<-genmat.copem(cbind(datos[,indice],datos[,indice_2]))

```



```
#obteniendo pseudo-observaciones condicionales parte 1
cat("Obteniendo pseudo-observaciones
    condicionales de primera particion...\n")
for (i in 1:length(u1_u2_u_3_P1[,indice]))
{
  pseudo_obs_cond_p1_dividida_en_indice[i,1]<-
    cv.du(u1_u2_u_3_P1[i,indice],u1_u2_u_3_P1[i,indice_1],
          matriz.copem_1)
}
for (i in 1:length(u1_u2_u_3_P1[,indice]))
{
  pseudo_obs_cond_p1_dividida_en_indice[i,2]<-
    cv.du(u1_u2_u_3_P1[i,indice], u1_u2_u_3_P1[i,indice_2],
          matriz.copem_2)
}
cat("Listo...\n")

#obteniendo pseudo-observaciones condicionales parte 2
cat("Obteniendo pseudo-observaciones
    condicionales de segunda particion...\n")
for (i in 1:length(u1_u2_u_3_P2[,indice]))
{
  pseudo_obs_cond_p2_dividida_en_indice[i,1]<-
    cv.du(u1_u2_u_3_P2[i,indice],
          u1_u2_u_3_P2[i,indice_1],matriz.copem_1)
}
for (i in 1:length(u1_u2_u_3_P2[,indice]))
{
  pseudo_obs_cond_p2_dividida_en_indice[i,2]<-
    cv.du(u1_u2_u_3_P2[i,indice],
          u1_u2_u_3_P2[i,indice_2],matriz.copem_2)
}
cat("Listo...\n")
```

```
cat("Ensamblando ambos tipo de pseudo-observaciones
    en una matriz por cada tipo...\n")
pseudo_obs_cond<-rbind(pseudo_obs_cond_p1_dividida_en_indice,
                      pseudo_obs_cond_p2_dividida_en_indice)
pseudo_obs_no_cond<-rbind(u1_u2_u_3_P1,u1_u2_u_3_P2)
cat("Listo...\n")

return(list(pseudo_obs_cond_p1_dividida_en_indice,
           pseudo_obs_cond_p2_dividida_en_indice,
           pseudo_obs_cond,pseudo_obs_no_cond))
}

#Simula las observaciones empleando el método
#no paramétrico mostrado en la tesis
simulacion_cond<-
function(numero_simulaciones,indice,datos)
{
  if(indice==1)
  {
    indice_1<-2
    indice_2<-3
  }
  if(indice==2)
  {
    indice_1<-1
    indice_2<-3
  }
  if(indice==3)
  {
    indice_1<-1
    indice_2<-2
  }
  V1<-V2<-c()
```

```

u<-runif(numero_simulaciones)
matriz.copem<-
  genmat.copem(cbind(datos[,indice],datos[,indice_1]))
#obteniendo la pseudo inversas utilizando como V. Condicional PHIV
for (i in 1:numero_simulaciones){
  unif<-runif(1,min=0,max=1)
  V1[i]<-
    uniroot(function(t,un=unif,mc=matriz.copem) un - cv.du(u[i],t,mc),
            interval=c(0,1))$root
  cat("Obteniendo V1 ",i,"->",V1[i],"\n")
}
matriz.copem<-
  genmat.copem(cbind(datos[,indice],datos[,indice_2]))
for (i in 1:numero_simulaciones){
  unif<-runif(1,min=0,max=1)
  V2[i]<-
    uniroot(function(t,un=unif,mc=matriz.copem) un - cv.du(u[i],t,mc),
            interval=c(0,1))$root
  cat("Obteniendo V2 ",i,"->",V2[i],"\n")
}
Y1<-Y2<-X3<-c()
for (i in 1:length(u)){
  cat(i,": u -> ",u[i])
  X3[i]<-Fn.inv.Bernshtein(u[i], datos[,indice])
  cat(", X3 -> ",X3[i],"\n")
}
for (i in 1:length(V1)){
  cat(i,": V1 -> ",V1[i])
  Y1[i]<-Fn.inv.Bernshtein(V1[i], datos[,indice_1])
  cat(", Y1 -> ",Y1[i],"\n")
}
for (i in 1:length(V2)){
  cat(i,": V2-> ",V2[i])
}

```

```

    Y2[i]<-Fn.inv.Bernshtein(V2[i], datos[,indice_2])
    cat(", Y2-> ",Y2[i],"\n")
  }
  if(indice==1)
  {
    return(cbind(X3,Y1,Y2))
  }

  if(indice==2)
  {
    return(cbind(Y1,X3,Y2))
  }

  if(indice==3)
  {
    return(cbind(Y1,Y2,X3))
  }
}

#Permite simular un conjunto de observaciones trivariadas a partir
#del método de regresión lineal
simulaciones_lineales<-function(numero_simulaciones,muestra){
  #simulación de los datos usando ajuste de regresion lineal
#simulando la variable VP
  sim.reg.PHIV<-simulacion_cond(numero_simulaciones,1,muestra)[,1]
  lm.VP.phiv<-lm(muestra[,2]~muestra[,1])
  alpha_vp<-coefficients(lm.VP.phiv)[[1]]
  beta_vp<-coefficients(lm.VP.phiv)[[2]]
  res<-residuals(lm.VP.phiv)
  mean.res<-mean(res)
  var.res<-var(res)
  error<-rnorm(numero_simulaciones,mean.res,sqrt(var.res))
  sim.reg.VP<-alpha_vp+beta_vp*sim.reg.PHIV+error

```

```

shapiro.test(res)
#simulando la variable K en base a sim.reg.VP y PHIV
lm.K.phiv.vp<-lm(muestra[,3]~muestra[,2]+muestra[,1])
delta.k<-coefficients(lm.K.phiv.vp)[[1]]
beta_vp2<-coefficients(lm.K.phiv.vp)[[2]]
beta_phiv<-coefficients(lm.K.phiv.vp)[[3]]
res.2<-residuals(lm.K.phiv.vp)
mean.res2<-mean(res.2)
var.res2<-var(res.2)
error_2<-rnorm(numero_simulaciones,mean.res2,sqrt(var.res2))
sim.reg.K<-delta.k+beta_vp2*sim.reg.VP+beta_phiv*sim.reg.PHIV+error_2
shapiro.test(res.2)
#simulando la variable K en base a sim.reg.VP y PHIV
#tomando en cuenta el termino de interaccion VP y PHIV
lm.K.phiv.interaccion.vp<-lm(muestra$K~muestra$VP*muestra$PHIV)
delta.k<-coefficients(lm.K.phiv.interaccion.vp)[[1]]
beta_vp2<-coefficients(lm.K.phiv.interaccion.vp)[[2]]
beta_phiv<-coefficients(lm.K.phiv.interaccion.vp)[[3]]
beta_interaccion_vp_phiv<-coefficients(lm.K.phiv.interaccion.vp)[[4]]
res.2<-residuals(lm.K.phiv.vp)
mean.res2<-mean(res.2)
var.res2<-var(res.2)
error_2<-rnorm(numero_simulaciones,mean.res2,sqrt(var.res2))
sim.reg.K.interaccion<-delta.k+
  beta_vp2*sim.reg.VP+beta_phiv*sim.reg.PHIV+
  beta_interaccion_vp_phiv*(sim.reg.PHIV*sim.reg.VP)+error_2
shapiro.test(res.2)
return(cbind(sim.reg.PHIV,sim.reg.VP,sim.reg.K,sim.reg.K.interaccion))}

```

```
#Comparación entre medidas de dependencia

#Obtiene la rho de Spearman
rho.spearman<-function(x,y)
{
  matriz.copem<<-genmat.copem(cbind(x,y))
  n<-length(x)
  u<-v<-seq(0.0001,0.9999,length=n)
  Cop_bern<-genmat.copem.Bernshtein(u,v)$copemB
  matriz_producto<-outer(u,v,function(aa=u,bb=v) aa*bb)
  return(12/(n^2-1)*(sum(Cop_bern-matriz_producto)))
}

#Obtiene la sigma de Wolff-Schwizer
sigma.schweizer<-function(x,y)
{
  matriz.copem<<-genmat.copem(cbind(x,y))
  n<-length(x)
  u<-v<-seq(0.0001,0.9999,length=n)
  Cop_bern<-genmat.copem.Bernshtein(u,v)$copemB
  matriz_producto<-outer(u,v,function(aa=u,bb=v) aa*bb)
  return(12/(n^2-1)*(sum(abs(Cop_bern-matriz_producto))))
}
```

A.2. Código para generar imágenes y de las tablas del documento

```

#Gráfica de los histogramas de cada variable aleatoria simulada
#del modelo paramétrico para cualquier parametro theta
graficar_histograma_simulacion<-function(datos,param_theta){
  png(filename=paste("histograma_simul_theta_",param_theta,".png"),
       width=1176,height=484)
  layout(matrix(c(1,0,1,2,0,2,3,0,3),3,3,
  byrow=F),respect=T,width=c(3,3,3))
  #Primera va
  hist(datos[,1],freq=F,main="Histograma v.a X1",xlab="X1")
  z<-seq(min(datos[,1]),max(datos[,1]),length=1000)
  lines(z,dunif(z),col="red",lwd=5)
  #Segunda va
  hist(datos[,2],freq=F,main="Histograma v.a X2",xlab="X2")
  z<-seq(min(datos[,2]),max(datos[,2]),length=1000)
  lines(z,f.x2(z),col="red",lwd=5)
  #Tercera va
  hist(datos[,3],freq=F,main="Histograma v.a X3",xlab="X3")
  z<-seq(min(datos[,3]),max(datos[,3]),length=1000)
  lines(z,f.x3(z,param_theta),col="red",lwd=5)
  dev.off()
}

#Gráfica de las funciones de densidad bivariada
#del modelo paramétrico en 3D
#para cualquier parámetro theta
grafica_densidades_bivariadas_persp<-function(param_theta){
  png(filename=
    paste("densidades_bivar_persp_simul_theta_",
    param_theta,".png"),

```

```

        width=1176,height=484)
x1<-x2<-x3<-seq(0.05,.9999,length=100)
par(mfrow=c(1,3))
z<-outer(x1,x2,function(a,b) mapply(f.1.2,a,b))
persp(z,theta=60,xlab="x1",ylab="x2",zlab="f12",main="f(x1,x2)")
z<-outer(x1,x3,function(a,b) mapply(f.1.3,a,b,
                                   MoreArgs=list(theta=param_theta)))
persp(z,theta=60,xlab="x1",ylab="x3",zlab="f13",main="f(x1,x3)")
z<-outer(x2,x3,function(a,b) mapply(f.2.3,a,b,
                                   MoreArgs=list(theta=param_theta)))
persp(z,theta=60,xlab="x2",ylab="x3",zlab="f23",main="f(x2,x3)")
dev.off()
}

#Gráfica las funciones de densidad bivariada
#del modelo paramétrico en forma de curvas de nivel
#para cualquier parámetro theta
grafica_densidades_bivariadas_image<-function(param_theta){
  png(filename=
        paste("densidades_bivar_image_simul_theta_",param_theta,".png"),
        width=1176,height=400)
x1<-x2<-x3<-seq(0.05,.9999,length=100)
par(mfrow=c(1,3))
z<-outer(x1,x2,function(a,b) mapply(f.1.2,a,b))
image(z,xlab="x1",ylab="x2",main="f(x1,x2)")
z<-outer(x1,x3,function(a,b) mapply(f.1.3,a,b,
                                   MoreArgs=list(theta=param_theta)))
image(z,xlab="x1",ylab="x3",main="f(x1,x3)")
z<-outer(x2,x3,function(a,b) mapply(f.2.3,a,b,
                                   MoreArgs=list(theta=param_theta)))
image(z,xlab="x2",ylab="x3",main="f(x2,x3)")
dev.off()
}

```



```
#Gráfica las observaciones simuladas del modelo teorico cuando  
#theta=0  
grafica_observaciones_bivariadas_plot<-function(datos){  
  png(filename=  
    paste("observaciones_bivar_plot_simul_theta_0.png"),  
    width=1176,height=400)  
  par(mfrow=c(1,3))  
  plot(datos[,1],datos[,2],xlab="x1",ylab="x2",cex.lab=1.7)  
  plot(datos[,1],datos[,2],xlab="x1",ylab="x3",cex.lab=1.7)  
  plot(datos[,2],datos[,3],xlab="x2",ylab="x3",cex.lab=1.7)  
  dev.off()  
}  
  
#Gráfica e las densidades bivariadas, las curvas de nivel y  
#los diagramas de dispersión del modelo teórico  
grafica_densidades_bivariadas_persp_image_plot<-  
  function(va_uno,va_dos,datos,param_theta){  
    if(va_uno==1 & va_dos==2){  
      png(filename=  
        paste("densidades_bivar_persp_image_plot_simul_1_2_theta_",  
          param_theta, ".png"),width=1176,height=400)  
      x1<-x2<-x3<-seq(0.05,.9999,length=100)  
      par(mfrow=c(1,3))  
      z<-outer(x1,x2,function(a,b) mapply(f.1.2,a,b))  
      persp(z,theta=60,xlab="x1",ylab="x2",zlab="f12",  
        main=paste("Theta: ",param_theta," , f(x1,x2)"))  
      image(z,xlab="x1",ylab="x2")  
      plot(datos,xlab="x1",ylab="x2")  
      dev.off()  
    }  
    if(va_uno==1 & va_dos==3){  
      png(filename=
```

```

paste("densidades_bivar_persp_image_plot_simul_1_3_theta_",
param_theta, ".png"),width=1176,height=400)
x1<-x2<-x3<-seq(0.05, .9999,length=100)
par(mfrow=c(1,3))
z<-outer(x1,x3,function(a,b)
  mapply(f.1.3,a,b,MoreArgs=list(theta=param_theta)))
persp(z,theta=60,xlab="x1",ylab="x3",zlab="f13",
main=paste("Theta: ",param_theta," , f(x1,x3)"))
image(z,xlab="x1",ylab="x3")
plot(datos,xlab="x1",ylab="x3")
dev.off()
}
if(va_uno==2 & va_dos==3){
  png(filename=
paste("densidades_bivar_persp_image_plot_simul_2_3_theta_",
param_theta, ".png"),width=1176,height=400)
x1<-x2<-x3<-seq(0.05, .9999,length=100)
par(mfrow=c(1,3))
z<-outer(x2,x3,function(a,b)
  mapply(f.2.3,a,b,MoreArgs=list(theta=param_theta)))
persp(z,theta=60,xlab="x2",ylab="x3",zlab="f23",
main=paste("Theta: ",param_theta," , f(x2,x3)"))
image(z,xlab="x2",ylab="x3")
plot(datos,xlab="x2",ylab="x3")
dev.off()
}
}

#Gráfica las correlaciones condicionales y no condicionales del
#modelo teórico para cualquier parametro theta
graficar_corr_cond_vs_no_cond_param_theta<-function(param_theta)
{
  png(paste("corr_cond_vs_no_cond_theta_",param_theta, ".png"),

```

```

width=1176,height=420)
par(mfrow=c(1,3),mar=c(6,6,3,3),oma=c(0,0,3,0))
# rho(x1,x3)
X2<-seq(0.01,.999,length=1000)
rho13.cond<-sapply(X2,rho13d2,theta=param_theta)
rho13.no_cond<-sapply(X2,rho.x1.x3,theta=param_theta)
plot(X2,as.vector(t(rho13.cond)),type="l",col="red",
      ylim=c(0,1),xlab="X2",
      ylab=substitute(paste(rho,"(",X[1],",",X[3], "|",X[2],
                           ", ",theta,"=",param_theta,")")),
      cex.lab=2.5,cex.axis=1.5,lwd=3)
mtext("Correlaciones condicionales vs No condicionales",
      NORTH<-3, line=0,
      adj=0.5, cex=2.2, outer=TRUE)
lines(X2,rho13.no_cond,col="blue",lwd=3)
legend(.57,1,c("No condicional","Condicional"),
      lty=c(1,1), lwd=c(2.5,2.5),
      col=c("blue","red"),cex=1.5)
#rho(x1,x2)
X3<-seq(0.01,.999,length=100)
rho12.cond<-sapply(X3,rho12d3,theta=param_theta)
rho12.no_cond<-rep(rho.x1.x2,length(X3))
plot(X3,rho12.cond,type="l",col="red",ylim=c(0,1),
      ylab=substitute(paste(rho,
                           "(",X[1],",",X[2], "|",X[3],",",theta,"=",param_theta,")")),
      cex.lab=2.5,cex.axis=1.5,lwd=3)
lines(X3,rho12.no_cond,col="blue",lwd=3)
legend(.57,1,c("No condicional","Condicional"),
      lty=c(1,1), lwd=c(2.5,2.5),
      col=c("blue","red"),cex=1.5)

#valor de de las absisas
X1<-seq(0.01,.999,length=100)

```

```

rho23.cond<-sapply(X1,rho23d1,theta=param_theta)
rho23.no_cond<-sapply(X1,rho.x2.x3,theta=param_theta)
plot(X1,as.vector(t(rho23.cond)),type="l",
      col="red",ylim=c(0,1),xlab="X1",
      ylab=substitute(paste(rho,"(",X[2],",",
X[3],")|",X[1],",",theta,"=",param_theta,")")),
      cex.lab=2.5,cex.axis=1.5,lwd=3)
lines(X1,rho23.no_cond,col="blue",lwd=3)
legend(.57,1,c("No condicional","Condicional"),
      lty=c(1,1), lwd=c(2.5,2.5),
      col=c("blue","red"),cex=1.5)
dev.off()
}

#Grafica las densidades de cópula
graficar.dCop<-
function(tt,x_cond,tamano_main,
tamano_sub,tamano_etiquetas,tamano_ejes,marg,
outer_marg){
u<-seq(0.05,0.95,length=100)
v<-w<-seq(0.05,0.95,length=100)

png(filename=
paste("dens_Cop_theta_",tt,".png"),width=1176,height=1200)
par(mfrow=c(3,4),mar=marg,oma=outer_marg)
image(dC.12.analitica(u,v),sub="Modelo no condicionado",
      main=paste("dC12, theta = ",tt),cex.main=tamano_main,
      cex.sub=tamano_sub,cex.lab=tamano_etiquetas[2],
      cex.axis=tamano_ejes)
image(dC.12d3.analitica(u,v,w=x_cond[1],theta=tt),
      sub=paste("X3 condicionada en:",x_cond[1]),
      cex.main=tamano_main,
      cex.lab=tamano_etiquetas[2],

```

```
cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.12d3.analitica(u,v,w=x_cond[2],theta=tt),
      sub=paste("X3 condicionada en:",x_cond[2]),
      cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.12d3.analitica(u,v,w=x_cond[3],theta=tt),
      sub=paste("X3 condicionada en:",x_cond[3]),
      cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.13.analitica(u,w,theta=tt),sub="Modelo no condicionado",
      main=paste("dC13, theta = ",tt),
      cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.13d2.analitica(u,w,v=x_cond[1],theta=tt),
      sub=paste("X2 condicionada en:",x_cond[1]),
      cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.13d2.analitica(u,w,v=x_cond[2],theta=tt),
      sub=paste("X2 condicionada en:",x_cond[2]),
      cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.13d2.analitica(u,w,v=x_cond[3],theta=tt),
      sub=paste("X2 condicionada en:",x_cond[3]),
      cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.23.analitica(v,w,theta=tt),
      sub="Modelo no condicionado",
```

```

main=paste("dC23, theta = ",tt),
cex.main=tamanyo_main,
cex.lab=tamanyo_etiquetas[2],
cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.23d1.analitica(v,w,u=x_cond[1],theta=tt),
sub=paste("X1 condicionada en:",x_cond[1]),
cex.main=tamanyo_main,
cex.lab=tamanyo_etiquetas[2],
cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.23d1.analitica(v,w,u=x_cond[2],theta=tt),
sub=paste("X1 condicionada en:",x_cond[2]),
cex.main=tamanyo_main,
cex.lab=tamanyo_etiquetas[2],
cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
image(dC.23d1.analitica(v,w,u=x_cond[3],theta=tt),
sub=paste("X1 condicionada en:",x_cond[3]),
cex.main=tamanyo_main,
cex.lab=tamanyo_etiquetas[2],
cex.sub=tamanyo_sub,cex.axis=tamanyo_ejes)
dev.off()}

#Gráfica de las pseudo-observaciones no condicionales
#del modelo teórico
graficar_pseudo_obs_no_cond<-function(flag_muestra_sim,theta,
pseudo_obs_nocond,tamanyo_main,tamayo_label,
tamanyo_axis){
  pseudo_obs.1.2<-cbind(pseudo_obs_nocond[,1],pseudo_obs_nocond[,2])
  pseudo_obs.1.3<-cbind(pseudo_obs_nocond[,1],pseudo_obs_nocond[,3])
  pseudo_obs.2.3<-cbind(pseudo_obs_nocond[,2],pseudo_obs_nocond[,3])
  if(flag_muestra_sim == 0)
  {
    png(paste("pseudo_obs_no_cond_datos_theta_",theta,".png"),width=1200

```

```
,height=850)
  string_muestra_sim<-"muestra"
}
else
{
  png("pseudo_obs_no_cond_datos_simulacion.png",width=1200,
      height=850)
  string_muestra_sim<-"simulacion"
}
par(mfrow=c(2,3))
u<-v<-seq(0.05,0.95,length=300)
matriz.copem<<-genmat.copem(pseudo_obs.1.2)
C12.bern<-genmat.copem.Bernshtein(u,v)$copemB
matriz.copem<<-genmat.copem(pseudo_obs.1.3)
C13.bern<-genmat.copem.Bernshtein(u,v)$copemB
matriz.copem<<-genmat.copem(pseudo_obs.2.3)
C23.bern<-genmat.copem.Bernshtein(u,v)$copemB
plot(pseudo_obs_nocond[,1],pseudo_obs_nocond[,2],xlab="X1",ylab="X2",
main=paste("Pseudo-Obs ",string_muestra_sim),cex.main=tamanyo_main,
cex.lab=tamayo_label,cex.axis=tamanyo_axis)
plot(pseudo_obs_nocond[,1],pseudo_obs_nocond[,3],xlab="X1",ylab="X3",
cex.main=tamanyo_main,cex.lab=tamayo_label,cex.axis=tamanyo_axis,
cex.lab=tamayo_label,cex.axis=tamanyo_axis)
plot(pseudo_obs_nocond[,2],pseudo_obs_nocond[,3],xlab="X2",ylab="X3",
cex.main=tamanyo_main,cex.lab=tamayo_label,cex.axis=tamanyo_axis,
cex.lab=tamayo_label,cex.axis=tamanyo_axis)
contour(C12.bern,xlab="u",ylab="v",
cex.main=tamanyo_main,cex.lab=tamayo_label,
cex.axis=tamanyo_axis)
contour(C13.bern,xlab="u",ylab="v",
cex.main=tamanyo_main,cex.lab=tamayo_label,
cex.axis=tamanyo_axis)
contour(C23.bern,xlab="u",ylab="v",
```

```

cex.main=tamanyo_main,cex.lab=tamayo_label,
cex.axis=tamanyo_axis)
  dev.off()
}

#Gráfica de los tres pares de pseudo-obs condicionales
#del modelo teórico
graficar_pseudo_obs_cond<-
function(pseudo_obs_cond.1.2,pseudo_obs_cond.1.3,
pseudo_obs_cond.2.3,param_theta,tamanyo_main,
tamanyo_sub,tamanyo_etiquetas,tamanyo_ejes,marg,
outer_marg){
u<-v<-seq(0.05,0.95,length=300)
matriz.copem<<-genmat.copem(pseudo_obs_cond.1.2[[3]])
C12_cond.bern<-genmat.copem.Bernshtein(u,v)$copemB

matriz.copem<<-genmat.copem(pseudo_obs_cond.1.3[[3]])
C13_cond.bern<-genmat.copem.Bernshtein(u,v)$copemB

matriz.copem<<-genmat.copem(pseudo_obs_cond.2.3[[3]])
C23_cond.bern<-genmat.copem.Bernshtein(u,v)$copemB

png(paste("pseudo_obs_cond_",param_theta,".png"),width=1200,
height=850)
par(mfrow=c(2,3),mar=marg,outer_marg)
plot(pseudo_obs_cond.1.2[[3]],type="p",
main=bquote(paste(theta," = ",.(param_theta),
", Pseudo Observaciones Cond en X3")),xlab="u1i",ylab="u2i",
cex.main=2.3,cex.lab=tamanyo_etiquetas[2],
cex.axis=tamanyo_ejes)
plot(pseudo_obs_cond.1.3[[3]],type="p",
main="Pseudo Observaciones Cond en X2",
xlab="u1i",ylab="u2i",

```



```
cex.main=2,cex.lab=tamanyo_etiquetas[2],
cex.axis=tamanyo_ejes)
plot(pseudo_obs_cond.2.3[[3]],type="p",
     main="Pseudo Observaciones Cond en X1",
     xlab="u1i",ylab="u2i",
     cex.main=2,cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
contour(C12_cond.bern)
contour(C13_cond.bern)
contour(C23_cond.bern)
dev.off()
png(paste("pseudo_obs_cond_divididas_",param_theta,"_dado_X3.png"),
    width=1200,height=1200)
par(mfrow=c(3,3),mar=marg,oma=outer_marg)
plot(pseudo_obs_cond.1.2[[1]],type="p",
     main=bquote(paste(theta," = ",.(param_theta),
                       ", Cond en X3",sep="")),
     col="red",xlab="",ylab="u2i",
     cex.main=tamanyo_main,
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,
      cex=tamanyo_etiquetas[1])
mtext(side = 1, text = "< Mediana(Xi)",
      line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.2[[2]],type="p",
     col="blue",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = ">= Mediana(Xi)",
      line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.2[[1]],type="p",
```

```

col="red",xlab="",ylab="u2i",
cex.lab=tamanyo_etiquetas[2],
cex.axis=tamanyo_ejes)
lines(pseudo_obs_cond.1.2[[2]],type="p",
      col="blue",xlab="u1i",ylab="u2i")
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
plot(pseudo_obs_cond.1.3[[1]],type="p",main=bquote(paste(theta," = ",
.(param_theta)," , Cond en X2",sep="")),
      col="red",xlab="",ylab="u2i",cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = "< Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.3[[2]],type="p",
      col="blue",xlab="",ylab="u2i",
      cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = ">= Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.3[[1]],type="p",
      col="red",xlab="",ylab="u2i",
      cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
lines(pseudo_obs_cond.1.3[[2]],type="p",
      col="blue",xlab="",ylab="u2i")
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
plot(pseudo_obs_cond.2.3[[1]],type="p",
      main=bquote(paste(theta," = ",
.(param_theta)," , Cond en X1",sep="")),
      col="red",xlab="",ylab="u2i",cex.main=tamanyo_main,
      cex.lab=tamanyo_etiquetas[2],
      cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = "< Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.2.3[[2]],type="p",

```

```

    col="blue",xlab="",ylab="u2i",
    cex.lab=tamanyo_etiquetas[2],
    cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = ">= Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.2.3[[1]],type="p",
     col="red",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
lines(pseudo_obs_cond.2.3[[2]],type="p",
      col="blue",xlab="u1i",ylab="u2i")
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
dev.off()
}
#Gráficas de las pseudo-obs no condicionales de los
#datos petrofísicos
graficar_pseudo_obs_no_cond_datos_petrofisicos<-
function(flag_muestra_sim,pseudo_obs_nocond,tamanyo_main,
tamayo_label,tamanyo_axis,marg,outer_marg){
pseudo_obs.1.2<-
  cbind(pseudo_obs_nocond[,1],pseudo_obs_nocond[,2])
pseudo_obs.1.3<-
  cbind(pseudo_obs_nocond[,1],pseudo_obs_nocond[,3])
pseudo_obs.2.3<-
  cbind(pseudo_obs_nocond[,2],pseudo_obs_nocond[,3])
if(flag_muestra_sim == 0)
{
  png("pseudo_obs_no_cond_datos_petrofisicos_muestra.png",
      width=1200,height=850)
  string_muestra_sim<-"muestra"
}
else
{

```

```
png("pseudo_obs_no_cond_datos_petrofisicos_simulacion.png",
    width=1200,height=850)
string_muestra_sim<-"simulacion"
}
par(mfrow=c(2,3),mar=marg,outer_marg)
u<-v<-seq(0.05,0.95,length=50)
matriz.copem<<-genmat.copem(pseudo_obs.1.2)
C12.bern<-genmat.copem.Bernshtein(u,v)$copemB
matriz.copem<<-genmat.copem(pseudo_obs.1.3)
C13.bern<-genmat.copem.Bernshtein(u,v)$copemB
matriz.copem<<-genmat.copem(pseudo_obs.2.3)
C23.bern<-genmat.copem.Bernshtein(u,v)$copemB
plot(pseudo_obs_nocond[,1],pseudo_obs_nocond[,2],
     xlab="PHIV",ylab="VP",
     main=paste("Pseudo-Obs ",string_muestra_sim),
     cex.main=tamanyo_main,cex.lab=tamayo_label,
     cex.axis=tamanyo_axis)
plot(pseudo_obs_nocond[,1],pseudo_obs_nocond[,3],xlab="PHIV",ylab="K",
     cex.main=tamanyo_main,cex.lab=tamayo_label,
     cex.axis=tamanyo_axis)
plot(pseudo_obs_nocond[,2],pseudo_obs_nocond[,3],xlab="VP",ylab="K",
     cex.main=tamanyo_main,cex.lab=tamayo_label,
     cex.axis=tamanyo_axis)
contour(C12.bern,xlab="u",ylab="v",cex.main=tamanyo_main,
        cex.lab=tamayo_label,cex.axis=tamanyo_axis)
contour(C13.bern,xlab="u",ylab="v",cex.main=tamanyo_main,
        cex.lab=tamayo_label,cex.axis=tamanyo_axis)
contour(C23.bern,xlab="u",ylab="v",cex.main=tamanyo_main,
        cex.lab=tamayo_label,cex.axis=tamanyo_axis)
dev.off()
}
```

```

#Gráficas de las pseudo-obs condicionales de los
#datos petrofísicos
graficar_pseudo_obs_cond_datos_petrofisicos<-
  function(flag_muestra_sim,pseudo_obs_cond.1.2,pseudo_obs_cond.1.3,
           pseudo_obs_cond.2.3,tamanyo_main,tamanyo_sub,
           tamanyo_etiquetas,tamanyo_ejes,
           marg,outer_marg){
u<-v<-seq(0.05,0.95,length=300)
matriz.copem<<-genmat.copem(pseudo_obs_cond.1.2[[3]])
C12_cond.bern<-genmat.copem.Bernshtein(u,v)$copemB
matriz.copem<<-genmat.copem(pseudo_obs_cond.1.3[[3]])
C13_cond.bern<-genmat.copem.Bernshtein(u,v)$copemB
matriz.copem<<-genmat.copem(pseudo_obs_cond.2.3[[3]])
C23_cond.bern<-genmat.copem.Bernshtein(u,v)$copemB
if(flag_muestra_sim == 0)
{
  png("pseudo_obs_cond_datos_petrofisicos_muestra.png",
      width=1200,height=850)
}
else
{
  png("pseudo_obs_cond_datos_petrofisicos_simulacion.png",
      width=1200,height=850)
}
par(mfrow=c(2,3),mar=marg,outer_marg)
plot(pseudo_obs_cond.1.2[[3]],type="p",
     main="Pseudo Observaciones Cond en K",xlab="u1i",
     ylab="u2i",cex.main=2,cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
plot(pseudo_obs_cond.1.3[[3]],type="p",
     main="Pseudo Observaciones Cond en VP",xlab="u1i",
     ylab="u2i",cex.main=2,cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)

```

```
plot(pseudo_obs_cond.2.3[[3]],type="p",
     main="Pseudo Observaciones Cond en PHIV",xlab="u1i",
     ylab="u2i",cex.main=2,cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
contour(C12_cond.bern,xlab="u1i",ylab="u2i",cex.main=2,
        cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
contour(C13_cond.bern,xlab="u1i",ylab="u2i",cex.main=2,
        cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
contour(C23_cond.bern,xlab="u1i",ylab="u2i",cex.main=2,
        cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
dev.off()
png("pseudo_obs_cond_divididas_datos_petrofisicos.png",
     width=1200,height=1200)
par(mfrow=c(3,3),mar=marg,oma=outer_marg)
plot(pseudo_obs_cond.1.2[[1]],
     type="p",main="Cond en K",col="red",xlab="",
     ylab="u2i",cex.main=tamanyo_main,cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i",
       line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = "< Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.2[[2]],type="p",col="blue",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,
       cex=tamanyo_etiquetas[1])
mtext(side = 1, text = ">= Mediana(Xi)",
       line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.2[[1]],type="p",
     col="red",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
lines(pseudo_obs_cond.1.2[[2]],type="p",
      col="blue",xlab="u1i",ylab="u2i")
```

```
mtext(side = 1, text = "u1i", line = 3,
      cex=tamanyo_etiquetas[1])
plot(pseudo_obs_cond.1.3[[1]],type="p",main="Cond en VP",col="red",
     xlab="",ylab="u2i",cex.main=tamanyo_main,
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,
      cex=tamanyo_etiquetas[1])
mtext(side = 1, text = "< Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.3[[2]],type="p",
     col="blue",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
mtext(side = 1, text = ">= Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.1.3[[1]],type="p",
     col="red",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
lines(pseudo_obs_cond.1.3[[2]],type="p",
      col="blue",xlab="",ylab="u2i")
mtext(side = 1, text = "u1i", line = 3,cex=tamanyo_etiquetas[1])
plot(pseudo_obs_cond.2.3[[1]],type="p",
     main="Cond en PHIV",col="red",
     xlab="",ylab="u2i",cex.main=tamanyo_main,
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,
      cex=tamanyo_etiquetas[1])
mtext(side = 1, text = "< Mediana(Xi)",
      line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.2.3[[2]],type="p",
     col="blue",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
mtext(side = 1, text = "u1i", line = 3,
```

```

    cex=tamanyo_etiquetas[1])
mtext(side = 1, text = ">= Mediana(Xi)", line = 5,cex=tamanyo_sub)
plot(pseudo_obs_cond.2.3[[1]],
     type="p",col="red",xlab="",ylab="u2i",
     cex.lab=tamanyo_etiquetas[2],
     cex.axis=tamanyo_ejes)
lines(pseudo_obs_cond.2.3[[2]],
      type="p",col="blue",xlab="u1i",ylab="u2i")
mtext(side = 1, text = "u1i", line = 3,
      cex=tamanyo_etiquetas[1])
dev.off()
}

# Gráficas de comparación entre modelos

#Grafica datos muestrales contra simulación no paramétrica
#y contra simulación via regresión lineal
graficar_muestra_original_simnoparam_simreglin<-
  function(etiquetas,muestra,simnoparam,simreglin,
           tamanyo_etiquetas,tamanyo_ejes,marg,outer_marg){
png("datos_muestrales_vs_simulacion_no_param_reg_lineal.png",
    width=1300,height=1700)
  par(mfrow=c(4,3),mar=marg,oma=outer_marg)
  plot(muestra[,1],muestra[,2],xlab=etiquetas[1],ylab=etiquetas[2],
       main="Muestra",cex.main=2.3,
       cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
  plot(muestra[,1],muestra[,3],xlab=etiquetas[1],ylab=etiquetas[3],
       cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
  plot(muestra[,2],muestra[,3],xlab=etiquetas[2],ylab=etiquetas[3],
       cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
  plot(simnoparam[,1],simnoparam[,2],xlab=etiquetas[1],
       ylab=etiquetas[2],
       main="Sim. no parametrica",cex.main=2.3,

```



```
    cex.lab=tamanyo_etiquetas[2],
    cex.axis=tamanyo_ejes)
plot(simnoparam[,1],simnoparam[,3],xlab=etiquetas[1],
     ylab=etiquetas[3],
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
plot(simnoparam[,2],simnoparam[,3],xlab=etiquetas[2],
     ylab=etiquetas[3],
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
plot(simreglin[,1],simreglin[,2],xlab=etiquetas[1],
     ylab=etiquetas[2],
     main="Simul.Reg.Lineal",ylim=c(1.5,4.5),cex.main=2.3,
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
plot(simreglin[,1],simreglin[,3],xlab=etiquetas[1],
     ylab=etiquetas[3],
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes,
     ylim=c(min(muestra[,2]),max(muestra[,2])))
plot(simreglin[,2],simreglin[,3],xlab=etiquetas[2],
     ylab=etiquetas[3],
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes,
     ylim=c(min(muestra[,3]),max(muestra[,3])))
plot(simreglin[,1],simreglin[,2],xlab=etiquetas[1],
     ylab=etiquetas[2],
     main="Simul.Reg.Lineal c/n Inter.",ylim=c(1.5,4.5),cex.main=2.3,
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
plot(simreglin[,1],simreglin[,4],xlab=etiquetas[1],
     ylab=etiquetas[3],
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes,
     ylim=c(min(muestra[,2]),max(muestra[,2])))
plot(simreglin[,2],simreglin[,4],xlab=etiquetas[2],
     ylab=etiquetas[3],
     cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes,
     ylim=c(min(muestra[,3]),max(muestra[,3])))
dev.off()}
```

```

#Grafica datos muestrales contra simulación no paramétrica
#y contra simulación via regresión lineal en 3D
graficar_muestra_original_simnoparam_simreglin_3D<-
  function(etiquetas,muestra,simnoparam,simreglin,tamanyo_main,
           tamanyo_sub,tamanyo_etiquetas,tamanyo_ejes,marg,
           outer_marg){
    library(scatterplot3d)
    png("datos_muestrales_vs_simulacion_no_param_reg_lineal_cn_inter.png"
        ,width=1100,height=1200)
    par(mfrow=c(2,2),mar=marg,oma=outer_marg)
    scatterplot3d(muestra[,1],muestra[,2],muestra[,3],color="red",
                  xlab=etiquetas[1],ylab=etiquetas[2],
                  zlab=etiquetas[3],
                  ylim=c(1.5,4.5),main="Muestra Original",
                  cex.main=tamanyo_main,
                  cex.lab=tamanyo_etiquetas[2],
                  cex.axis=tamanyo_ejes)
    scatterplot3d(simnoparam[,1],simnoparam[,2],simnoparam[,3],
                  color="purple",
                  xlab=etiquetas[1],ylab=etiquetas[2],
                  zlab=etiquetas[3],
                  main="Simul. Metodo no parametrico",
                  cex.main=tamanyo_main,
                  cex.lab=tamanyo_etiquetas[2],cex.axis=tamanyo_ejes)
    scatterplot3d(simreglin[,1],simreglin[,2],simreglin[,3],
                  color="orange",xlab=etiquetas[2],
                  ylab=etiquetas[2],zlab=etiquetas[3],
                  main="Simul.Reg.Lineal",
                  cex.main=tamanyo_main,cex.lab=tamanyo_etiquetas[2],
                  cex.axis=tamanyo_ejes,xlim=c(min(muestra[,1]),
                  max(muestra[,1])),ylim=c(min(muestra[,2]),
                  max(muestra[,2])),
  }

```

```
      zlim=c(min(muestra[,3]),max(muestra[,3]))
scatterplot3d(simreglin[,1],simreglin[,2],simreglin[,4],
  color="blue",
  xlab=etiquetas[2],ylab=etiquetas[2],
  zlab=etiquetas[3],main="Simul.Reg.Lineal c/n Inter.",
  cex.main=tamano_main,cex.lab=tamano_etiquetas[2],
  cex.axis=tamano_ejes,
  xlim=c(min(muestra[,1]),max(muestra[,1])),
  ylim=c(min(muestra[,2]),max(muestra[,2])),
  zlim=c(min(muestra[,3]),max(muestra[,3])))

dev.off()
}

#Tabla comparacion Wolff-Schwizer de
#pseudo-observaciones no condicionales
#de los datos petrofísicos
imprimir_coeficientes<-function(datos)
{
  cat("Coeficiente de Spearman: ",
    rho.spearman(datos[,1],datos[,2]),
    " Coeficiente de Schweizer",
    sigma.schweizer(datos[,1],datos[,2]),"\n")
}
```

A.2.1. Desarrollo de la tesis

```

#Capítulo uno: Marco Teórico

#Cotas Frechet Hoefding y Cópula producto
#sobre las Cotas de Frechet-Hoeffding
#y funcion monotona creciente de X
u1 <- seq(0,1,length=10)
v1 <- seq(0,1,length=10)
u1_cota_sup <- seq(0,1,length=1000)
v1_cota_sup <- seq(0,1,length=1000)
Cxy_W<-function(u,v) max(u+v-1,0)
Cxy_M<-function(u,v) min(u,v)
Cxy_Indep<-function(u,v) u*v
z_W<-outer(u1,v1,function(uu=u1,vv=v1) mapply(Cxy_W,uu,vv))
z_M<-outer(u1,v1,function(uu=u1,vv=v1) mapply(Cxy_M,uu,vv))
z_M_cota_Sup<-outer(u1_cota_sup,v1_cota_sup,
                    function(uu=u1_cota_sup,
                              vv=v1_cota_sup) mapply(Cxy_M,uu,vv))
z_Indep<-outer(u1,v1,function(uu=u1,vv=v1) mapply(Cxy_Indep,uu,vv))

#Ilustración 1.1: Cotas de Frechet-Hoeffding
dev.new()
par(mfrow=c(2,3),mai=c(0.2772,0.2772,0.2772,0.2772))
contour(u1,v1,z_W,main="Cota inferior F-H",xlab="u",ylab="v")
contour(u1_cota_sup,v1_cota_sup,z_M_cota_Sup,
        main="Cota Superior F-H",xlab="u",ylab="v")
contour(u1,v1,z_Indep,main="Cópula Producto",xlab="u",ylab="v")
persp(u1,v1,z_W,xlab="u",ylab="v",zlab="W(u,v)",theta=-30)
persp(u1,v1,z_M,xlab="u",ylab="v",zlab="M(u,v)",theta=-30)
persp(u1,v1,z_Indep,xlab="u",ylab="v",zlab="Pi(u,v)",theta=-30)

```

```
#Capítulo dos: Metodología de cópulas en racimo

#Tabla 2.I sobre resumen de estadística descriptiva
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
summary(muestra)

#Ilustración 2.1: Histogramas contra funciones marginales
#de densidad
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
graficar_histograma_simulacion(muestra,0)

#Ilustración 2.2: Gráficos 3D de función densidad bivariada
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
grafica_densidades_bivariadas_persp(0)

#Ilustración 2.3: Curvas de nivel a color de función de
#densidad bivariada
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
grafica_densidades_bivariadas_image(0)

#Ilustración 2.4: Diagramas de dispersión de observaciones
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
grafica_observaciones_bivariadas_plot(muestra)

#Ilustración 2.5: Diagramas de dispersión de observaciones simuladas
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
```

```
sim_0<-gen.vec.unif(tamano_muestra,0)
grafica_observaciones_bivariadas_plot(sim_0)

#Tabla de Coeficiente de correlación teorico v.s muestral
cor(muestra[,1],muestra[,2])
rho.x1.x2
cor(muestra[,1],muestra[,3])
rho.x1.x3(0,0)
cor(muestra[,2],muestra[,3])
rho.x2.x3(0,0)

#Modelo teórico cuando theta = 0
#cargar todo a continuacion
#el codigo antes de ejecutar
#lo correspondiente a las imagenes
#2.6 - 2.10
tamano_muestra<-1000
muestra<-gen.vec.unif(tamano_muestra,0)
simul<-gen.vec.unif(500,0)

#Convirtiendo la simulacion en un data frame
simul.df<-as.data.frame(simul)

#creando el modelo de regresión múltiple
regresion.mult<-lm(formula=simul.df[,1]~simul.df[,2]+simul.df[,3],
                  data=simul.df)
beta<-coef(regresion.mult)
varianza_reg_lineal<-(summary(regresion.mult)$sigma)**2

#Crea los valores de x2 y x3
x1<-seq(min(simul[,1]),max(simul[,1]),length=300)
x2<-seq(min(simul[,2]),max(simul[,2]),length=300)
x3<-seq(min(simul[,3]),max(simul[,3]),length=300)
```

```
f<-function(x2,x3) beta[1]+beta[2]*x2+beta[3]*x3

#Fórmula esperanza condicional
esper<-function(a,b)
  -(max(a,b))/(1-max(a,b))*log(max(a,b))*(a<1)*
  (a>0)*(b<1)*(b>0)

#Fórmula varianza condicional
varianza<-function(a,b)
  (max(a,b)-(max(a,b)^2/(1-max(a,b))^2)*log(max(a,b))^2)*
  (0<a)*(1>a)*(0<b)*(1>b)

#Fórmula f_1/2,3
f1g23<-
  function(a,b,c) max(b,c)/((a^2)*(1-max(b,c)))*
  (b>0)*(b<a)*(a>c)*(0<c)*(0<a)*(a<1)
  x1.reg<-outer(x2,x3,f)
  x1.esper<-outer(x2,x3,function(a,b) mapply(esper,a,b))
  x1.var<-outer(x2,x3,function(a,b) mapply(varianza,a,b))

#2.6: Esperanza condicional vs Regresion Lineal
dev.new()
persp3d(x2, x3, x1.reg, aspect=c(1, 1, 1), col = "lightblue",
        xlab = "x2", ylab = "x3",
        zlab = "Regresion vs Esperanza Condicional")
persp3d(x2, x3, x1.esper, col="red",add=T)

x1.var<-outer(x2,x3,function(a,b) mapply(varianza,a,b))
x1.var_corr_lineal<-matrix(rep(varianza_reg_lineal,
                              length=(length(x2)*length(x3))),
                          nrow=length(x2),ncol=length(x3))
```

```

#2.7: Varianza condicional vs Varianza de Regresion Lineal
persp3d(x2,x3,x1.var,theta=30,phi=20,xlab="x2",ylab="x3",zlab="x1",
        sub="Varianza condicional",col="red")
persp3d(x2,x3,x1.var_corr_lineal,col="blue",add=T)

dev.new()
par(mfrow=c(1,2))
image(x1.var,sub=expression(paste("Var(",X[1],"|",X[2],"",X[3],")")))
image(x1.var_corr_lineal,sub=expression(sigma^2))

#2.8: Prueba Shapiro-Wilk
#Corroborando el p-value al aplicar
#la prueba shaphiro wilk a los residuales
shapiro.test(residuals(regresion.mult))

#2.9: Histograma de residuales y Normal ajustada
#Ajustando los residuales a una normal y graficando un histograma
#con un ajuste a una normal
normal<-fitdistr(residuals(regresion.mult),"normal")
intervalo<-seq(min(residuals(regresion.mult)),
               max(residuals(regresion.mult)),length=100)
ajuste<-dnorm(intervalo,normal$estimate[[1]],
               normal$estimate[[2]])
dev.new()
hist(residuals(regresion.mult),freq=F)
lines(intervalo,ajuste)

#2.10:  $f_{1/23}(x_1 | 0.9, 0.8)$  contra
#Normal ( $\beta_1 + 0.9\beta_2 + 0.8\beta_3, \sigma^2$ )
#Graficando  $f_{1/23}(x_1 | .9, .8)$  y
#una densidad Normal( $\beta_1 + 0.9\beta_2 + 0.8\beta_3, \sigma^2$ )
media<-
regresion.mult$coef[1]+0.9*

```



```
regresion.mult$coef[2]+0.8*regresion.mult$coef[3]
desv<-summary(regresion.mult)$sigma
intervalo<-seq(0.001,2,length=1000)
dens.f1g23<-sapply(intervalo,f1g23,b=.9,c=.8)
dens.normal<-dnorm(intervalo,media,desv^2)
dev.new()
plot(intervalo,dens.normal,xlab="",ylab="")
lines(intervalo,dens.f1g23,col="red")

#Simula los datos del modelo paramétrico
numero_cols_sim<-3
tamano_muestra<-1000
theta_vect<-c(0.01,0.49,0.51,0.99,1.01,1.98,2.03)
tamano_vector_theta<-length(theta_vect)
sim<-array(rep(0,tamano_muestra*numero_cols_sim*tamano_vector_theta),
           dim=c(tamano_muestra,numero_cols_sim,tamano_vector_theta))
i<-0
for(i in 1:tamano_vector_theta){
sim[, ,i]<-gen.vec.unif(tamano_muestra,theta_vect[i])
}

#Tabla 2.III a 2.VI sobre resumen de estadística descriptiva
summary(sim[, ,1])
summary(sim[, ,2])
summary(sim[, ,3])
summary(sim[, ,4])
summary(sim[, ,5])
summary(sim[, ,6])
summary(sim[, ,7])
```

```

#Ilustraciones 2.11 - 2.19
#Graficando los histogramas, las funciones de densidad
#conjunta y los diagramas de dispersion
#del modelo teorico evaluado en cada valor de theta
for(i in 1:tamanyo_vector_theta){
graficar_histograma_simulacion(sim[, ,i],theta_vect[i])
grafica_densidades_bivariadas_persp(theta_vect[i])
grafica_densidades_bivariadas_image(theta_vect[i])
grafica_densidades_bivariadas_persp_image_plot(1,2,
        cbind(sim[,1,i],sim[,2,i]),theta_vect[i])
grafica_densidades_bivariadas_persp_image_plot(1,3,
        cbind(sim[,1,i],sim[,3,i]),theta_vect[i])
grafica_densidades_bivariadas_persp_image_plot(2,3,
        cbind(sim[,2,i],sim[,3,i]),theta_vect[i])
}

#Tabla 2.VII comparacion entre correlaciones
#teóricas y muestrales
for(i in 1:tamanyo_vector_theta){
  cor1 <- cor(cbind(sim[,1,i],sim[,2,i]),method="pearson")[1,2]
  cor2 <- cor(cbind(sim[,1,i],sim[,3,i]),method="pearson")[1,2]
  cor3 <- cor(cbind(sim[,2,i],sim[,3,i]),method="pearson")[1,2]
  cat("----Inicio----\n", "")
  cat("Correlacion rho(x1,x2|theta=",theta_vect[i],") -> teorica: ",
        rho.x1.x2," muestral: ",cor1,"\n")
  cat("Correlacion rho(x1,x3|theta=",theta_vect[i],") -> teorica: ",
        rho.x1.x3(0,theta_vect[i])," muestral: ",cor2,"\n")
  cat("Correlacion rho(x2,x3|theta=",theta_vect[i],") -> teorica: ",
        rho.x2.x3(0,theta_vect[i])," muestral: ",cor3,"\n")
  cat("----Fin----\n", "")
}

```

```
#Ilustración 2.20: comparacion entre  
#correlaciones teóricas vs muestrales  
theta<-seq(0,5,length=300)  
correlacion1<-  
  rep(rho.x1.x2,300)  
correlacion2<-  
  sapply(theta,rho.x1.x3,x2=0)  
correlacion3<-  
  sapply(theta,rho.x2.x3,x1=0)  
dev.new()  
plot(theta,correlacion1,col="blue",type="l",ylim=c(0.3,0.8),  
      xlab=expression(theta),  
      ylab=expression(rho[xy]), cex.main=1.25,  
      cex.lab=1.5, cex.axis=0.75,lwd=2)  
lines(theta,correlacion2,col="red",type="l",lwd=2)  
lines(theta,correlacion3,col="black",type="l",lwd=2)  
legend(4,.8,c(expression(rho(x[1],x[2])),expression(rho(x[1],x[3])),  
              expression(rho(x[2],x[3]))),lty=c(1,1), lwd=c(2.5,2.5),  
        col=c("blue","red","black"))  
  
#Ilustraciones 2.21,2.23,2.25,2.27,2.29,2.31,2.33  
#Correlaciones condicionales  
graficar_corr_cond_vs_no_cond_param_theta(0.01)  
graficar_corr_cond_vs_no_cond_param_theta(0.49)  
graficar_corr_cond_vs_no_cond_param_theta(0.51)  
graficar_corr_cond_vs_no_cond_param_theta(0.99)  
graficar_corr_cond_vs_no_cond_param_theta(1.01)  
graficar_corr_cond_vs_no_cond_param_theta(1.98)  
graficar_corr_cond_vs_no_cond_param_theta(2.03)
```

```

#Ilustraciones 2.22,2.24,2.26,2.28,2.30, 2.32, 2.34:
#Densidades de Cópula condicionales
#gráfica dependiendo del valor de theta como primer argumento
graficar.dCop(0.01,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))
graficar.dCop(0.49,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))
graficar.dCop(0.51,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))
graficar.dCop(0.99,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))
graficar.dCop(1.01,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))
graficar.dCop(1.98,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))

graficar.dCop(2.03,c(0.2,0.5,0.8),2.5,2.5,c(1.8,2.8),1.5,
              c(6,2,6,1),c(3,3,1,1))

#Ilustraciones 2.35 a 2.47 y Tabla Remillard Scalliet 2.IX
#de las pseudo observaciones
#no condicionales condicionales de los Datos Petrofísicos
for(i in 1:tamanyo_vector_theta){
  pseudo_obs_no_cond_muestra<-
    get_pseudo_observaciones_no_cond(0,sim[, ,i],sim[, ,i])
  graficar_pseudo_obs_no_cond(0,theta_vect[i],pseudo_obs_no_cond_muestra,
    tamanyo_main=1.5,tamayo_label=1.8,tamanyo_axis=1.8)
}

for(i in 1:tamanyo_vector_theta){
  pseudo_obs_cond_muestra.1.2<-
    get_pseudo_observaciones_cond(3,sim[, ,i])
  pseudo_obs_cond_muestra.1.3<-

```

```
  get_pseudo_observaciones_cond(2,sim[, ,i])
pseudo_obs_cond_muestra.2.3<-
  get_pseudo_observaciones_cond(1,sim[, ,i])
graficar_pseudo_obs_cond(pseudo_obs_cond_muestra.1.2,
  pseudo_obs_cond_muestra.1.3,
  pseudo_obs_cond_muestra.2.3,
  theta_vect[i],3.5,1.8,c(1.8,2.8),1.5,
  c(6,6,6,6),c(2,2,3,2))

#p-values de prueba de contraste de hipotesis
#usando el paquete TwoCop
#util para evaluar cual de las permutaciones de las
#variables aleatorias
#permite cumplir con la simplifying assumption
pvalue.Condicional.C12d3<-
  TwoCop(pseudo_obs_cond_muestra.1.2[[1]],
    pseudo_obs_cond_muestra.1.2[[2]])$pvalue
pvalue.Condicional.C13d2<-
  TwoCop(pseudo_obs_cond_muestra.1.3[[1]],
    pseudo_obs_cond_muestra.1.3[[2]])$pvalue
pvalue.Condicional.C23d1<-
  TwoCop(pseudo_obs_cond_muestra.2.3[[1]],
    pseudo_obs_cond_muestra.2.3[[2]])$pvalue
cat("theta: ",theta_vect[i],"C12d3: ",
  pvalue.Condicional.C12d3,"C13d2: ",
  pvalue.Condicional.C13d2,"C23d1: ",
  pvalue.Condicional.C23d1,"\n")
}
```

```
#Capítulo 3: Aplicación a un caso de estudio

#Archivo con la muestra de los datos petrofísicos
#Reescribir la ruta dependiendo de la ubicación
#del archivo
muestra<-
  read.csv("datos3_ordenados en VP.csv",header=T)

#Tabla 3.I de analisis descriptivo de datos petrofísicos
summary(muestra)

#Ilustración 3.1
dev.new()
boxplot(muestra)

#Ilustración 3.2
dev.new()
par(mfrow=c(1,3))
plot(muestra[,1],muestra[,2],xlab="X1",ylab="X2")
plot(muestra[,1],muestra[,3],xlab="X1",ylab="X3")
plot(muestra[,2],muestra[,3],xlab="X2",ylab="X3")
dev.new()
scatterplot3d(muestra[,1],muestra[,2],muestra[,3],xlab="PHIV",
  ylab="VP",zlab="K",main="Muestra de datos petrofísicos",
  color="red",ylim=c(min(muestra[,2]),max(muestra[,2])))

#Ilustración 3.3
#histogramas de las tres variables aleatorias
dev.new()
par(mfrow=c(1,3),mar=c(4,6,0,0),oma=c(0,0,0,0))
hist(muestra[,1],freq=F,xlab="PHIV",
  ylab="densidad",main="",cex.lab=2)
hist(muestra[,2],freq=F,xlab="VP",
```

```

    ylab="densidad",main="",cex.lab=2)
hist(muestra[,3],freq=F,xlab="K",
     ylab="densidad",main="",cex.lab=2)

#Pseudo observaciones condicional
pseudo_obs_cond.1.2_muestra<-
  get_pseudo_observaciones_cond(3,muestra)
pseudo_obs_cond.1.3_muestra<-
  get_pseudo_observaciones_cond(2,muestra)
pseudo_obs_cond.2.3_muestra<-
  get_pseudo_observaciones_cond(1,muestra)
pseudo_obs_no_cond_muestra<-
  get_pseudo_observaciones_no_cond(0,muestra,muestra)

#Gráfica de pseudo-observaciones condicionales modelo
# de caso datos petrofisicos
#Lista de parametros
#->pseudo_obs_cond.1.3
#->pseudo_obs_cond.2.3
#->pseudo_obs_cond.1.2
#->tamaño main
#->tamaño sub
#->tamaño labels de los ejes (tamano de eje y ,tamano de eje x)
#->tamaño de los numeros de los ejes
#->tamaño margenes (vector mar)
#->tamaño de los outer margins (vector oma)

#Ilustración 3.4
graficar_pseudo_obs_no_cond_datos_petrofisicos(0,
  pseudo_obs_no_cond_muestra,
  tamaño_main=2.3,tamayo_label=2,
  tamaño_axis=2,c(6,6,6,6),c(2,2,3,2))

```

```
#Ilustración 3.5
dev.new()
scatterplot3d(pseudo_obs_no_cond_muestra[,1],
              pseudo_obs_no_cond_muestra[,2],
              pseudo_obs_no_cond_muestra[,3],
              color="red",xlab="u",ylab="v",
              zlab="w")

#Tabla 3.II de comparacion Wolff-Schwizer de
#pseudo-observaciones no condicionales
#de los datos petrofisicos

#Aplicando coeficiente de correlacion y sigma de Wolff-Schwizer
#a pseudo observaciones
imprimir_coeficientes(cbind(pseudo_obs_no_cond_muestra[,1],
                             pseudo_obs_no_cond_muestra[,2]))
cor(cbind(pseudo_obs_no_cond_muestra[,1],
          pseudo_obs_no_cond_muestra[,2]),
    method="spearman")

imprimir_coeficientes(cbind(pseudo_obs_no_cond_muestra[,1],
                             pseudo_obs_no_cond_muestra[,3]))
cor(cbind(pseudo_obs_no_cond_muestra[,1],
          pseudo_obs_no_cond_muestra[,3]),
    method="spearman")

imprimir_coeficientes(cbind(pseudo_obs_no_cond_muestra[,2],
                             pseudo_obs_no_cond_muestra[,3]))
cor(cbind(pseudo_obs_no_cond_muestra[,2],
          pseudo_obs_no_cond_muestra[,3]),
    method="spearman")
```


#Ilustración 3.6 y 3.7

```
graficar_pseudo_obs_cond_datos_petrofisicos(0,  
  pseudo_obs_cond.1.2_muestra,  
  pseudo_obs_cond.1.3_muestra,  
  pseudo_obs_cond.2.3_muestra,  
  3.5,1.8,c(1.8,2.8),1.5,cbind(6,6,6,6),  
  cbind(2,2,3,2))
```

*#Tabla 3.III comparación pseudo-observaciones condicionales
#de la muestra original de datos petrofísicos*

```
pvalue.Condicional.C12d3<-  
  TwoCop(pseudo_obs_cond.1.2_muestra[[1]],  
    pseudo_obs_cond.1.2_muestra[[2]])$pvalue  
pvalue.Condicional.C13d2<-  
  TwoCop(pseudo_obs_cond.1.3_muestra[[1]],  
    pseudo_obs_cond.1.3_muestra[[2]])$pvalue  
pvalue.Condicional.C23d1<-  
  TwoCop(pseudo_obs_cond.2.3_muestra[[1]],  
    pseudo_obs_cond.2.3_muestra[[2]])$pvalue
```

*#Tabla 3.IV de comparacion Wolff-Schwizer del mejor condicionamiento
#de los datos petrofisicos*

```
imprimir_coeficientes(pseudo_obs_cond.2.3_muestra[[3]])  
cor(pseudo_obs_cond.2.3_muestra[[3]],method="spearman")
```

```

#Simula observaciones en base a la variable aleatoria
# en que se condicione
#Parametros:
# ->indice: variable aleatoria condicionada
# ->datos : datos a analizar

#Simula datos petrofisicos
simulacion_cond_PHIV<-simulacion_cond(473,1,muestra)
pseudo_obs_no_cond_simulacion<-
  get_pseudo_observaciones_no_cond(0,
    simulacion_cond_PHIV,simulacion_cond_PHIV)

pseudo_obs_cond.1.2_simulacion<-
  get_pseudo_observaciones_cond(3,simulacion_cond_PHIV)
pseudo_obs_cond.1.3_simulacion<-
  get_pseudo_observaciones_cond(2,simulacion_cond_PHIV)
pseudo_obs_cond.2.3_simulacion<-
  get_pseudo_observaciones_cond(1,simulacion_cond_PHIV)

#Ilustración 3.8
graficar_pseudo_obs_no_cond_datos_petrofisicos(1,
  pseudo_obs_no_cond_simulacion,
  tamanyo_main=2.3,tamayo_label=2,tamanyo_axis=2
  ,c(6,6,6,6),c(2,2,3,2))

#Ilustración 3.9
graficar_pseudo_obs_cond_datos_petrofisicos(1,
  pseudo_obs_cond.1.2_simulacion,
  pseudo_obs_cond.1.3_simulacion,
  pseudo_obs_cond.2.3_simulacion,3.5,1.8,
  c(1.8,2.8),1.5,c(6,6,6,6),c(2,2,3,2))

```

```

#Ilustración 3.10
#Gráfica de los datos muestrales contra simulacion
# no parametrica dado (VP,K/PHIV)
dev.new()
plot(muestra[,1],muestra[,2],xlab="PHIV",ylab="VP",main="Originales")
plot(muestra[,1],muestra[,3],xlab="PHIV",ylab="K")
plot(muestra[,2],muestra[,3],xlab="VP",ylab="K")
plot(simulacion_cond_PHIV[,1],simulacion_cond_PHIV[,2],
      xlab="PHIV",ylab="VP",
      main="Simulacion")
plot(simulacion_cond_PHIV[,1],simulacion_cond_PHIV[,3],
      xlab="PHIV",ylab="K")
plot(simulacion_cond_PHIV[,2],simulacion_cond_PHIV[,3],
      xlab="VP",ylab="K")

#Tabla 3.V p-values prueba Remillard - Scaillet muestra vs simulacion
Comp_cop_PHIV.vs.VP<-TwoCop(cbind(muestra[,1],muestra[,2]),
                             cbind(simulacion_cond_PHIV[,1],
                                     simulacion_cond_PHIV[,2]))$pvalue
Comp_cop_PHIV.vs.K<-TwoCop(cbind(muestra[,1],muestra[,3]),
                             cbind(simulacion_cond_PHIV[,1],
                                     simulacion_cond_PHIV[,3]))$pvalue
Comp_cop_VP.vs.K<-TwoCop(cbind(muestra[,2],muestra[,3]),
                             cbind(simulacion_cond_PHIV[,2],
                                     simulacion_cond_PHIV[,3]))$pvalue
Comp_cop_PHIV.vs.VP.vs.K<-TwoCop(cbind(muestra[,1],muestra[,2],
                                         muestra[,3]),cbind(simulacion_cond_PHIV[,1],
                                                             simulacion_cond_PHIV[,2],
                                                             simulacion_cond_PHIV[,3]))$pvalue
sim_no_param<-simulacion_cond(1000,1,muestra)
simulacion_reg_lin<-simulaciones_lineales(1000,muestra)

```

#Ilustración 3.11

```
graficar_muestra_original_simnparam_simreglin(c("PHIV","VP","K"),  
muestra,sim_no_param,simulacion_reg_lin,  
c(1.8,1.8),1.5,c(6,6,6,6),c(2,2,3,2))
```

#Ilustración 3.12

```
graficar_muestra_original_simnparam_simreglin_3D(c("PHIV","VP","K"),  
muestra,sim_no_param,  
simulacion_reg_lin,2.3,  
1.8,c(1.5,1.5),0.8,c(6,6,2,2),  
c(2,2,4,2))
```

Bibliografía

- Aas, K., Czado, C., Frigessi, A. y Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and economics*, 44(2), 182-198.
- Bedford, T. y Cooke, R. M. (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32(1-4), 245-268.
- Bedford, T. y Cooke, R. M. (2002, agosto). Vines—a new graphical model for dependent random variables. *Ann. Statist.* 30(4), 1031-1068.
- Czado, C. (2011). The world of vines. Munich, Germany: Technische Universität München.
- Embrechts, P., McNeil, A. y Straumann, D. (1999). Correlation: pitfalls and alternatives. *Risk Magazine*, (5), 61-71.
- Erdely, A. (2009). Cópulas y dependencia de variables aleatorias: una introducción. *Miscelanea Matemática*, (48), 7-28.
- Erdely, A. y Díaz-Viera, M. (2010). Nonparametric and semiparametric bivariate modeling of petrophysical porosity-permeability dependence from well log data. En *Copula theory and its applications*. Lectures in statistics 198. Berlin: Springer-Verlag.
- Joe, H. (1996). Families of m -variate distributions with given margins and $m(m - 1)/2$ bivariate dependence parameters. En L. Rüschendorf, B. Schweizer y M. D. Taylor (Eds.), *Distributions with fixed marginals and related topics* (Vol. Volume 28, pp. 120-141). Lecture Notes—Monograph Series. Hayward, CA: Institute of Mathematical Statistics.
- Kurowicka, D. y Cooke, R. (2006). *Uncertainty analysis with high dimensional dependence modelling*. John Wiley y Sons, Ltd.
- Kurowicka, D. y Joe, H. (2012). *Front matter*. World Scientific.
- Nelsen, R. (2006). *An introduction to copulas* (2nd). La primera edición es de 1999. New York: Springer-Verlag.

- Perez, J. M. y Palacín, A. F. (1987). Estimating the quantile function by bernstein polynomials. *Computational Statistics, Data Analysis*, 5(4), 391-397.
- Rémillard, B. y Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100(3), 377-386.
- Sancetta, A. y Satchell, S. (2004). The bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric theory*, (20), 535-562.
- Schweizer, B. y Wolff, E. F. (1981, julio). On nonparametric measures of dependence for random variables. *Ann. Statist.* 9(4), 879-885.