Probabilidad - II 2024
Universidad Nacional de Colombia - Dic 14

Tutor: Carlos E. Alonso–Malaver.

***ö.* Para masticar-rumiar**:

Yo soy yo mismo,
Pero no el mismo...
Reguetonero - Brazil

Distribuciones Multivariadas

Contexto General

1. De Forma sencilla
   - Sea $U = (\Omega, \mathfrak{F}, P)$ espacio de probabilidad.
   - Las variables aleatorias $X_1, X_2, \ldots, X_p$ están definidas sobre U, $p \geq 2$.
   - Definido el vector $\mathbf{X} := (X_1, X_2, \ldots, X_p)$
   - El propósito es Estudiar $P(\mathbf{X} \in B)$, $B \subset \mathbb{R}^p$.

2. Si has entendido las distribuciones bivariadas, lo básico para trabajar probabilidad en el mundo multivariado $p \geq 2$ se presenta a continuación:
   - Dado un vector discreto, i.e. sus componentes son v.a. discretas,
     - Su función de masa de probabilidad conjunta $f_{\mathbf{X}}$, nos dice:

       $$f_{\mathbf{X}}(\mathbf{x}) = P(\mathbf{X} = \mathbf{x}) = P(X_1 = x_1, X_2 = x_2, \ldots, X_p = x_p) \quad \mathbf{x} \in \mathbb{R}^p.$$

     - De lo anterior, dado $B \subset \mathbb{R}^p$,

       $$P(\mathbf{X} \in B) = \sum_{\mathbf{x} \in B} f_{\mathbf{X}}(\mathbf{x}).$$

- En particular $F_{\mathbf{X}}(\mathbf{x}_0) = P(\mathbf{X} \leq \mathbf{x}_0) = \sum\limits_{\mathbf{x} \leq \mathbf{x}_0} f_{\mathbf{X}}(\mathbf{x})$.

- En el caso de un vector aleatorio continuo, i.e. sus componentes son v.a. continuas,
  - Se hace uso de una función de densidad conjunta $f_{\mathbf{X}}$
  - Dado un Boreliano $B \subset \mathbb{R}^p$,

$$P(\mathbf{X} \in B) = \int\limits_{\mathbf{x} \in B} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x}$$

  - En particular

$$\begin{aligned} F_{\mathbf{X}}(\mathbf{x}) = P(\mathbf{X} \leq \mathbf{x}) &= \int\limits_{\infty}^{\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} \\ &= \int\limits_{-\infty}^{x_1} \int\limits_{-\infty}^{x_2} \cdots \int\limits_{-\infty}^{x_p} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_p. \end{aligned}$$

- Ejemplo 1: Dadas X, Y, Z v.a. continuas con f.d. conjunta dada por:

$$f_{X,Y,Z}(x, y, z) = \frac{12}{7}(x + 2y + 3z) I_{\{0 < x \leq y \leq z < 1\}}(x, y, z).$$

  Asume se desea llegar $g_{Z|X=x,Y=y}(z)$. Este objetivos nos permite trabajar los conceptos ya desarrollos en el caso bivariado.
  Previo, se requiere $f_{X,Y}(x, y)$, veamos:

$$\begin{aligned} f_{X,Y}(x, y) &= \frac{12}{7} \int\limits_{y}^{1} (x + 2y + 3z) \, \mathrm{d}z \\ &= \frac{12}{7} \left[x + (2 - x)y - \frac{7y^2}{2} + \frac{3}{2} \right] I_{\{0 < x \leq y < 1\}}(x, y) \end{aligned}$$

Usando lo anterior:

$$\begin{aligned}
g_{Z|X=x,Y=y}(z) &= \frac{f_{X,Y,Z}(x,y,z)}{f_{X,Y}(x,y)} \\
&= \frac{\frac{12}{7}(x+2y+3z)}{\frac{12}{7}\left[x+(2-x)y-\frac{7y^2}{2}+\frac{3}{2}\right]} \\
&= \frac{x+2y+3z}{x+(2-x)y-\frac{7y^2}{2}+\frac{3}{2}} I_{\{y\leq z<1\}}(z), \quad 0 \leq x \leq y
\end{aligned}$$

Lo anterior para, $x = \frac{1}{3}$ y $y = \frac{1}{2}$

$$\begin{aligned}
g_{Z|X=\frac{1}{4},Y=\frac{1}{2}}(z) &= \frac{\frac{1}{4}+2\frac{1}{2}+3z}{\frac{1}{4}+(2-\frac{1}{4})\frac{1}{2}-\frac{7\frac{1}{2^2}}{2}+\frac{3}{2}} I_{\{\frac{1}{2}\leq z<1\}}(z) \\
&= \frac{4}{7}\left(\frac{5}{4}+3z\right) I_{\{\frac{1}{2}\leq z<1\}}(z)
\end{aligned}$$

Verifica que $g_{Z|X=\frac{1}{4},Y=\frac{1}{2}}(z)$ es función de densidad.

- De forma análoga se puede hallar $f_X(x), f_Y(y), f_Z(z)$,

$$\begin{aligned}
f_X(x) &= \frac{12}{7}\int_x^1\int_y^1 x+2y+3z\mathrm{d}z\mathrm{d}y \\
&= \frac{12}{7}\left(\frac{4}{3}-x-2x^2+\frac{5}{3}x^3\right) I_{\{0<x<1\}}(x)
\end{aligned}$$

- Obtenidas $f_X(x)$ y $f_{X,Y,Z}(x,y,z)$ podemos hallar

$$g_{Y,Z|X=x}(y,z) = \frac{f_{X,Y,Z}(x,y,z)}{f_X(x)},$$

  que es una función de densidad bivariada condicional.

- **En resumen**: si tenemos $f_{X,Y,Z}(x,y,z)$, podemos derivar la siguiente información:
  - Las marginales bivariadas $f_{X,Y}(x,y)$, $f_{X,Z}(x,z)$ y $f_{Y,Z}(y,z)$.
  - Las marginales univariadas $f_X(x)$, $f_Y(y)$ y $f_Z(z)$.

- Las condicionales bivariadas $g_{X,Y|Z=z}(x,y)$, $g_{X,Z|Y=y}(x,z)$ y $g_{Y,Z|X=x}(y,z)$.
- Las condicionales univariadas $g_{X|Y=y,Z=z}(x)$, $g_{Y|X=x,Z=z}(y)$ y $g_{Z|X=x,Y=y}(z)$.

- Una forma más general del resultado anterior es, si se tiene un vector aleatorio $\mathbf{X} = (X_1, X_2, \ldots, X_p)$ con función de masa o densidad conjunta $f_{\mathbf{X}}(\mathbf{x})$, si particionamos el vector $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, para cada componente se puede obtener
  - Las funciones de masa o densidad marginales, i.e. $f_{\mathbf{X}_j}(\mathbf{x})$, $j = 1, 2$.
  - Las funciones de masa o densidad condicional, i.e. $g_{\mathbf{X}_i|\mathbf{X}_j=u}(x|u)$; $i, j = 1, 2$, $i \neq j$.
- En el caso continuo, i.e. el vector aleatorio $\mathbf{X} = (X_1, X_2, \ldots, X_p)$ continuo, con funciones de densidad y de distribución conjunta $f_{\mathbf{X}}(\cdot)$ y $F_{\mathbf{X}}(\cdot)$, se cumple:
  - $F_{\mathbf{X}}(\mathbf{x}) = \int\limits_{-\infty}^{\mathbf{x}} f_{\mathbf{X}}(\mathbf{u}) d\mathbf{u}$.
  - $f_{\mathbf{X}}(\mathbf{x}) = \frac{\partial^p}{\partial x_p \ldots \partial x_1} F_{\mathbf{X}}(\mathbf{x})$.

3. Distribuciones Multivariadas - Independencia: si las v.a. $X_1, X_2, \ldots, X_p$, son independientes entonces:

   - $$F_{\mathbf{X}}(\mathbf{x}) = F_{X_1}(x_1) F_{X_2}(x_2) \ldots F_{X_p}(x_p) \tag{1}$$
   - $$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) f_{X_2}(x_2) \ldots f_{X_p}(x_p) \tag{2}$$

   Donde $\mathbf{X} = (X_1, X_2, \ldots, X_p)$. Si se cumplen las Ec. 1 y 2 entonces $X_1, X_2, \ldots, X_p$ son v.a. independientes.

Transformación de Variables Aleatorias

1. De Forma sencilla:

$$(\Omega, \mathfrak{F}, P) \quad \xrightarrow{X} \quad (\mathbb{R}, \mathfrak{B}, P_X) \quad \xrightarrow{Y:=r(X)} \quad (\mathbb{R}, \mathfrak{B})$$

   - Asume se tiene una v.a. X para la que se conoce su ley de distribución P_X.

- El interés recae sobre una transformación de X, esto es $Y = r(X)$.
- El propósito inicial es hallar P_Y.

2. Caso Discreto: en el caso de una variable aleatoria discreta el truco es devolverse al espacio inicial (usando inversa de conjuntos), lo que significa:

   - Dada una variable aleatoria discreta X, con función de masa de probabilidad $f_X(x) = P(X = x)$.
   - Si se define $Y = r(X)$, se tiene que la función de masa de probabilidad de Y, $f_Y(y) = P(Y = y)$, está dada por:

   $$f_Y(y) = P(Y = y) = \sum_{x:r(x)=y} P(X = x) = \sum_{x:r(x)=y} f_X(x)$$

   - Ejemplo 2. Veamos lo anterior mediante un ejemplo pedagógico.
     - X variable aleatoria que toma valores en en conjunto $\{1, 2, 3, \ldots, 20\}$, con función de masa de probabilidad

     $$f_X(x) = P(X = x) = \frac{x}{210}.$$

     - Si se define $Y = r(X)$, donde $r(x)$:= *Número de divisores del número* x, hallemos $f_Y(\cdot)$.
     - Veamos:
       - Los divisores del número $x = 7$ son $\{1, 7\}$ de donde $Y|_{X=7} = r(7) = 2$
       - Los divisores del número $x = 12$ son $\{1, 2, 3, 4, 6, 12\}$ de donde $Y|_{X=12} = r(12) = 6$
     - Observa y colige:
       - $Y = 1, 2, \ldots, 6$, posibles valores de la v.a. Y
       - $Y^{-1}(\{1\}) = \{1\}$, aquí piensa en los números que tiene un sólo divisor.
       - $Y^{-1}(\{2\}) = \{2, 3, 5, 7, 11, 13, 17, 19\}$ (números primos menores de 20)
       - $Y^{-1}(\{3\}) = \{4, 9\}$, análogo, piensa en los números que tienen tres divisores.

     Si lo anterior es claro:

- $f_Y(1) = f_X(1) = \frac{1}{210}$
- $$\begin{aligned} f_Y(2) &= f_X(2) + f_X(3) + f_X(5) + f_X(7) + f_X(11) \\ &\quad + f_X(13) + f_X(17) + f_X(19) \\ &= \frac{2}{210} + \frac{3}{210} + \ldots + \frac{19}{210} = \frac{77}{210} \end{aligned}$$
- $f_Y(3) = f_X(4) + f_X(9) = \frac{13}{210}$
- $f_Y(5) = f_X(16) + f_X(20) = \frac{36}{210}$

En resumen:

$$f_X(x) = \begin{cases} \frac{x}{210} & \text{si} \quad x = 1, 2, \ldots, 20 \\ 0 & \text{e.o.c.} \end{cases}$$

Si se define $Y := r(X)$, con $r(X)$:=número de divisores de X, entonces

$$f_Y(y) = \begin{cases} \frac{1}{210} & \text{si} \quad y = 1 \\ \frac{77}{210} & \text{si} \quad y = 2 \\ \frac{13}{210} & \text{si} \quad y = 3 \\ \frac{53}{210} & \text{si} \quad y = 4 \\ \frac{36}{210} & \text{si} \quad y = 5 \\ \frac{30}{210} & \text{si} \quad y = 6 \\ 0 & \text{e.o.c.} \end{cases}$$

3. Caso Continuo: Iniciaremos éste caso mediante un ejemplo,

- Sea $U \sim U(0, 1)$.
- Nos interesa la variable aleatoria $Y = -\beta \ln(1 - U)$, donde $\beta > 0$.
- Nuestro objetivo es: hallar F_Y o f_Y (hallaremos ambas)
- De partida se tiene:

$$f_U(u) := I_{(0,1)}(u) \quad \text{y} \quad F_U(u) := P(U \leq u) = \begin{cases} 0 & \text{si} \quad u < 0 \\ x & \text{si} \quad 0 \leq u < 1 \\ 1 & \text{si} \quad u > 1 \end{cases}$$

Para el comportamiento de Y observa

a) Si $U \in (0,1)$ entonces $1-U \in (0,1)$.
b) De lo anterior $\ln(1-U) \in (-\infty, 0)$.
c) Entonces $Y = -\beta \ln(1-U) \in (0,\infty)$, $\beta > 0$.

Ahora:

$$
\begin{aligned}
F_{Y}(y) &:= P(Y \leq y) = P(-\beta \ln(1-U) \leq y) \\
&= P(\ln(1-U) \geq -\frac{y}{\beta}) = P(1-U \geq e^{-\frac{y}{\beta}}) \\
&= P(U \leq 1-e^{-\frac{y}{\beta}}) = F_{U}(1-e^{-\frac{y}{\beta}}) \\
&= \begin{cases} 0 \quad \text{si} \quad 1-e^{-\frac{y}{\beta}} < 0 \\ 1-e^{-\frac{y}{\beta}} \quad \text{si} \quad 0 \leq 1-e^{-\frac{y}{\beta}} < 1 \\ 1 \quad \text{si} \quad 1-e^{-\frac{y}{\beta}} > 1 \end{cases} \\
&= \left(1-e^{-\frac{y}{\beta}}\right) I_{(0,\infty)}(y)
\end{aligned}
$$

De las propiedades de función de densidad:

$$
\begin{aligned}
f_{Y}(y) = \frac{\partial F_{Y}(y)}{\partial y} &= \frac{\partial F_{U}(1-e^{-\frac{y}{\beta}})}{\partial y} \\
&= f_{U}(1-e^{-\frac{y}{\beta}}) * \left(\frac{1}{\beta}e^{-\frac{y}{\beta}}\right) \\
&= \left(\frac{1}{\beta}e^{-\frac{y}{\beta}}\right) I_{(0,\infty)}(y) \qquad (3)
\end{aligned}
$$

Si una v.a. Y tiene densidad dada por la función en la Ec. (3) entonces la v.a. Y tiene distribución exponencial, $Y \sim Exp(\beta)$.

- Lo anterior es útil porque nos permite hacer simulación de Monte Carlo de una v.a. con distribución exponencial. Veamos los pasos para ello,
  a) Genera u de acuerdo a una distribución uniforme $(0,1)$
  b) Calcula $y = -\beta \ln(1-u)$
  c) Los resultados anteriores nos permiten pensar que y (así definido), proviene de una distribución exponencial de parámetro β.
- **Generalizando**: En el ejemplo anterior se tiene que $Y = r(U)$, con $r(u)$ una función 1-a-1 creciente (como función de u), lo que podemos hallar

$r^{-1}(y) = s(y)$, lo que a su vez nos permite generalizar lo hecho, como sigue:
Función Creciente: Dada X v.a. continua con función de densidad $f_X(\cdot)$. Si $Y = r(U)$, con $r(u)$ función monótona creciente entonces

$$\begin{aligned} F_Y(y) &:= P(Y \leq y) = P(r(U) \leq y) = P(U \leq s(y)) \\ &= F_U(s(y)) \end{aligned}$$

Donde $s(\cdot) = r^{-1}(\cdot)$, de lo anterior

$$\begin{aligned} f_Y(y) &= \frac{\partial F_Y(y)}{\partial y} = \frac{\partial F_U(s(y))}{\partial y} \\ &= f_U(s(y))\frac{\partial s(y)}{\partial y} = f_U(s(y))s'(y) \end{aligned} \tag{4}$$

Función Decreciente: para el caso $Y = r(U)$ función 1-a-1 decreciente (como función de U), lo anterior es como sigue

$$\begin{aligned} F_Y(y) &:= P(Y \leq y) = P(r(U) \geq y) = 1 - P(U < s(y)) \\ &= 1 - F_U(s(y)) \end{aligned}$$

De donde:

$$\begin{aligned} f_Y(y) &= \frac{\partial F_Y(y)}{\partial y} = \frac{\partial\left[1 - F_U(s(y))\right]}{\partial y} \\ &= -f_U(s(y))\frac{\partial s(y)}{\partial y} = f_U(s(y))(-s'(y)) \end{aligned} \tag{5}$$

En éste caso observa s es decreciente y por ende $s'(y)$ es una función de valor negativo.
De lo anterior, Ec. (4) y (5), **En general**: si $Y = r(U)$, con $r(u)$ una función absolutamente monótona (creciente o decreciente) diferenciable, se tiene que:

$$f_Y(y) = f_U(s(y))\,|s'(y)| \tag{6}$$

Donde $|\cdot|$ indica valor absoluto.

- Ejemplo 3: Dada X con función de densidad

$$f_X(x) = e^{-x} I_{(0,\infty)},$$

es de interés hallar la densidad de $Y = -X^{\frac{1}{2}}$.
Aplicando el resultado que se presenta en la Ecuación (6):

  - $r(u) = -u^{\frac{1}{2}}$, observa es monótona decreciente.
  - $s(y) = r^{-1}(y) = (-y)^2$ y $\frac{\partial s(y)}{\partial y} = 2y < 0$.
  - De donde:

$$f_Y(y) = f_U(s(y)) |s'(y)| = -2ye^{y^2} I_{(-\infty,0)}(y)$$

- **Un Camino Alterno**:Un camino alterno es trabajar directamente y buscar una solución analítica. Veamos Sea X v.a. con función de densidad dada por

$$f_X(x) = \frac{1}{18} x^2 I_{(-3,3)}(x)$$

Asume se desea la función de densidad de $Y = X^2$.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) = P(|X| \leq y^{\frac{1}{2}}) \\ &= P(-y^{\frac{1}{2}} \leq X \leq y^{\frac{1}{2}}) = P(X \leq y^{\frac{1}{2}}) - P(X \leq -y^{\frac{1}{2}}) \\ &= F_X(y^{\frac{1}{2}}) - F_X(-y^{\frac{1}{2}}) \end{aligned}$$

De lo anterior:

$$\begin{aligned} f_Y(y) = \frac{\partial F_Y(y)}{\partial y} &= \frac{\partial \left[F_X(y^{\frac{1}{2}}) - F_X(-y^{\frac{1}{2}}) \right]}{\partial y} \\ &= f_X(y^{\frac{1}{2}}) \frac{1}{2} y^{-\frac{1}{2}} + f_X(-y^{\frac{1}{2}}) \frac{1}{2} y^{-\frac{1}{2}} \\ &= \frac{1}{18} y^{\frac{1}{2}} I_{(0,9)}(y) \end{aligned}$$

Transformación de Vectores Aleatorios

1. Iniciemos con dos Ejemplos **Distribuciones del Máximo y del Mínimo**: Asume que se tiene una muestra aleatoria $X_1, X_2, \ldots, X_n$ de una población con función de distribución acumulada $F_X(\cdot)$, y se desea conocer la distribución del máximo y del mínimo valor observados.
Notación:

   - $Y = X_{(n)} = \max\{X_1, X_2, \ldots, X_n\}$ (Estadística de orden n) y
   - $Z = X_{(1)} = \min\{X_1, X_2, \ldots, X_n\}$ (Estadística de orden 1).

   Distribución del Máximo:

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X_1 \leq y, X_2 \leq y, \ldots, X_n \leq y) \\ &= P(X_1 \leq y)P(X_2 \leq y)\ldots P(X_n \leq y) \\ &= F_X^n(y) \end{aligned}$$

   En el caso X v.a. continua con función de densidad $f_X(\cdot)$, se tiene:

$$f_Y(y) = nF_X^{n-1}(y)f_X(y).$$

   Distribución del Mínimo

$$\begin{aligned} F_Z(z) &= P(Z \leq z) = 1 - P(Z > z) \\ &= 1 - P(X_1 > z, X_2 > z, \ldots, X_n \leq z) \\ &= 1 - P(X_1 > z)P(X_2 > z)\ldots P(X_n \leq z) \\ &= 1 - [1 - P(X_1 \leq z)][1 - P(X_2 \leq z)]\ldots[1 - P(X_n \leq z)] \\ &= 1 - [1 - F_X(z)]^n \end{aligned}$$

   De nuevo, en el caso X v.a. continua con función de densidad $f_X(\cdot)$, se tiene:

$$f_Z(z) = n[1 - F_X(z)]^{n-1}f_X(z).$$

   Aplicación:

   - Un caso aplicado de lo anterior es asumir que se tiene un sistema en línea que depende de n componentes (ejemplo resistencias) construídos con la misma tecnología.
   - El sistema falla al fallar al menos un componente.

- Es decir, si los tiempos de sobrevida de los n componentes son $X_1, X_2, \ldots, X_n$ v.a (i.i.d.).
- El tiempo de sobrevida del sistema es

$$Z = X_{(1)} = \text{mín}\{X_1, X_2, \ldots, X_n\}.$$

- Si asumimos X_j tiene distribución exponencial, con:

$$f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} I_{(0,\infty)}(x).$$

$$F_X(x) = \left[1 - e^{-\frac{x}{\beta}}\right] I_{(0,\infty)}(x).$$

  El tiempo antes de la primera falla del sistema tiene una distribución con función de de densidad:

$$\begin{aligned} f_Z(z) &= n[1 - F_X(z)]^{n-1} f_X(z) \\ &= \frac{n}{\beta} e^{-\frac{n}{\beta} z} \end{aligned}$$

  Nota: Lo anterior, explicado con un ejemplo sencillo, es como sigue.

  - Si se tiene una resistencia cuyo tiempo antes de la primera falla, X, tiene distribución exponencial, $X \sim exp(12)$ (en meses). Se espera que X falle a los dos 12 meses.
  - Si se monta un sistema (en linea) con 6 resistencias de las planteadas en el ítem anterior, entonces $Z = X_{(1)} \sim \exp\left(\frac{\beta}{6}\right) = \exp(2)$, de donde se espera que el sistema falle a los dos meses.

2. **Transformaciones 1-1**. Lo que sigue es pensar en transformaciones vectoriales 1-1, concepto que debes entender-recordar. Veamos dos ejemplos.

Ejemplo 4: Asumiendo se tiene dos variables aleatorias X, Y, con las que formamos el vector $\mathbf{X} = (X, Y)$, y conocemos la distribución de $\mathbf{X}$.

- A las variables X, Y, las llamaremos variables de inicio.
- Definimos (es de interés) las v.a. $Z := r_1(\mathbf{X}) = r_1(X, Y) = X + Y$ y $W := r_2(\mathbf{X}) = r_2(X, Y) = X - Y$, que llamaremos variables de llegada.
- Observa $\mathbf{Z} = (Z, W) = (r_1(\mathbf{X}), r_2(\mathbf{X})) = r(\mathbf{X})$.

- Se tiene: $X = \frac{Z+W}{2}$ y $Y = \frac{Z-W}{2}$.
- Podemos devolvernos, $\mathbf{X} = (X, Y) = (s_1(\mathbf{Z}), s_2(\mathbf{Z})) = s(\mathbf{Z})$.
- Para la transformación

$$r : (x, y) \to (z, w),$$

hemos hallado la transformación

$$s : (z, w) \to (x, y).$$

Entonces la transformación $r : \mathbb{R}^2 \to \mathbb{R}^2$ es 1-1.

Ejemplo 5: Bajo los mismos supuestos del ejemplo anterior.

- Asumiendo $X > 0$ y $Y > 0$.
- $Z = \frac{X}{Y} = r_1(\mathbf{X})$ y $W = XY = r_2(\mathbf{X})$, entonces $r(\mathbf{X}) = (r_1(\mathbf{X}), r_2(\mathbf{X}))$.
- Observa:

$$X = [ZW]^{\frac{1}{2}} \quad \text{y} \quad Y = \left[\frac{W}{Z}\right]^{\frac{1}{2}}.$$

- De nuevo: la transformación $r : (0, \infty)^2 \to (0, \infty)^2$ es 1-1

3. **De forma sencilla**: lo que sigue es trabajar transformación de vectores, para transformaciones 1-1. Análogo a lo hecho en v.a. tenemos un vector aleatorio $\mathbb{X}$,

$$\mathbb{X} : (\Omega, \mathfrak{F}, P) \quad \to \quad (\mathbb{R}, \mathfrak{B}, P_X)$$

- Se conoce la distribución (conjunta) del vector $\mathbb{X}$, $P_{\mathbb{X}}$.
- El interés recae sobre una transformación de $\mathbb{X}$, esto es $\mathbb{Y} = \mathbb{T}(\mathbb{X})$.
- El propósito es hallar $P_{\mathbb{Y}}$.

Caso Discreto: En general tenemos:

- Un vector aleatorio discreto $\mathbb{X} = (X_1, X_2, \ldots, X_n)$, cuya función de masa de probabilidad conjunta $f_{\mathbb{X}}(\mathbf{x}) = P(\mathbb{X} = \mathbf{x})$, conocida.
- Si se define $\mathbb{Y} = \mathbb{T}(\mathbb{X})$,

$$\mathbb{T} : \mathbb{R}^n \to \mathbb{R}^p$$

Lo que debes pensar es:

$$\mathbb{Y} = (Y_1, Y_2, \dots, Y_p) = \mathbb{T}(\mathbb{X}) = (T_1(\mathbb{X}), T_2(\mathbb{X}), \dots, T_p(\mathbb{X})).$$

En general $p \leq n$. Dado lo anterior, en el caso de un vector aleatorio discreto el truco es devolverse al espacio inicial (idéntico a los hecho con v.a. discretas), lo que significa:

- Es decir, la función de masa de probabilidad de $\mathbb{Y}$, $f_{\mathbb{Y}}(y)$, está dada por:

$$f_{\mathbb{Y}}(y) = P(\mathbb{Y} = y) = \sum_{\mathbf{x}:\mathbb{T}(\mathbf{x})=y} P(\mathbb{X} = \mathbf{x}) = \sum_{\mathbf{x}:\mathbb{T}(\mathbf{x})=y} f_{\mathbb{X}}(\mathbf{x})$$

Caso Continuo: En el caso de los vectores aleatorios continuos se busca tener transformaciones 1-a-1, ítem que ya ha sido explicado.

- Se tiene un vector aleatorio $\mathbb{X} = (X_1, X_2, \dots, X_n)$, cuya función de densidad de probabilidad conjunta $f_{\mathbb{X}}(\mathbf{x})$, conocida, y una transformación $\mathbb{Y} = (Y_1, Y_2, \dots, Y_n) = \mathbb{T}(\mathbb{X}) = (T_1(\mathbb{X}), T_2(\mathbb{X}), \dots, T_n(\mathbb{X}))$
- Dado que $\mathbb{T}$ es 1-a-1, existe una transformación inversa $\mathbb{S}(\mathbb{Y}) = (S_1(\mathbb{Y}), S_2(\mathbb{Y}), \dots, S_n(\mathbb{Y})) = (X_1, X_2, \dots, X_n)$.
- A partir de lo anterior es posible definir (Jacobiano):

$$J = \det\left(\frac{\partial \mathbb{S}}{\partial \mathbf{y}}\right) = \det\begin{pmatrix} \frac{\partial S_1}{\partial y_1} & \frac{\partial S_1}{\partial y_2} & \cdots & \frac{\partial S_1}{\partial y_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial S_n}{\partial y_1} & \frac{\partial S_n}{\partial y_2} & \cdots & \frac{\partial S_n}{\partial y_n} \end{pmatrix}$$

- La función de densidad conjunta del vector $\mathbb{Y} = \mathbb{T}(\mathbb{X})$, está dada por:

$$f_{\mathbb{Y}}(\mathbf{y}) = \begin{cases} f_{\mathbb{X}}(\mathbf{s})|J| & \text{si} \quad \mathbf{s} \in \Theta \\ 0 & \text{en otro caso} \end{cases}.$$

Donde Θ es el recorrido de la transformación $\mathbb{T}$ y $|\cdot|$ indica valor absoluto. Entencer lo anterior es más claro viendo un ejemplo.

4. Ejemplo 6: Asume que se tiene un vector aleatorio $\mathbb{X} = (X_1, X_2)$ con f.d. dada por:

$$f_{\mathbb{X}}(\mathbf{x}) = f_{X_1,X_2}(x_1, x_2) = 4x_1x_2I_{(0,1)\times(0,1)}(x_1, x_2). \tag{7}$$

Asume se desea la función de densidad de $Y = X_1 + X_2$, pero ésta transformación no es 1-a-1. Por lo anterior podemos seguir el siguiente **camino**:

- $\mathbb{Y} = (Y, Z) = (X_1 + X_2, X_1 - X_2)$
- De lo anterior: $\mathbb{X} = (X_1, X_2) = \left(\frac{Y+Z}{2}, \frac{Y-Z}{2}\right)$,
- Es decir la transformación $\mathbb{Y}$ es 1-a-1, y, los resultados presentados nos permiten llegar a $f_{\mathbb{Y}}(\mathbf{y})$.
- Luego hallaremos $f_Y(y)$, como una marginal.

El camino a seguir es hallar $f_{\mathbb{Y}}(\mathbf{y})$, paso-a-paso es como sigue:

- $y = T_1(\mathbf{x}) = x_1 + x_2 \in (0, 2)$ y $z = T_2(\mathbf{x}) = x_1 - x_2 \in (-1, 1)$.
- Observa $z = x_1 - x_2 \leq x_1 \leq x_1 + x_2 = y$.
- $S_1(\mathbf{y}) = \frac{y+z}{2}$ y $S_2(\mathbf{y}) = \frac{y-z}{2}$.
- De lo anterior
$$J = \det \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} = -\frac{1}{2}$$
- Hemos llegado:
$$f_{\mathbb{Y}}(\mathbf{y}) = \begin{cases} f_{\mathbb{X}}(\mathbf{s})\,|J| & \text{si} \quad 0 < y < 2,\ -1 < z < 1,\ z \leq y \\ 0 & \text{en otro caso.} \end{cases}$$
$$= \begin{cases} \frac{1}{2}(y+z)(y-z) & \text{si} \quad 0 < y < 2,\ -1 < z < 1,\ z \leq y \\ 0 & \text{en otro caso.} \end{cases}$$

5. Ejemplo 7: Se tiene un vector aleatorio $\mathbb{X} = (X_1, X_2)$ con f.d. conocida $f_{\mathbb{X}}(\mathbf{x})$.
   - Si se desea la función de densidad de $Y = X_1 * X_2$, el camino a seguir es definir $Z = X_1$, de donde se tiene $\mathbb{Y} = (Y, Z)$ y $\mathbb{X} = (Z, \frac{Y}{Z})$.
   - Si se desea la función de densidad de $Y = \frac{X_1}{X_2}$, el camino a seguir es definir $Z = X_2$, de donde se tiene $\mathbb{Y} = (Y, Z)$ y $\mathbb{X} = (YZ, Z)$.
6. **Lectura**: Paradoja de Borel - Kolmorov. Pg 171, Degroot(1986)

Fin - Clase del día Miércoles Dic. 18