Variables Aleatorias y Distribuciones de Probabilidad

UCR – ECCI

CI-0115 Probabilidad y Estadística

Prof. Kryscia Daviana Ramírez Benavides

Variable Aleatoria

- Una **variable aleatoria** es una función que asocia un número real con cada elemento del espacio muestral.
- Se utiliza una letra mayúscula, por ejemplo X, para denotar una variable aleatoria, y su correspondiente minúscula, x en este caso, para denotar a cada uno de sus valores.
  - Cada valor de X representa un evento que es un subconjunto del espacio muestral para el experimento dado.

Variable Aleatoria (cont.)

- **Ejemplo:** Se sacan dos bolas de manera sucesiva sin reemplazo de una urna que contiene cuatro bolas rojas y tres negras. Los posibles resultados y los valores *x* de la variable aleatoria *X*, donde *X* es el número de bolas rojas, son:

Espacio Muestral	***x***
RR	2
RN	1
NR	1
NN	0

Variable Aleatoria (cont.)

- Si un espacio muestral contiene un número finito de posibilidades, o una serie interminable con tantos elementos como números enteros existen, se llama **espacio muestral discreto**.
  - Una variable aleatoria se llama **variable aleatoria discreta** si se puede contar su conjunto de resultados posibles.
- Si un espacio muestral contiene un número infinito de posibilidades igual al número de puntos en un segmento de línea, se llama **espacio muestral continuo**.
  - Una variable aleatoria se llama **variable aleatoria continua** si puede tomar valores en una escala continua.

Variable Aleatoria (cont.)

- En la mayor parte de los problemas prácticos, las **variables aleatorias continuas representan datos medidos**, como todos los posibles pesos, alturas, temperaturas, distancias o periodos de vida.
- Mientras que las **variables aleatorias discretas representan datos contados**, como el número de artículos defectuosos en una muestra de *k* artículos o el número de accidentes de carretera por año en un país.

Distribuciones Discretas de Probabilidad

- El conjunto de pares ordenados $(x, f(x))$ es una **distribución de probabilidad** o **función de masa de probabilidad** de la variable aleatoria discreta X si, para cada resultado posible x,

$$1.\ f(x) \geq 0$$

$$2.\ \sum_{x} f(x) = 1$$

$$3.\ P(X = x) = f(x)$$

Distribuciones Discretas de Probabilidad (cont.)

- **Ejemplo #1:** Hay tres computadoras defectuosas de un conjunto de ocho similares en una tienda. Si una escuela hace una compra al azar de dos de estas computadoras, ¿cuál es la distribución de probabilidad para el número de defectuosas?
  - Sea *X* una variable aleatoria cuyos valores de *x* son la cantidad posible de computadoras defectuosas que compra la escuela. Entonces *x* puede ser cualquiera de los números 0, 1 y 2.

$$f(0)=P(X=0)=\frac{\binom{3}{0}\binom{5}{2}}{\binom{8}{2}}=\frac{10}{28} \quad f(1)=P(X=1)=\frac{\binom{3}{1}\binom{5}{1}}{\binom{8}{2}}=\frac{15}{28} \quad f(2)=P(X=2)=\frac{\binom{3}{2}\binom{5}{0}}{\binom{8}{2}}=\frac{3}{28}$$

x	0	1	2
$f(x)$	$\frac{10}{28}$	$\frac{15}{28}$	$\frac{3}{28}$

Distribuciones Discretas de Probabilidad (cont.)

- **Ejemplo #2:** Si una agencia de autos vende 50% de su inventario de cierto vehículo extranjero equipado con bolsas de aire, encontrar una fórmula para la distribución de probabilidad del número de autos con bolsas de aire entre los siguientes cuatro vehículos que venda la agencia.
  - Sea X una variable aleatoria cuyos valores de x son la cantidad posible de los cuatro autos vendidos que tienen bolsa de aire. Entonces x puede ser cualquiera de los números 0, 1, 2, 3 y 4.

$$f(x) = P(X = x) = \frac{\binom{4}{x}}{2^4} = \frac{\binom{4}{x}}{16} \quad \text{para } x = 0,1,2,3,4$$

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Distribuciones Discretas de Probabilidad (cont.)

- Hay muchos problemas en donde se desea calcular la probabilidad de que el valor observado de una variable aleatoria *X* sea menor o igual que algún número real *x*.
- La **distribución acumulada** o **función de distribución** $F(x)$ de una variable aleatoria discreta *X* con distribución de probabilidad $f(x)$ es

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad -\infty < x < +\infty$$

- Como consecuencia

$$f(x) = F(x) - F(x-1)$$

Distribuciones Discretas de Probabilidad (cont.)

- **Ejemplo:** Para el ejemplo #2 anterior, se calcula las distribuciones acumuladas.

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$F(0) = f(0) = \frac{1}{16}$$

$$F(1) = f(0) + f(1) = \frac{5}{16}$$

$$F(2) = f(0) + f(1) + f(2) = \frac{11}{16}$$

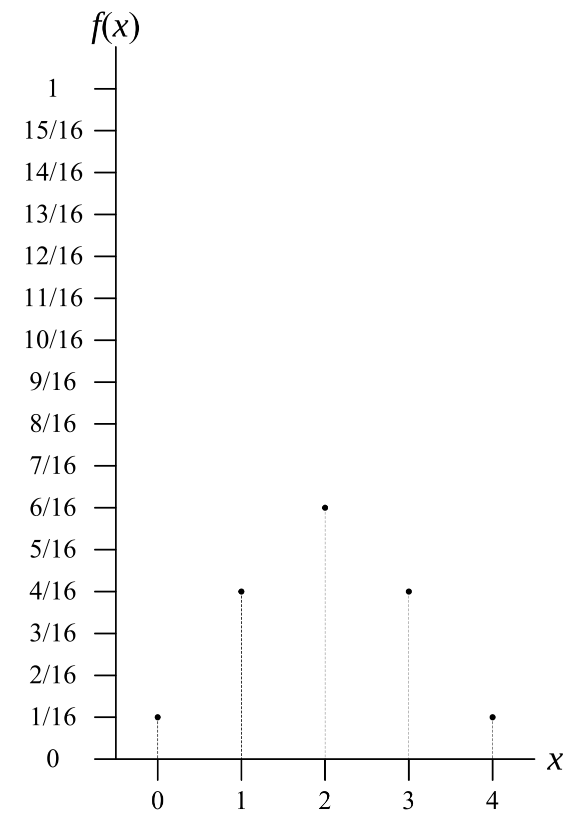
$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16}$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = \frac{16}{16} = 1$$

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/16 & 0 \le x < 1 \\ 5/16 & 1 \le x < 2 \\ 11/16 & 2 \le x < 3 \\ 15/16 & 3 \le x < 4 \\ 1 & x \ge 4 \end{cases}$$

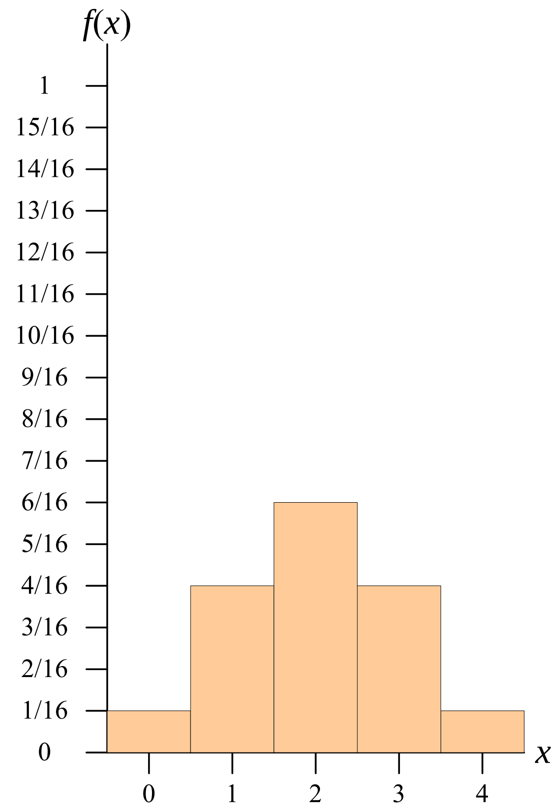
Distribuciones Discretas de Probabilidad (cont.)

- Es útil ver de forma gráfica una distribución de probabilidad.
- Los puntos $(x, f(x))$ se pueden graficar mediante un gráfico de barras. El gráfico permite observar de forma fácil qué valores de X tienen más probabilidad de ocurrencia.



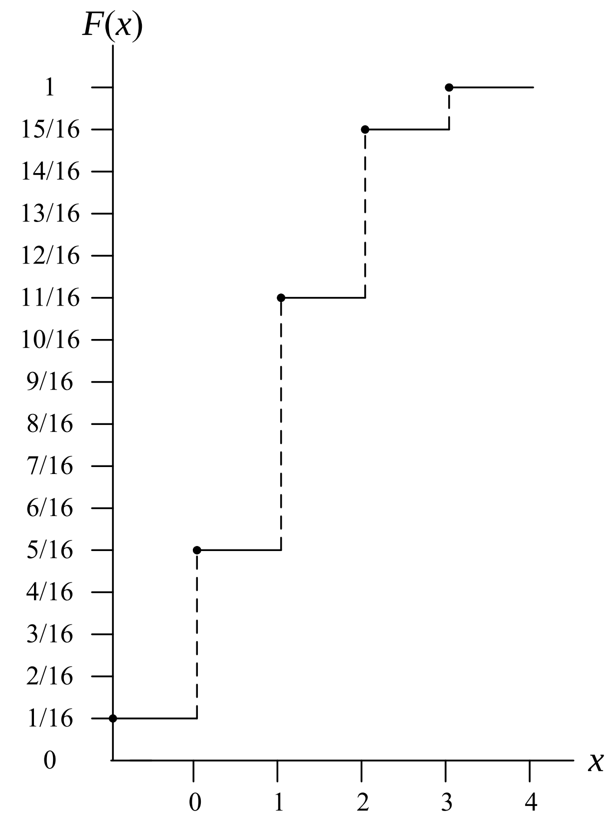
Distribuciones Discretas de Probabilidad (cont.)

- Los puntos (x, $f(x)$) se pueden graficar mediante un gráfico llamado **histograma de probabilidad**.
- Este se utiliza para las distribuciones de probabilidad de v.a. continuas, ya que el ancho de los rectángulos representan intervalos y esto áreas.



Distribuciones Discretas de Probabilidad (cont.)

- También es útil ver de forma gráfica una distribución acumulada.
- Los puntos (x, $F(x)$) se pueden graficar mediante un **diagrama de frecuencias acumuladas**.



Distribuciones Continuas de Probabilidad

- Una variable aleatoria continua tiene una probabilidad cero de tomar exactamente cualquiera de sus valores. Por lo cual, su distribución de probabilidad no se puede dar en forma tabular, pero se puede establecer como una fórmula.
- En este tipo de variables se trata con un intervalo en lugar de un valor puntual de nuestra variable aleatoria.
- Se trata el cálculo de probabilidades para varios intervalos de variables aleatorias continuas como $P(a < X < b)$, $P(X > a)$, etc.

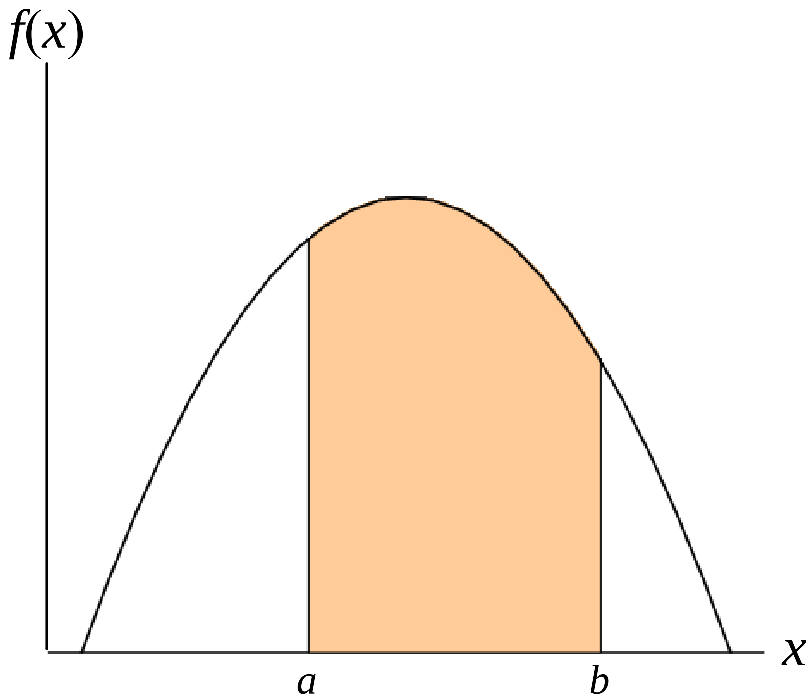
Distribuciones Continuas de Probabilidad (cont.)

- Cuando *X* es continua, se puede notar que

$$P(a < X \leq b) = P(a < X < b) + P(X = b) = P(a < X < b)$$

  - No importa si incluimos o no un extremo del intervalo. Esto no es cierto cuando *X* es discreta.
- La fórmula que se utiliza para una variable aleatoria continua, será función de los valores numéricos de la variable *X* y, como tal, se representará mediante la notación $f(x)$ y se le llama función de densidad de probabilidad, o simplemente función de densidad de *X*.
  - Se utilizan áreas para representar probabilidades y son valores numéricos positivos, la función de densidad debe estar completamente por arriba del eje x.

Distribuciones Continuas de Probabilidad (cont.)



$$P(a < X < b)$$

Distribuciones Continuas de Probabilidad (cont.)

- La función $f(x)$ es una **función de densidad de probabilidad** (fdp) para la variable aleatoria continua X, definida en el conjunto de números reales R, si

$$1.\ f(x) \geq 0 \quad x \in R$$

$$2.\ \int_{-\infty}^{+\infty} f(x)\,dx = 1$$

$$3.\ P(a < X < b) = \int_{a}^{b} f(x)\,dx$$

Distribuciones Continuas de Probabilidad (cont.)

- **Ejemplo:** El error en la temperatura de reacción, en ºC, para un experimento de laboratorio controlado es una variable aleatoria continua X que tiene *fdp*

$$f(x) = \begin{cases} \dfrac{x^2}{3} & -1 < x < 2 \\ 0 & \text{otro caso} \end{cases}$$

- Verificar la condición 2 de la definición y encontrar $P(0 < X \leq 1)$.

Distribuciones Continuas de Probabilidad (cont.)

- **Ejemplo:** Experimento de laboratorio controlado es una variable aleatoria continua X. Verificar la condición 2 de la definición y encontrar $P(0 < X \leq 1)$.

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3}dx = \left.\frac{x^3}{9}\right|_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = \frac{9}{9} = 1$$

$$P(0 < X \leq 1) = \int_{0}^{1} \frac{x^2}{3}dx = \left.\frac{x^3}{9}\right|_{0}^{1} = \frac{1}{9} - \frac{0}{9} = \frac{1}{9}$$

Distribuciones Continuas de Probabilidad (cont.)

- La **distribución acumulada de probabilidad** $F(x)$ de una variable aleatoria continua X con *fdp* $f(x)$ es

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt \qquad -\infty < x < +\infty$$

- Como consecuencia

$$P(a < X < b) = F(b) - F(a)$$

$$f(x) = \frac{dF(x)}{dx}$$

si existe la derivada

Distribuciones Continuas de Probabilidad (cont.)

- **Ejemplo:** Para la *fdp* del ejemplo anterior, encontrar $F(x)$, y utilizarla para evaluar $P(0 < X \leq 1)$.

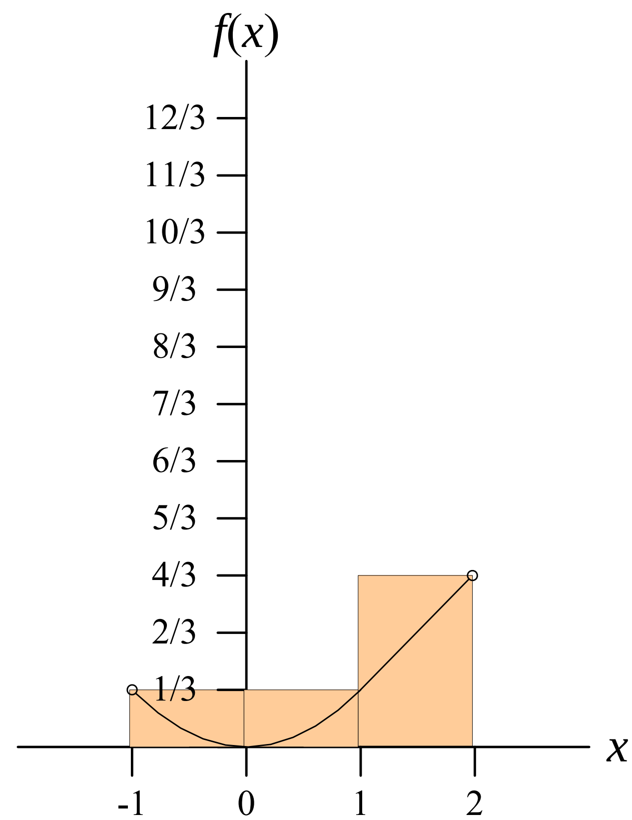
$$F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-1}^{x} \frac{t^2}{3}dt = \left.\frac{t^3}{9}\right|_{-1}^{x} = \frac{x^3+1}{9}$$

$$F(x) = \begin{cases} 0 & x \leq -1 \\ \dfrac{x^3+1}{9} & -1 < x < 2 \\ 1 & x \geq 2 \end{cases}$$

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9}$$

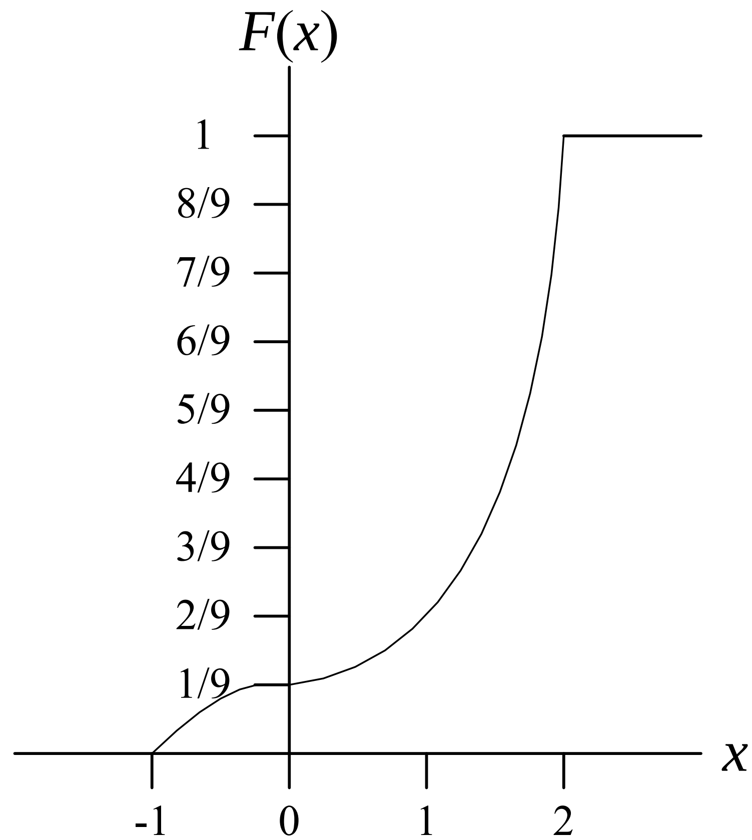
Distribuciones Continuas de Probabilidad (cont.)

- Los puntos (x, $f(x)$) se pueden graficar mediante un gráfico llamado **histograma de probabilidad**.
- El ancho de los rectángulos representan intervalos y esto áreas.



Distribuciones Continuas de Probabilidad (cont.)

- También es útil ver de forma gráfica una distribución acumulada.
- Los puntos (x, $F(x)$) se pueden graficar mediante un **diagrama de frecuencias acumuladas**.



Distribuciones de Probabilidad

- Otro cálculos de probabilidad:

$$P(X \geq x) = 1 - P(X < x)$$
$$P(X > x) = 1 - P(X \leq x)$$

  - En el caso continuo no importa si incluimos o no un extremo del intervalo.
  - En el caso discreto sí importa si incluimos o no un extremo del intervalo.

Distribuciones de Probabilidad Conjunta

- Hasta el momento se ha visto el estudio de variables aleatorias y sus distribuciones de probabilidad que se restringe a espacios muestrales unidimensionales, en los que se registra los resultados de un experimento como valores que toma una sola variable aleatoria.
- Habrá situaciones, sin embargo, donde se puede encontrar que es deseable registrar los resultados simultáneos de diversas variables aleatorias.

Distribuciones de Probabilidad Conjunta (cont.)

- La función $f(x,y)$ es una **distribución de probabilidad conjunta** o **función de masa de probabilidad** de las variables aleatorias discretas X y Y si

$$1.\ f(x,y) \geq 0 \quad \forall (x,y)$$

$$2.\ \sum_x \sum_y f(x,y) = 1$$

$$3.\ P(X = x, Y = y) = f(x,y)$$

- Para cualquier región A en el plano xy, $P[(X,Y) \in \sum\sum f(x,y)]$.

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo:** Se seleccionan al azar dos repuestos para un bolígrafo de una caja que contiene tres repuestos azules, dos rojos y tres verdes. Si X es el número de repuestos azules y Y es el número de repuestos rojos que se seleccionan, encontrar la función de probabilidad conjunta $f(x,y)$ y $P[(X,Y) \in A]$, donde A es la región $\{(x,y) \mid x + y \leq 1\}$.
  - El número total de formas posibles de seleccionar dos repuestos de ocho es C(8,2) = 28.
  - Los posibles pares de valores (x,y) son: (0,0), (0,1), (1,0), (1,1), (0,2) y (2,0).
    - El valor (0,0) corresponde a seleccionar al azar 0 azules, 0 rojos y 2 verdes; o sea, C(3,0)C(2,0)C(3,2) = 3.
      - Entonces, $f(0,0)$ representa la probabilidad de seleccionar cero azules y rojos, y dos verdes; o sea, $f(0,0) = 3/28$.

Distribuciones de Probabilidad Conjunta (cont.)

$f(x,y)$		**x = 0**	**x = 1**	**x = 2**	**Totales por Fila**
y	**0**	3/28	9/28	3/28	**15/28**
	1	6/28	6/28	-	**12/28**
	2	1/28	-	-	**1/28**
Totales por Columna		**10/28**	**15/28**	**3/28**	**1**

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo:** Se seleccionan al azar dos repuestos para un bolígrafo de una caja que contiene tres repuestos azules, dos rojos y tres verdes. Si X es el número de repuestos azules y Y es el número de repuestos rojos que se seleccionan, encontrar la función de probabilidad conjunta $f(x,y)$ y $P[(X,Y) \in A]$, donde A es la región $\{(x,y) \mid x + y \leq 1\}$.
  - La probabilidad conjunta, entonces, se puede representar mediante la fórmula:

$$f(x,y) = \frac{\binom{3}{x}\binom{2}{y}\binom{3}{2-x-y}}{\binom{8}{2}} \qquad \begin{array}{c} x = 0,1,2 \\ y = 0,1,2 \\ 0 \leq x + y \leq 2 \end{array}$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo:** Se seleccionan al azar dos repuestos para un bolígrafo de una caja que contiene tres repuestos azules, dos rojos y tres verdes. Si X es el número de repuestos azules y Y es el número de repuestos rojos que se seleccionan, encontrar la función de probabilidad conjunta $f(x,y)$ y $P[(X,Y) \in A]$, donde A es la región $\{(x,y) \mid x + y \leq 1\}$.
  - El cálculo de $P[(X,Y) \in A]$ se realiza de la siguiente manera:

$$
\begin{aligned}
P[(X,Y) \in A] &= P(X + Y \leq 1) \\
&= f(0,0) + f(1,0) + f(0,1) \\
&= \frac{3}{28} + \frac{6}{28} + \frac{9}{28}
\end{aligned}
$$

$$P[(X,Y) \in A] = \frac{18}{28} = \frac{9}{14}$$

Distribuciones de Probabilidad Conjunta (cont.)

- La función $f(x,y)$ es una **función de densidad de probabilidad conjunta** de las variables aleatorias continuas X y Y si

$$1.\ f(x,y) \geq 0 \quad \forall (x,y)$$

$$2. \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1$$

$$3. P[(X,Y) \in A] = \int_A \int f(x,y)\,dx\,dy$$

para cualquier región A en el plano xy.

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo:** Una fábrica de dulces distribuye cajas de chocolates con un surtido de cremas, chiclosos y nueces, cubiertas de con chocolate claro y oscuro. Para una caja seleccionada al azar, sean X y Y, respectivamente, las proporciones de chocolates claro y oscuro que son cremas, y suponga que la función de densidad conjunta es

$$f(x,y)=\begin{cases}\frac{2}{5}(2x+3y) & 0\le x\le 1, 0\le y\le 1\\ 0 & \text{otro caso}\end{cases}$$

Verificar la condición 2 de la definición y encontrar $P[(X,Y) \in A]$, donde A está la región $\{(x,y) \mid 0 < x < ½, ¼ < y < ½\}$.

Distribuciones de Probabilidad Conjunta (cont.)

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = \int_0^1\int_0^1 \frac{2}{5}(2x+3y)dxdy$$

$$= \int_0^1 \left.\frac{2x^2}{5}+\frac{6xy}{5}\right|_{x=0}^{x=1} dy$$

$$= \int_0^1 \left(\frac{2}{5}+\frac{6y}{5}\right)dy$$

$$= \left.\frac{2y}{5}+\frac{3y^2}{5}\right|_{y=0}^{y=1}$$

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = \frac{2}{5}+\frac{3}{5}=1$$

Distribuciones de Probabilidad Conjunta (cont.)

$$P[(X,Y)\in A] = P\left(0 < X < \frac{1}{2}, \frac{1}{4} < Y < \frac{1}{2}\right)$$

$$= \int_{1/4}^{1/2}\int_{0}^{1/2} \frac{2}{5}(2x+3y)\,dxdy$$

$$= \int_{1/4}^{1/2} \left.\frac{2x^2}{5} + \frac{6xy}{5}\right|_{x=0}^{x=1/2} dy$$

$$= \int_{1/4}^{1/2}\left(\frac{1}{10} + \frac{3y}{5}\right)dy$$

$$= \left.\frac{y}{10} + \frac{3y^2}{10}\right|_{y=1/4}^{y=1/2}$$

$$P[(X,Y)\in A] = \frac{1}{10}\left[\left(\frac{1}{2}+\frac{3}{4}\right) - \left(\frac{1}{4}+\frac{3}{16}\right)\right] = \frac{13}{160}$$

Distribuciones de Probabilidad Conjunta (cont.)

- Las **distribuciones marginales** de *X* sola y *Y* sola son:
  - Para el caso discreto, $g(x)$ y $h(y)$ son las distribuciones de probabilidad de las variables aleatorias *X* y *Y*.

$$f_x(x) = g(x) = \sum_y f(x,y) \qquad f_y(y) = h(y) = \sum_x f(x,y)$$

  - Para el caso continuo, $g(x)$ y $h(y)$ son las funciones de densidad de las variables aleatorias *X* y *Y*

$$f_x(x) = g(x) = \int_{-\infty}^{+\infty} f(x,y)dy \qquad f_y(y) = h(y) = \int_{-\infty}^{+\infty} f(x,y)dx$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #1:** Obtener las distribuciones marginales de *X* sola y *Y* sola del ejemplo de la distribución conjunta de las variables discretas.

$$P(X=0)=f_x(0)=g(0)=\sum_{y=0}^{2} f(0,y)=f(0,0)+f(0,1)+f(0,2)$$

$$P(X=0)=f_x(0)=g(0)=\frac{3}{28}+\frac{6}{28}+\frac{1}{28}=\frac{10}{28}=\frac{5}{14}$$

$$P(X=1)=f_x(1)=g(1)=\sum_{y=0}^{2} f(1,y)=f(1,0)+f(1,1)+f(1,2)$$

$$P(X=1)=f_x(1)=g(1)=\frac{9}{28}+\frac{6}{28}+0=\frac{15}{28}$$

$$P(X=2)=f_x(2)=g(2)=\sum_{y=0}^{2} f(2,y)=f(2,0)+f(2,1)+f(2,2)$$

$$P(X=2)=f_x(2)=g(2)=\frac{3}{28}+0+0=\frac{3}{28}$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #1:** Obtener las distribuciones marginales de *X* sola y *Y* sola del ejemplo de la distribución conjunta de las variables discretas.

$$P(Y=0)=f_y(0)=h(0)=\sum_{x=0}^{2} f(x,0)=f(0,0)+f(1,0)+f(2,0)$$

$$P(Y=0)=f_y(0)=h(0)=\frac{3}{28}+\frac{9}{28}+\frac{3}{28}=\frac{15}{28}$$

$$P(Y=1)=f_y(1)=h(1)=\sum_{x=0}^{2} f(x,1)=f(0,1)+f(1,1)+f(2,1)$$

$$P(Y=1)=f_y(1)=h(1)=\frac{6}{28}+\frac{6}{28}+0=\frac{12}{28}=\frac{3}{7}$$

$$P(Y=2)=f_y(2)=h(2)=\sum_{x=0}^{2} f(x,2)=f(0,2)+f(1,2)+f(2,2)$$

$$P(Y=2)=f_y(2)=h(2)=\frac{1}{28}+0+0=\frac{1}{28}$$

Distribuciones de Probabilidad Conjunta (cont.)

$f(x,y)$		x			Totales por Fila
		0	**1**	**2**	
y	**0**	3/28	9/28	3/28	**15/28**
	1	6/28	6/28	-	**12/28**
	2	1/28	-	-	**1/28**
Totales por Columna		**10/28**	**15/28**	**3/28**	**1**

Distribuciones de Probabilidad Conjunta (cont.)

- Se puede observar que los resultados de las distribuciones marginales corresponden a los totales de las columnas y filas respectivos de los valores $f(x,y)$.
- Y la suma individual de los valores de $g(x)$ y $h(x)$ dan 1, cumpliendo la definición de distribución de probabilidad.

x	**0**	**1**	**2**
$g(x)$	5/14	15/28	3/28

y	**0**	**1**	**2**
$h(y)$	15/28	3/7	1/28

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** Obtener las distribuciones marginales de *X* sola y *Y* sola del ejemplo de la distribución conjunta de las variables continuas.

$$f_x(x) = g(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy$$

$$= \int_0^1 \frac{2}{5}(2x+3y)\,dy$$

$$= \left.\frac{4xy}{5} + \frac{6y^2}{10}\right|_{y=0}^{y=1}$$

$$f_x(x) = g(x) = \frac{4x+3}{5}$$

$$f_y(y) = h(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx$$

$$= \int_0^1 \frac{2}{5}(2x+3y)\,dx$$

$$= \left.\frac{2x^2}{5} + \frac{6xy}{5}\right|_{x=0}^{x=1}$$

$$f_y(y) = h(y) = \frac{2+6y}{5}$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** Obtener las distribuciones marginales de X sola y Y sola del ejemplo de la distribución conjunta de las variables continuas.

$$g(x) = \begin{cases} \dfrac{4x+3}{5} & 0 \le x \le 1 \\ 0 & \text{otro caso} \end{cases}$$

$$h(y) = \begin{cases} \dfrac{2+6y}{5} & 0 \le y \le 1 \\ 0 & \text{otro caso} \end{cases}$$

Distribuciones de Probabilidad Conjunta (cont.)

- Se puede observar que las distribuciones marginales *g*(*x*) y *h*(*x*) son distribuciones de probabilidades de las variables individuales *X* y *Y*, cumpliendo la definición de distribución de probabilidad (*fdp*).

$$\int_{-\infty}^{+\infty} g(x)dx = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dydx = 1$$

$$P(a < X < b) = P(a < X < b, -\infty < Y < +\infty)$$

$$P(a < X < b) = \int_{a}^{b}\int_{-\infty}^{+\infty} f(x,y)dydx = \int_{a}^{b} g(x)dx$$

Distribuciones de Probabilidad Conjunta (cont.)

- Se puede observar que las distribuciones marginales $g(x)$ y $h(x)$ son distribuciones de probabilidades de las variables individuales X y Y, cumpliendo la definición de distribución de probabilidad (*fdp*).

$$\int_{-\infty}^{+\infty} h(y)dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$$

$$P(a < Y < b) = P(-\infty < X < +\infty, a < Y < b)$$

$$P(a < Y < b) = \int_{a}^{b}\int_{-\infty}^{+\infty} f(x,y)dxdy = \int_{a}^{b} h(y)dy$$

Distribuciones de Probabilidad Conjunta (cont.)

- Sean X y Y dos variables aleatorias, discretas o continuas.
  - La **distribución de probabilidad condicional** de la variable aleatoria Y, dado que $X = x$, es:

$$f(y|x) = \frac{f(x,y)}{g(x)}, \quad g(x) > 0$$

  - La **distribución de probabilidad condicional** de la variable aleatoria X, dado que $Y = y$, es:

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad h(y) > 0$$

Distribuciones de Probabilidad Conjunta (cont.)

- Si se desea encontrar la probabilidad de que la variable aleatoria discreta X caiga entre a y b cuando se sabe que la variable aleatoria discreta $Y = y$, se evalúa:

$$P(a < X < b | Y = y) = \sum_x f(x|y)$$

- Si se desea encontrar la probabilidad de que la variable aleatoria discreta Y caiga entre a y b cuando se sabe que la variable aleatoria discreta $X = x$, se evalúa:

$$P(a < Y < b | X = x) = \sum_y f(y|x)$$

Distribuciones de Probabilidad Conjunta (cont.)

- Si se desea encontrar la probabilidad de que la variable aleatoria continua X caiga entre a y b cuando se sabe que la variable aleatoria continua $Y = y$, se evalúa:

$$P(a < X < b \mid Y = y) = \int_a^b f(x \mid y)\,dx$$

- Si se desea encontrar la probabilidad de que la variable aleatoria continua Y caiga entre a y b cuando se sabe que la variable aleatoria continua $X = x$, se evalúa:

$$P(a < Y < b \mid X = x) = \int_a^b f(y \mid x)\,dy$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #1:** Obtener la distribución condicional de X, dado que $Y = 1$, y utilizarla para determinar $P(X = 0 \mid Y = 1)$, del ejemplo de la distribución conjunta de las variables discretas.
  - Se debe encontrar $f(x|y)$, donde $y = 1$, por lo tanto se debe encontrar primero $h(1)$:

$$h(1) = \sum_{x=0}^{2} f(x,1) = \frac{6}{28} + \frac{6}{28} + 0 = \frac{12}{28} = \frac{3}{7}$$

$$f(x|1) = \frac{f(x,1)}{h(1)} = \frac{f(x,1)}{3/7} = \frac{7}{3} f(x,1) \quad x = 0,1,2$$

Distribuciones de Probabilidad Conjunta (cont.)

$f(x,y)$		x = 0	x = 1	x = 2	**Totales por Fila**
y	**0**	3/28	9/28	3/28	**15/28**
	1	6/28	6/28	-	**12/28**
	2	1/28	-	-	**1/28**
Totales por Columna		**10/28**	**15/28**	**3/28**	**1**

Distribuciones de Probabilidad Conjunta (cont.)

x	**0**	**1**	**2**
$g(x)$	5/14	15/28	3/28

y	**0**	**1**	**2**
$h(y)$	15/28	3/7	1/28

Distribuciones de Probabilidad Conjunta (cont.)

$$f(x|1) = \frac{f(x,1)}{h(1)} = \frac{f(x,1)}{3/7} = \frac{7}{3} f(x,1) \quad x = 0,1,2$$

$$f(0|1) = \frac{7}{3} f(0,1) = \frac{7}{3} \times \frac{6}{28} = \frac{42}{84} = \frac{1}{2}$$

$$f(1|1) = \frac{7}{3} f(1,1) = \frac{7}{3} \times \frac{6}{28} = \frac{42}{84} = \frac{1}{2}$$

$$f(2|1) = \frac{7}{3} f(2,1) = \frac{7}{3} \times 0 = 0$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #1:** La distribución condicional de X, dado que $Y = 1$ es:

x	**0**	**1**	**2**
$f(x\|1)$	1/2	1/2	0

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #1:** La $P(X = 0 \mid Y = 1)$, utilizando la distribución condicional de X, dado que $Y = 1$, es:

$$P(X = 0 | Y = 1) = f(0|1) = \frac{1}{2}$$

  - Por lo tanto, si sabe que uno de los repuestos seleccionados es rojo, se tiene una probabilidad de ½ de que el otro repuesto no sea azul.

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** La densidad conjunta para las variables aleatorias (*X*,*Y*), donde *X* es el cambio de temperatura unitario y *Y* es la proporción de desplazamiento espectral que produce cierta partícula atómica es:

$$f(x,y)=\begin{cases}10xy^2 & 0<x<y<1\\ 0 & \text{otro caso}\end{cases}$$

  - Encontrar las densidades marginales $g(x)$ y $h(y)$, y la densidad condicional $f(y|x)$.
  - Encontrar la probabilidad de que el espectro se desplace más de la mitad de las observaciones totales, dado que la temperatura aumenta a 0.25 de unidad.

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** Las densidades marginales *g*(*x*) y *h*(*y*), y la densidad condicional *f*(*y*|*x*) son:

$$g(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_{x}^{1}10xy^2dy=\frac{10}{3}xy^3\Big|_{y=x}^{y=1}=\frac{10}{3}x(1-x^3),\quad 0<x<1$$

$$h(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_{0}^{y}10xy^2dx=5x^2y^2\Big|_{x=0}^{x=y}=5y^4,\quad 0<y<1$$

$$f(y|x)=\frac{f(x,y)}{g(x)}=\frac{10xy^2}{\frac{10}{3}x(1-x^3)}=\frac{3y^2}{(1-x^3)}\quad 0<x<y<1$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** La probabilidad de que el espectro se desplace más de la mitad de las observaciones totales, dado que la temperatura aumenta a 0.25 de unidad es:

$$P\left(Y > \frac{1}{2} \middle| X = 0.25\right) = \int_{1/2}^{1} f\left(y \middle| x = 0.25\right) dy = \int_{1/2}^{1} \frac{3y^2}{\left(1 - 0.25^3\right)} dy$$

$$= \left. \frac{y^3}{(63/64)} \right|_{1/2}^{1} = \frac{64}{63} - \frac{8}{63}$$

$$P\left(Y > \frac{1}{2} \middle| X = 0.25\right) = \frac{56}{63} = \frac{8}{9}$$

Distribuciones de Probabilidad Conjunta (cont.)

- Sean X y Y dos variables aleatorias, discretas o continuas, con distribución de probabilidad conjunta $f(x,y)$ y distribuciones marginales $g(x)$ y $h(y)$, respectivamente. Se dice que las variables aleatorias X y Y son **estadísticamente independientes** si y sólo si

$$f(x, y) = g(x)h(y)$$

para toda (x,y) dentro de sus rangos.

Distribuciones de Probabilidad Conjunta (cont.)

- **Prueba:** Si $f(x|y)$ no depende de y, entonces $f(x|y) = g(x)$ y $f(x,y) = g(x)h(y)$. Demostración:

$$f(x,y) = f(x|y)h(y)$$

$$g(x) = \sum_y f(x,y) = \sum_y f(x|y)h(y)$$

$$g(x) = f(x|y)\sum_y h(y)$$

$$\sum_y h(y) = 1$$

$$g(x) = f(x|y) \quad f(x,y) = g(x)h(y)$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Prueba:** Si $f(x|y)$ no depende de y, entonces $f(x|y) = g(x)$ y $f(x,y) = g(x)h(y)$. Demostración:

$$f(x,y) = f(x|y)h(y)$$

$$g(x) = \int_{-\infty}^{+\infty} f(x,y)dy = \int_{-\infty}^{+\infty} f(x|y)h(y)dy$$

$$g(x) = f(x|y)\int_{-\infty}^{+\infty} h(y)dy$$

$$\int_{-\infty}^{+\infty} h(y)dy = 1$$

$$g(x) = f(x|y) \quad f(x,y) = g(x)h(y)$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #1:** Mostrar que las variables aleatorias del ejemplo de la distribución conjunta de las variables discretas no son estadísticamente independientes.
  - Se considera el punto (0,1). De los cálculos y la tabla anteriores se encuentra que las tres probabilidades $f(0,1)$, $g(0)$ y $h(1)$ son

$$f(0,1) = \frac{6}{28} = \frac{3}{14}$$

$$g(0) = \sum_{y=0}^{2} f(0,y) = \frac{3}{28} + \frac{6}{28} + \frac{1}{28} = \frac{10}{28} = \frac{5}{14}$$

$$h(1) = \sum_{x=0}^{2} f(x,1) = \frac{6}{28} + \frac{6}{28} + 0 = \frac{12}{28} = \frac{3}{7}$$

$$\Rightarrow$$

$$f(0,1) \neq g(0)h(1)$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** La densidad conjunta para las variables aleatorias (X,Y) es:

$$f(x,y)=\begin{cases}\dfrac{x\left(1+3y^2\right)}{4} & 0<x<2,0<y<1\\[2ex] 0 & \text{otro caso}\end{cases}$$

  - Encontrar las densidades marginales $g(x)$ y $h(y)$, y la densidad condicional $f(x|y)$.
  - Demostrar que son estadísticamente independientes.

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** Las densidades marginales $g(x)$ y $h(y)$, y la densidad condicional $f(x|y)$ son:

$$g(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy = \int_0^1 \frac{x(1+3y^2)}{4}\,dy = \left.\frac{xy}{4} + \frac{xy^3}{4}\right|_{y=0}^{y=1} = \frac{x}{2}, \quad 0 < x < 2$$

$$h(y) = \int_{-\infty}^{+\infty} f(x,y)\,dx = \int_0^2 \frac{x(1+3y^2)}{4}\,dx = \left.\frac{x^2}{8} + \frac{3x^2y^2}{8}\right|_{x=0}^{x=2} = \frac{1+3y^2}{2}, \quad 0 < y < 1$$

$$f(x|y) = \frac{f(x,y)}{h(y)} = \frac{\dfrac{x(1+3y^2)}{4}}{\dfrac{(1+3y^2)}{2}} = \frac{x}{2} \quad 0 < x < 2$$

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo #2:** Las variables aleatorias *X* y *Y* son estadísticamente independientes:

$$f(x, y) = g(x)h(y)$$

$$\frac{x(1+3y^2)}{4} = \frac{x}{2} \times \frac{1+3y^2}{2}$$

$$\Rightarrow$$

$$\frac{x(1+3y^2)}{4} = \frac{x(1+3y^2)}{4}$$

Distribuciones de Probabilidad Conjunta (cont.)

- Todas las definiciones anteriores respecto a dos variables aleatorias se pueden generalizar al caso de *n* variables aleatorias.
- Sea $f(x_1,x_2,\ldots,x_n)$ la función de probabilidad conjunta de las variables aleatorias $x_1, x_2, \ldots, x_n$. La distribución marginal de X_1, por ejemplo, es

$$g(x_1) = \sum_{x_2} \ldots \sum_{x_n} f(x_1, x_2, \ldots, x_n), \quad \text{caso discreto}$$

$$g(x_1) = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f(x_1, x_2, \ldots, x_n) dx_2 dx_3 \ldots dx_n, \quad \text{caso continuo}$$

Distribuciones de Probabilidad Conjunta (cont.)

- Y se puede obtener **distribuciones marginales conjuntas** como $\phi(x_1,x_2)$, donde

$$\phi(x_1,x_2)=\begin{cases}\displaystyle\sum_{x_3}\cdots\sum_{x_n} f(x_1,x_2,\ldots,x_n) & \text{caso discreto}\\ \displaystyle\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} f(x_1,x_2,\ldots,x_n)\,dx_3\,dx_4\ldots dx_n & \text{caso continuo}\end{cases}$$

Distribuciones de Probabilidad Conjunta (cont.)

- Se pueden considerar numerosas distribuciones condicionales. Por ejemplo, la distribución condicional conjunta de X_1, X_2 y X_3, dado que $X_4 = x_4$, $X_5 = x_5$, …, $X_n = x_n$, se escribe como

$$f\left(x_1, x_2, x_3 \middle| x_4, x_5, \ldots, x_n\right) = \frac{f\left(x_1, x_2, \ldots, x_n\right)}{g\left(x_4, x_5, \ldots, x_n\right)}$$

- Donde $g(x_4,x_5,\ldots,x_n)$ es la distribución marginal conjunta de las variables aleatorias X_4, X_5,…, X_n.

Distribuciones de Probabilidad Conjunta (cont.)

- Sean $X_1, X_2, \ldots, X_n$ n variables aleatorias, discretas o continuas, con distribución de probabilidad conjunta $f(x_1, x_2, \ldots, x_n)$ y distribuciones marginales $f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$, respectivamente. Se dice que las variables aleatorias $X_1, X_2, \ldots, X_n$ son **estadísticamente independientes** mutuamente si y sólo si

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n)$$

para toda $(x_1, x_2, \ldots, x_n)$ dentro de sus rangos.

Distribuciones de Probabilidad Conjunta (cont.)

- **Ejemplo:** El tiempo de vida, en años, de cierto producto alimenticio perecedero empacado en cajas de cartón es una variable aleatoria cuya función de densidad de probabilidad está dada por

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{otro caso} \end{cases}$$

  Sean X_1, X_2 y X_3 los tiempos de vida para tres de estas cajas que se seleccionaron de forma independiente.
  - Encontrar $P(X_1 < 2,\ 1 < X_2 < 3,\ X_3 > 2)$.

Distribuciones de Probabilidad Conjunta (cont.)

- **Solución:** Como las cajas se seleccionan de forma independiente, se puede suponer que las variables aleatorias X_1, X_2 y X_3 son estadísticamente independientes y que tienen la densidad de probabilidad conjunta

$$f(x_1,x_2,x_3) = f(x_1)f(x_2)f(x_3) = e^{-x_1}e^{-x_2}e^{-x_3}$$

$$f(x_1,x_2,x_3) = \begin{cases} e^{-x_1-x_2-x_3} & x_1>0, x_2>0, x_3>0 \\ 0 & \text{otro caso} \end{cases}$$

$$P(X_1<2, 1<X_2<3, X_3>2) = \int_2^{+\infty}\int_1^3\int_0^2 e^{-x_1-x_2-x_3}\,dx_1dx_2dx_3$$

$$= \left(1-e^{-2}\right)\left(e^{-1}-e^{-3}\right)e^{-2}$$

$$P(X_1<2, 1<X_2<3, X_3>2) = 0.0376$$

Referencias Bibliográficas

- Walpole, R.E.; Myers, R.H.; Myers, S.L. & Ye, K. "Probabilidad y estadística para ingeniería y ciencias". Octava Edición. Pearson Prentice-Hall. México, 2007.