UNIVERSIDAD TECNOLÓGICA NACIONAL

Facultad Regional San Nicolás

PROBABILIDAD y ESTADÍSTICA II

UNIDAD Nº1

Licenciatura en Enseñanza de la Matemática

Año 2011

Mg. Lucía C. Sacco

UNIDAD Nº1

Variables aleatorias bidimensionales y n-dimensionales. Distribuciones.

Variables aleatorias bidimensionales discretas y continuas. Definiciones y ejemplos. Distribución de probabilidad puntual conjunta. Distribuciones marginales. Variables aleatorias independientes.

Propósitos:

Brindar oportunidades para la construcción de herramientas que permitan:

- Definir variables aleatorias n-dimensionales.
- Identificar situaciones reales en las que intervienen variables aleatorias bidimensionales.
- Analizar distribuciones de probabilidades de variables aleatorias discretas y continuas.
- Resolver situaciones problemáticas de diferentes distribuciones de probabilidad de variables aleatorias discretas y continuas.

Bibliografía sugerida:

- Canavos, George. *Probabilidad y Estadística. Aplicaciones y Métodos*. México. McGraw Hill. 1988.
- Meyer Paul L. *Probabilidad y Aplicaciones Estadísticas*. México. Addison Wesley Iberoamericana .1993.
- Walpole Ronald, Myers Raymond. *Probabilidad y Estadística*. México. Pearson Educación. 1999.
- Zylberberg, Alejandro. *Probabilidad y Estadística P(X).* Nueva Librería.

1. Introducción

A veces es necesario trabajar con probabilidades que involucran a más de una variable aleatoria al mismo tiempo.

Ejemplos

Podemos querer calcular cuál es la probabilidad de que:

- una persona elegida al azar mida entre 1.70 y 1.80 m y pese entre 80 y 90 kg.
- una persona que pesa entre 70 y 80 kg mida menos de 1.60 m.

En esos casos usamos variables aleatorias bidimensionales. En general, las variables aleatorias pueden ser de dimensión n; hablamos entonces de variables n-dimensionales.

Las variables aleatorias que estudiamos en **Probabilidad y Estadística I** son un caso particular, denominado variable aleatoria unidimensional. En general, podemos pensar a las variables aleatorias n-dimensionales como vectores, siendo cada una de las componentes del vector una variable aleatoria en sí.

2. Variables aleatorias bidimensionales

En el caso de las variables aleatorias bidimensionales, podemos pensarlas como un vector con dos componentes, cada una de las cuales es una variable aleatoria unidimensional.

Así como los valores posibles de una variable unidimensional están contenidos en una recta, siendo la recta misma, o parte de ella, los valores posibles de una variable bidimensional están contenidos en un plano, siendo todo el plano o parte de él.

Las variables aleatorias bidimensionales pueden ser discretas o continuas. Como cada componente de una variable aleatoria de dimensión mayor a 1 es una variable aleatoria unidimensional, una variable aleatoria bidimensional puede tener sus dos componentes discretas, sus dos componentes continuas, o una discreta y una continua.

El ejemplo que dimos al principio, del peso y la altura de una persona, tiene sus dos componentes continuas.

Un ejemplo de una variable aleatoria bidimensional con una componente discreta y una continua, puede ser considerar la longitud de las rutas y la cantidad de estaciones de servicio que hay en ellas.

3. Tipos de distribuciones

Así como en las variables aleatorias unidimensionales nos interesa estudiar cómo se distribuye la probabilidad de cada uno de los valores posibles, en las variables aleatorias bidimensionales nos interesa lo mismo, con la salvedad de que ahora los valores posibles son pares de valores, o bien vectores de dimensión 2. Notemos que:

1) la probabilidad de un determinado par de valores no puede ser menor que cero.
2) la suma de las probabilidades de todos los pares de valores da 1, porque al hacer el experimento siempre sale uno de los pares posibles.

Cuando se estudian conjuntamente dos variables, surgen tres tipos de distribuciones, conjuntas, marginales y condicionadas.

3.1 Función Distribución Conjunta

En las variables aleatorias unidimensionales, la función de densidad de probabilidad es una función que le asigna a cada valor posible de la variable aleatoria un número real que consiste en la probabilidad de que ocurra.

En las variables aleatorias bidimensionales, la imagen de la función sigue siendo de dimensión 1 (porque la probabilidad es un número) pero el dominio es de dimensión 2. Se define Función Distribución Conjunta de la variable bidimensional (X, Y) como:

$$\begin{array}{llll} F_{XY}: & \Re^2 & \to & \Re \\ & (x,y)\in\Re^2 & \to & F_{XY}(x,y)=P(X\le x, Y\le y) \end{array}$$

- F_{XY} contiene toda la información probabilística de la variable aleatoria (X, Y).
- $F_{XY}(x,y)$ representa la probabilidad de que la variable aleatoria bidimensional tome valores en el cuadrante (x,y)

3.1.1 Distribuciones conjuntas de variables aleatorias discretas

- $F_{XY}(x,y)$ es una función que a cada par de valores posibles le asigna su probabilidad.
- $F_{XY}(x,y)$ es una función de probabilidad discreta conjunta si y solo si cumple con:

  i) $F_{XY}(x,y)\ge 0 \quad \forall(x,y)$

  ii) $\sum_x \sum_y F_{XY}(x,y)=1$

  iii) $Si\ A\subset D_{XY} \Rightarrow P(A)=\sum\sum_{(x,y)\in A} P(x,y)$

  Siendo A un subconjunto de D_{XY} (dominio de la función F_{XY} incluido en R^2)

Ejemplo 1

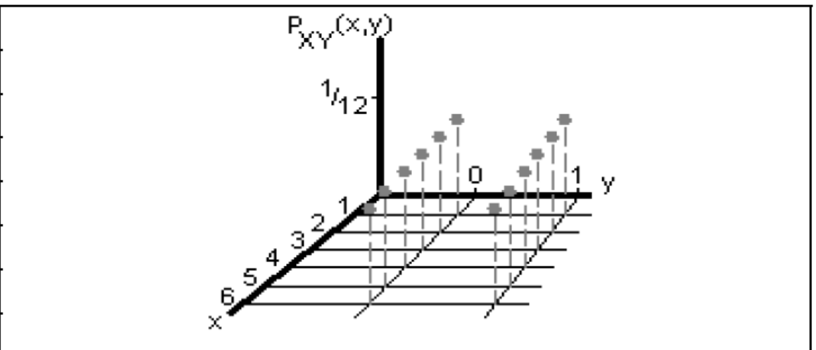
Siendo X: el número que sale al tirar un dado honesto e Y: la cantidad de caras que salen al tirar una moneda:

X: puede tomar del 1 a 6.

Y: cara (1) y ceca (0).

Hay doce eventos diferentes. Cada probabilidad de que salga un número y una de las caras de las monedas es 1/12.

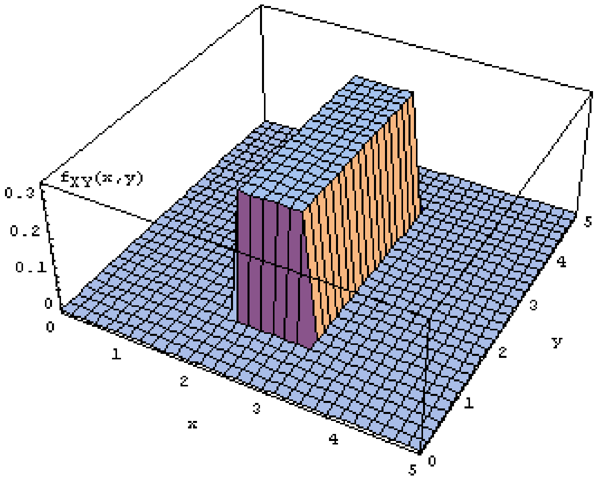
		Y	
	P_{XY}	**0**	**1**
X	**1**	$1/12$	$1/12$
	2	$1/12$	$1/12$
	3	$1/12$	$1/12$
	4	$1/12$	$1/12$
	5	$1/12$	$1/12$
	6	$1/12$	$1/12$



3.1.2 Distribuciones conjuntas de variables aleatorias continuas

Análogamente a la función de densidad de una variable aleatoria unidimensional, para obtener probabilidades a partir de la **función de densidad** de una variable aleatoria bidimensional debemos integrarla.

En vez de una integral simple, es una integral doble. Es decir, la integral de la función de densidad $f_{XY}(x,y)$ en un dominio D del plano xy, da la probabilidad de que la variable aleatoria XY asuma un valor comprendido en ese dominio.



$f_{XY}(x,y)$ es una función de densidad de probabilidad continua conjunta si y solo si cumple con:

i) $f_{XY}(x,y) \geq 0 \quad \forall (x,y)$

ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1$

iii) $P(A) = \iint_A f_{XY}(x,y)dxdy$

Siendo A un subconjunto de R^2 incluido en el dominio de $f_{XY}(x,y)$

Ejemplo 2

Se toma un punto al azar del plano XY, con la primera componente entre 2 y 3, y la segunda entre 1 y 4, y se toma la variable aleatoria X como la componente x del punto, y la variable aleatoria Y como la componente y del punto.

$$f_{XY}(x,y) = \begin{Bmatrix} \frac{1}{3} \; si \;\; 2 < x < 3; 1 < y < 4 \\ 0 \qquad \forall otro \; x,y \end{Bmatrix}$$

Luego la probabilidad de que el par *(x, y)* caiga en un determinado intervalo es la integral de la altura $f_{XY}(x,y)$ en dicho intervalo.

3.2 Distribuciones marginales

Cada componente de una variable aleatoria bidimensional es una variable aleatoria unidimensional en sí misma.

Es decir, cada una de las dos variables aleatorias que forman la variable aleatoria bidimensional es una variable aleatoria unidimensional común y corriente. Entonces nos puede interesar conocer la distribución de una componente por separado, sin tener en cuenta a la otra componente.

Eso se denomina "marginar", y la distribución de la variable unidimensional por separado se llama "distribución marginal".

3.2.1 Distribuciones marginales de variables aleatorias discretas

Sea la variable aleatoria bidimensional (X, Y) distribuida según $P_{XY}(x,y)$, la distribución de X, también llamada **distribución marginal de X,** es $P_X(x)=\sum_{y=-\infty}^{+\infty}P_{XY}(x,y)$ para cada valor x de la variable aleatoria X. Análogamente, la **distribución marginal de Y** es $P_Y(y)=\sum_{x=-\infty}^{+\infty}P_{XY}(x,y)$ para cada valor y de la variable aleatoria Y. Es decir, para cada valor posible de la variable aleatoria cuya distribución se desea hallar, se suman las probabilidades conjuntas de ese valor con cada uno de los valores posibles de la otra variable.

Ejemplo 3

Si la distribución conjunta de X e Y es la siguiente, hallar la distribución de X y la distribución de Y por separado.

P_{XY}		Y	
		20	**30**
X	**1**	0,1	0,3
	2	0,4	0,2

Primero enumeramos los valores posibles de X: 1 y 2. Para cada valor posible de X, aplicamos la fórmula:

$$P_X(1)=\sum_{y=-\infty}^{+\infty}P_{XY}(1,y)=P_{XY}(1,20)+P_{XY}(1,30)=0{,}1+0{,}3=0{,}4$$

$$P_X(2)=\sum_{y=-\infty}^{+\infty}P_{XY}(2,y)=P_{XY}(2,20)+P_{XY}(2,30)=0{,}4+0{,}2=0{,}6$$

Entonces se obtiene $P_X(x)=\begin{Bmatrix}0{,}4 & x=1\\ 0{,}6 & x=2\\ 0 & \forall\ otro\ x\end{Bmatrix}$

Luego, enumeramos los valores posibles de Y: 20 y 30. Para cada valor posible de Y, aplicamos la fórmula:

$$P_Y(20)=\sum_{x=-\infty}^{+\infty}P_{XY}(x,20)==P_{XY}(1,20)+P_{XY}(2,20)=0{,}1+0{,}4=0{,}5$$

$$P_Y(30)=\sum_{x=-\infty}^{+\infty}P_{XY}(x,30)==P_{XY}(1,30)+P_{XY}(2,30)=0{,}3+0{,}2=0{,}5$$

Entonces se obtiene

$$P_Y(y)=\begin{Bmatrix}0{,}5 & y=20\\ 0{,}5 & y=30\\ 0 & \forall\ otro\ y\end{Bmatrix}$$

Veamos lo que ocurre si en la tabla que usamos para escribir la distribución conjunta, agregamos los totales por fila y por columna:

P_{XY}		Y		
		20	**30**	
X	**1**	0,1	0,3	0,4
	2	0,4	0,2	0,6
		0,5	0,5	

Observamos que en los márgenes de la tabla no obtuvimos otra cosa que las distribuciones marginales de X y de Y. Esa es la razón por la cual las distribuciones de X e Y por separado se denominan "marginales".

3.2.2 Distribuciones marginales de variables aleatorias continuas

La marginación de variables continuas es análoga a la de las variables discretas, pero puede acarrear algunas dificultades adicionales.

Sea la variable aleatoria bidimensional (X, Y) distribuida según $f_{XY}(x,y)$:

- la distribución de X (también llamada distribución marginal de X) es $f_X = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$ para cada región del dominio de X donde no cambien los límites de integración de $f_{XY}(x,y)$ con respecto a Y.
- Análogamente, la distribución de Y es $f_Y = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$ para cada región del dominio de Y donde no cambien los límites de integración de $f_{XY}(x,y)$ con respecto a X.

Para calcular $f_X(x)$ es posible aplicar el siguiente método:

1) Subdividir el dominio de X de forma tal que en cada intervalo no cambien:

- Las ecuaciones que determinan los límites de integración de $f_{XY}(x,y)$ respecto de Y
- Las ecuaciones que determinan la separación de las ramas de $f_{XY}(x,y)$ (si las hay)

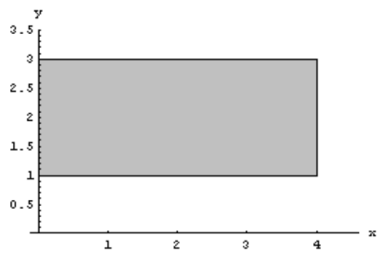
2) Para cada intervalo, calcular $f_X = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$, teniendo en cuenta que si en ese intervalo de X hay distintas ramas de $f_{XY}(x,y)$, la integral será la suma de distintas integrales.

3) Armar la $f_X(x)$ poniendo en cada intervalo lo calculado en el punto 2.

Ejemplo 4

Sea la siguiente distribución:

$$f_{XY}(x,y) = \begin{cases} \frac{1}{8} & si \quad 0 < x < 4,\ 1 < y < 3 \\ 0 & \forall \ otro \ \ x, y \end{cases}$$



Calcular $f_X = \int_{-\infty}^{\infty} f_{XY}(x,y)dy$ y $f_Y = \int_{-\infty}^{\infty} f_{XY}(x,y)dx$

Marginación de x:

Tenemos que subdividir el dominio de la X de forma tal que dentro de cada intervalo no cambien las ecuaciones que determinan los límites de integración respecto de Y, ni las que separan ramas de $f_{XY}(x,y)$. En esta $f_{XY}(x,y)$ no hay múltiples ramas, así que para dividir en intervalos el dominio de X, solamente tendremos en cuenta el comportamiento de la Y en cada intervalo:

Para -∞ < x < 0, la Y no aparece, para 0 < x < 4, la Y varía entre 1 y 3 y para 4 < x < +∞, la Y no aparece.

Entonces nos quedan 3 intervalos, de los cuales 2 son triviales, sabemos que la densidad marginal de la X será cero en ellos.

Entonces aplicamos la fórmula al único intervalo relevante (0 < x < 4) $f_X = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy = \int_1^3 \frac{1}{8} dy = \frac{1}{4}$

Luego construimos la función de densidad de la X, que tendrá solamente una rama porque hubo un solo intervalo relevante $f_X(x) = \begin{cases} \frac{1}{4} & 0 < x < 4 \\ 0 & \forall \ otro \ x \end{cases}$

Marginación de y:

En este caso la marginación de Y es muy similar a la de X. Como no hay múltiples ramas, solo vamos a observar el comportamiento de la X a la hora de tomar intervalos para la Y. Procedemos para -∞ < y < 1, la X no aparece, para 1 < y < 3, la X varía entre 0 y 4 y para 3 < y < +∞, la X no aparece.

Entonces nos quedan 3 intervalos, de los cuales 2 son triviales. Entonces aplicamos la fórmula al único intervalo relevante (1 < y < 3) $f_y = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dx = \int_0^4 \frac{1}{8} dx = \frac{1}{2}$

Luego construimos la función de densidad de la Y, que tendrá solamente una rama porque hubo un solo intervalo relevante $f_Y(y) = \begin{cases} \frac{1}{2} & 1 < y < 3 \\ 0 & \forall \ otro \ y \end{cases}$

3.3 Distribuciones condicionales

Ahora vamos a tomar lo estudiado en la **Unidad Nº2 de Probabilidad y Estadística I** sobre probabilidad condicional de sucesos, y lo vamos a extender a las variables aleatorias. En particular vamos a analizar cómo el hecho de que conozcamos el valor que asumió una variable aleatoria al hacer el experimento modifica la distribución de probabilidad de otra variable cuyo resultado aún no conocemos.

En dicha unidad vimos que en general, si sabemos que un **suceso** ocurre, eso modifica las probabilidades de los demás **sucesos**.

3.3.1 Distribución condicional para variables aleatorias continuas

Sean X, Y variables aleatorias continuas, se define:

$$f_{X/Y}(x,y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Donde $f_{X/Y}(x,y)$ se lee **función de densidad condicional de X dado Y**, y es una función de densidad de X, pero que es "genérica" porque además depende de Y, y para cada valor de Y, será una distribución en concreto para X.

Es decir, dado Y, tenemos una distribución para X. Dicho de otro modo, conociendo Y, tenemos una distribución para X.

La función de densidad condicional de X dado Y determina la correspondiente distribución condicional de probabilidades, es decir, nos dice cómo se distribuyen las probabilidades de los valores de X, una vez que se conoce el valor que ha tomado Y.

Según vemos en la fórmula, la función de densidad condicional de X se obtiene dividiendo la función de densidad conjunta por la función de densidad marginal de Y.

Notemos que esta fórmula es análoga a la fórmula que se dio en la **Unidad Nº2** para la probabilidad condicional: P (A/B) = P (A $\cap$ B) / P (B).

Por simplicidad, se trabaja con distribuciones muy sencillas, que seguramente no se ajustan mucho a la realidad, pero nos permitirán visualizar los conceptos.

Ejemplo 5

Se realiza el experimento de tomar una persona al azar y medir su peso y su altura. Se definen los sucesos A y B. Suceso A: La persona pesa más de 60kg. Suceso B: La persona mide 1.90 m. Llamemos X a la variable aleatoria peso e Y a la variable aleatoria altura.

Donde Y está expresada en metros y X está expresada en decenas de kg.

La distribución conjunta es $f_{XY}(x,y) = \begin{cases} \frac{1}{8}y & 0 < y < 2, \quad 3y < x < 3y+4 \\ 0 & \forall \ otro \ \ x,y \end{cases}$

Sabiendo que ocurre el suceso B, ¿cuál es la distribución condicional del peso?

Estudio analítico de la situación

En principio el suceso A puede ocurrir con probabilidad P(A). Pero si sabemos que el suceso B ocurrió, entonces la probabilidad de que ocurra A será seguramente mayor, porque si se sabe que la persona mide 1.90 m, que pese más de 60kg es más probable que si no conocemos la altura.

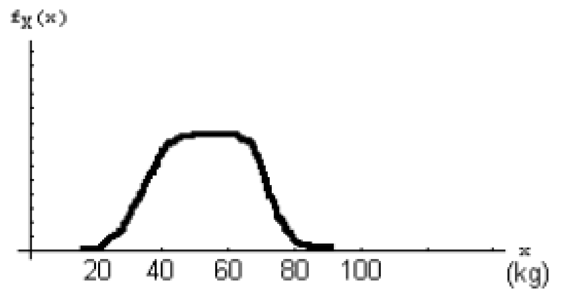
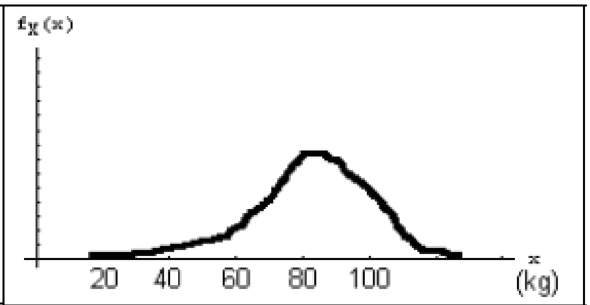
De hecho P(A/B) será un valor muy cercano a 1, porque es muy probable que una persona que *sabemos* que mide 1.90 m pese más de 60kg. Hasta aquí nada nuevo.

Ahora supongamos que el peso y la altura de la persona en realidad son **variables aleatorias**.

La conclusión inmediata es que si conocemos el valor que tomó una de las variables aleatorias al hacer el experimento, eso nos modificará la distribución de probabilidad de la otra variable aleatoria.

Tenemos la función de densidad conjunta de las dos variables aleatorias. Podemos, si queremos, obtener la distribución marginal del peso, es decir, la distribución de la variable peso, que no tiene en cuenta la altura.

Pero si conociéramos que la variable altura tomó el valor 1.90m, ¿la distribución marginal del peso que teníamos sigue siendo válida? No. Seguramente, la masa de probabilidad del peso tenderá a distribuirse más hacia los valores más altos. Un gráfico nos permitirá visualizar lo planteado y entenderlo más intuitivamente:

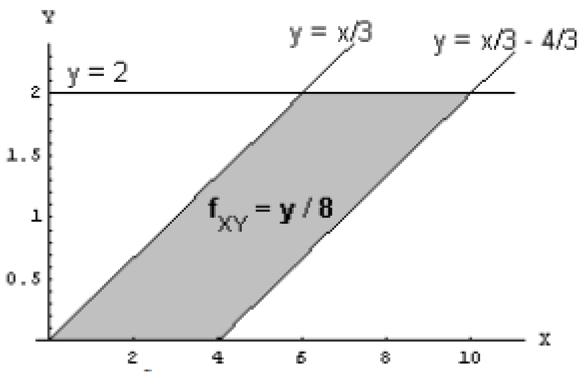
	
Distribución marginal de la variable peso, es decir, sin saber nada de la altura.	Distribución de la variable peso, sabiendo que la variable altura tomó el valor 1.90

Podemos repetir esto muchas veces para distintos valores de la altura, y obtendríamos distintas distribuciones para el peso. Esto nos lleva a pensar que podemos encontrar una distribución "genérica" del peso en función de la altura, es decir, una función de densidad para el peso en la cual también aparezca la variable altura, y entonces para cada valor que tome la variable altura, tendremos una función de densidad distinta para el peso.

Esa distribución del peso que es genérica porque además aparece la altura, y que se transforma en una distribución en particular al darle un valor a la altura, se denomina **distribución condicional del peso dada la altura**.

Cálculo de la distribución marginal del peso:

Realizamos en primer lugar el gráfico del dominio de la función densidad.



Del mismo, es posible observar que las variables se condicionan mutuamente.

Por ejemplo, si la altura es 2 m, el peso necesariamente está entre 60 y 100 kg (la probabilidad es nula fuera de ese intervalo). Si el peso es de 80 kg, la altura necesariamente debe estar entre 1,33 y 2m.

Vemos que en principio, saber qué resultado arrojó una variable nos condiciona acerca de cuáles son los valores posibles de la otra variable.

Por ejemplo en el gráfico vemos que la variable peso puede ir entre 0 y 100 kg (esto es, cuando no conocemos la altura). Pero si **conocemos** que por ejemplo la altura es 2 m, el peso ya no puede variar entre 0 y 100 kg, sino entre 60 y 100 kg.

Como en este ejemplo conocemos la distribución conjunta, podríamos, por ejemplo, proceder como estudiamos en la sección anterior, para encontrar las distribuciones marginales del peso y la altura.

De esa forma tendríamos f_X y f_Y, las distribuciones marginales de X e Y, es decir, las distribuciones de X y de Y que no tienen en cuenta a la otra variable. O sea, las distribuciones que tenemos para X e Y cuando no sabemos qué valor tomó la otra variable.

Pero en este caso nos interesa estudiar cómo se distribuye X (el peso) si conocemos, es decir, si es dato, el valor de Y (altura).

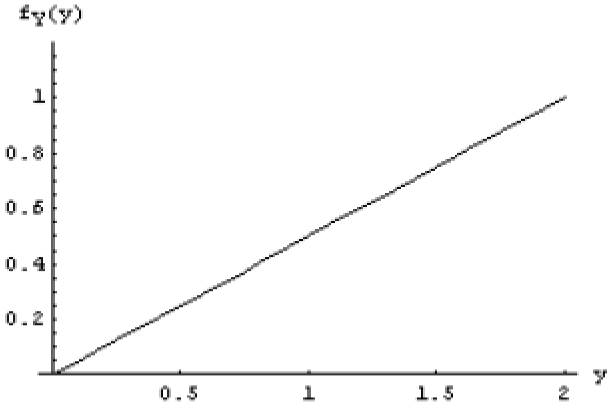
Vamos a usar la fórmula que vimos antes $f_{X/Y}(x,y)=\dfrac{f_{XY}(x,y)}{f_Y(y)}$

Para encontrar la distribución condicional de X dado Y, vamos a necesitar la conjunta de X e Y, y la marginal de Y. La conjunta de X e Y es dato; la marginal de Y la encontramos a partir de la conjunta según se estudió en la sección anterior:

$$f_Y(y)=\int_{-\infty}^{\infty} f_{XY}(x,y)dx=\int_{3y}^{3y+4}\frac{1}{8}dx=\frac{1}{2}$$

Lo cual vale para el intervalo $0<y<2$. Luego:

$$f_Y(y)=\begin{cases}\dfrac{y}{2} & 0<y<2\\ 0 & \forall\ otro\ \ y\end{cases}$$



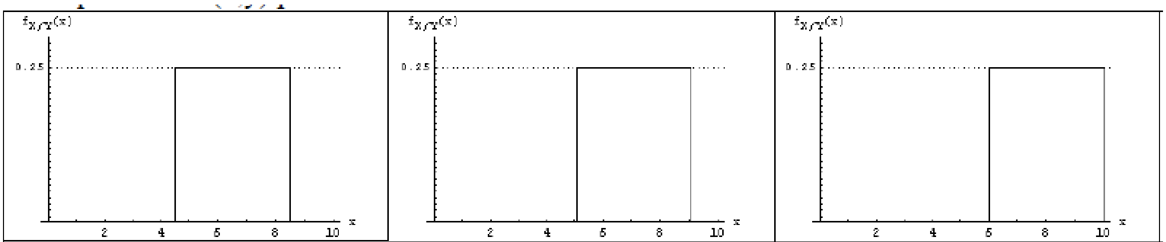
Ahora vamos a dividir la conjunta por la marginal de Y para encontrar la condicional de X dado Y.

Recordemos que para dividir dos funciones partidas lo que se hace es, rama a rama, dividir los valores, e intersecar los dominios. En este caso, (y/8) / (y/2) = 1/4, y el dominio donde esto es válido es la intersección de los dominios.

Dicha intersección coincide con el dominio de la conjunta, y en dicho dominio x varía entre *3y* y *3y+4*. En resumen queda $f_{X/Y}(x,y)=\begin{cases}\dfrac{1}{4} & 0<y<2,\ 3y<x<3y+4\\ 0 & \forall\ otro\ \ x,y\end{cases}$

Esa es la función de densidad condicional de X dado Y. En ella podemos poner cualquier valor permitido de Y, y obtendremos la distribución de probabilidades para X dado que conocemos el valor de Y. Por ejemplo, si en esa función ponemos y = 1.8, obtendremos la distribución del peso X de las personas que miden 1.80m.

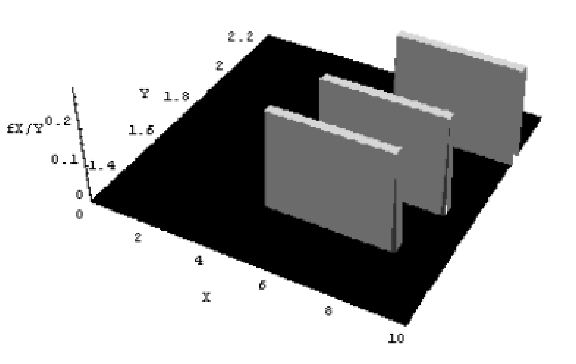
Grafiquemos $f_{X/Y}(x,y)=\begin{cases}\dfrac{1}{4} & 0<y<2,\ 3y<x<3y+4\\ 0 & \forall\ otro\ \ x,y\end{cases}$ para distintos valores de Y:



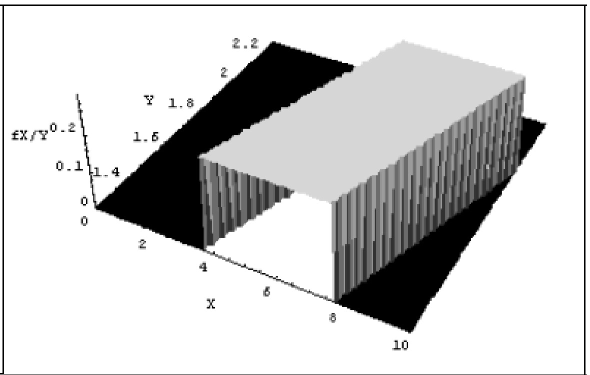
y = 1.50 m	y = 1.70 m	y = 2 m

Observamos que para distintos valores de la altura, las probabilidades de los valores posibles del peso son distintas. En este caso vemos que a medida que la altura aumenta, la masa de probabilidades de los pesos se va corriendo hacia los valores grandes.

Veamos ahora gráficos de $f_{X/Y}(x,y)=\begin{cases}\dfrac{1}{4} & 0<y<2,\ 3y<x<3y+4\\ 0 & \forall\ otro\ x,y\end{cases}$ en 3 dimensiones:



Estos son los mismos cortes de antes (y = 1.50, y = 1.70, y = 20) pero vistos en 3 dimensiones



Esta es la gráfica completa, sin hacer cortes. En ella se aprecia plenamente lo que el conocimiento de la Y le hace a la X.

3.3.2 Distribución condicional para variables aleatorias discretas

Sean X, Y variables aleatorias discretas, se define $P_{X/Y}(x,y)=\dfrac{P_{XY}(x,y)}{P_Y(y)}$

Donde $P_{X/Y}$ se lee **función de probabilidad condicional de X dado Y**, y es una función de probabilidad de X, pero que es "genérica" porque además depende de Y, y para cada valor de Y, será una distribución en concreto para X.

La función de probabilidad condicional de X dado Y determina la correspondiente distribución condicional de probabilidades, es decir, nos dice cómo se distribuyen las probabilidades de los valores de X, una vez que se conoce el valor que ha tomado Y.

Según vemos en la fórmula, la función de probabilidad condicional de X se obtiene dividiendo la función de probabilidad conjunta por la función de probabilidad marginal de Y.

Notemos nuevamente que esta fórmula es análoga a la fórmula que se dio en la **Unidad Nº2** para la probabilidad condicional: P (A/B) = P (A $\cap$ B) / P (B).

Ejemplo 6

Se tienen las variables aleatorias discretas X e Y, cuya distribución conjunta es la que se presenta en la siguiente tabla. Calcular $P_{X/Y}(x,y)$.

P_{XY}		Y		
		0	2	4
X	1	0,25	0,05	0,3
	2	0,15	0,1	0,15

Como podemos apreciar en la fórmula, vamos a necesitar la distribución marginal de Y. Le agregamos las distribuciones marginales a la tabla:

P_{XY}		**Y**			**P_X**
		0	**2**	**4**	
X	**1**	0,25	0,05	0,3	**0,6**
	2	0,15	0,1	0,15	**0,4**
P_Y		**0,4**	**0,15**	**0,45**	

Es decir, obtuvimos que las distribuciones marginales son:

$$P_X(x) = \begin{cases} 0{,}6 & x = 1 \\ 0{,}4 & x = 2 \\ 0 & \forall \ otro \ x \end{cases} \qquad P_Y(y) = \begin{cases} 0{,}4 & y = 0 \\ 0{,}15 & y = 2 \\ 0{,}45 & y = 4 \\ 0 & \forall \ otro \ y \end{cases}$$

Para encontrar $P_{X/Y}(x,y)$ utilizamos la fórmula, la cual nos dice que para cada x y cada y, la probabilidad de que X = x dado que Y = y se obtiene como el cociente de la conjunta evaluada en (x,y) y la marginal de Y evaluada en y. Es decir, cada probabilidad de la posición (x,y) de la tabla vale $P_{XY}(x,y)/P_Y(y)$:

$P_{X/Y}$		**Y**		
		0	**2**	**4**
X	**1**	0,25/0,4	0,05/0,15	0,3/0,45
	2	0,15/0,4	0,1/0,15	0,15/0,45

—> haciendo las cuentas

$P_{X/Y}$		**Y**		
		0	**2**	**4**
X	**1**	5/8	1/3	2/3
	2	3/8	2/3	1/3

Llegamos al resultado fácil y rápidamente. También podemos expresar el resultado mediante alguna notación con llaves, como por ejemplo:

$$P_{X/Y}(x,y) = \begin{cases} 5/8 & x = 1 \ \wedge \ y = 0 \\ 3/8 & x = 2 \ \wedge \ y = 0 \\ 1/3 & x = 1 \ \wedge \ y = 2 \\ 2/3 & x = 2 \ \wedge \ y = 2 \\ 2/3 & x = 1 \ \wedge \ y = 4 \\ 1/3 & x = 2 \ \wedge \ y = 4 \\ 0 & \forall \ otro \ x, y \end{cases}$$

4. Independencia de variables aleatorias

En **Probabilidad y Estadística I** se estudió el concepto de independencia de sucesos. Se estableció que dos sucesos son estadísticamente independientes si el conocimiento de que ocurrió uno de ellos no afecta la probabilidad de que el otro ocurra.

Si quisiéramos generalizar ese concepto a las variables aleatorias, tendríamos que decir que dos variables aleatorias son estadísticamente independientes si el conocimiento del valor que arrojó una de ellas no afecta la distribución de probabilidades de los valores que puede arrojar la otra.

Pensándolo, eso es lo mismo que decir que X e Y son independientes si $f_{X/Y}(x,y)$ es idéntica para todos los posibles valores de Y. Lo que es lo mismo que decir que $f_{X/Y}(x,y)$ no depende de Y.

El siguiente paso es darse cuenta de que si $f_{X/Y}(x,y)$ no depende de Y, entonces es en realidad $f_X(x)$, es decir, la **distribución marginal** de X, porque recordemos que hablar de distribución condicional de X tiene sentido solamente cuando en una función de densidad de X aparece alguna otra variable aleatoria.

Otra forma de pensarlo es que si $f_{X/Y}(x,y)$ es la distribución de X sabiendo el valor que arrojó Y, y $f_X(x)$ es la distribución de X cuando no se sabe qué valor arrojó Y, y el conocimiento de los valores de Y no afecta la distribución de X, entonces necesariamente $f_{X/Y}(x,y)$ debe ser igual a $f_X(x)$, porque si Y no afecta a X, entonces a la hora de distribuir X da lo mismo si conocemos el valor de Y que si no lo conocemos.

Llegamos entonces a la conclusión de que X e Y son independientes si $f_{X/Y}(x,y) = f_X(x)$. Si reemplazamos en esa ecuación $f_{X/Y}(x,y)$ por $f_{XY}(x,y)/f_Y(y)$, llegamos a la expresión equivalente $f_{XY}(x,y) = f_X(x).f_Y(y)$

Demos entonces la definición de independencia estadística de variables aleatorias:

Para X, Y variables aleatorias continuas	**Para X, Y variables aleatorias discretas**
X e Y son estadísticamente independientes	X e Y son estadísticamente independientes
<=>	<=>
$f_{X/Y}(x,y) = f_X(x)$	$P_{X/Y}(x,y) = P_X(x)$
<=>	<=>
$f_{X/Y}(x,y) = f_Y(y)$	$P_{X/Y}(x,y) = P_Y(y)$
<=>	<=>
$f_{XY}(x,y) = f_X(x).f_Y(y)$	$P_{XY}(x,y) = P_X(x).P_Y(y)$

4. 1 Variables continuas:

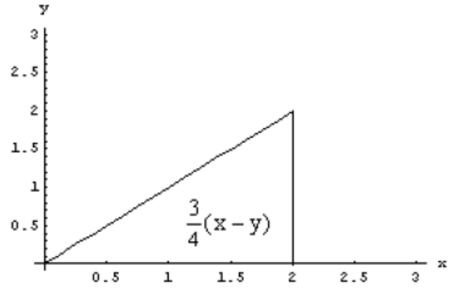
Ejemplo 7

Analizar para cada una de las siguientes distribuciones si las variables X e Y son independientes o no:

a) $f_{XY}(x,y) = \begin{cases} \frac{3}{4}(x-y) \;\; si \;\; 0<x<2; 0<y<x \\ 0 \qquad \forall otro \; x,y \end{cases}$

b) $f_{XY}(x,y)=\begin{cases}\frac{1}{72}xy^2 & si\ \ 0<x<4;0<y<3\\ 0 & \forall otro\ x,y\end{cases}$

a) La representación del dominio es:



Marginamos calculando:

$$f_X(x)=\int_{-\infty}^{\infty} f_{XY}(x,y)dy=\int_0^2 \frac{3}{4}(x-y)dy=\frac{3}{8}x^2$$

$$f_Y(y)=\int_{-\infty}^{\infty} f_{XY}(x,y)dx=\int_y^2 \frac{3}{4}(x-y)dx=\frac{3}{2}(\frac{1}{4}y^2-y+1)$$

Tenemos entonces:

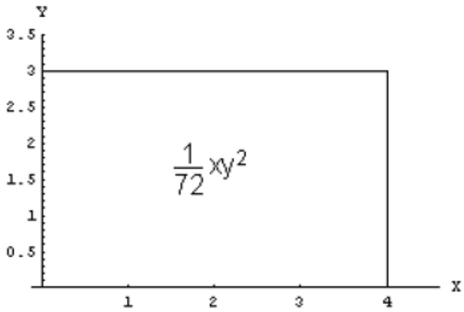
$$f_X(x)=\begin{cases}\frac{3}{8}x^2 & 0<x<2\\ 0 & \forall\ \ otro\ \ x\end{cases}\qquad f_Y(y)=\begin{cases}\frac{3}{2}(\frac{1}{4}y^2-y+1) & 0<y<2\\ 0 & \forall\ \ otro\ \ x\end{cases}$$

Multiplicándolas se obtiene que el valor es:

$$\frac{3}{8}x^2\ .\ \frac{3}{2}(\frac{1}{4}y^2-y+1)=\frac{9}{16}x^2(\frac{1}{4}y^2-y+1)$$

Y el dominio es $0<x<2\cap 0<y<2$. Se ve claramente que ni los valores ni el dominio coinciden con los de la función original. Luego, **X e Y no son independientes.**

b) La representación del dominio es:



Marginamos calculando:

$$f_X(x)=\int_{-\infty}^{\infty} f_{XY}(x,y)dy=\int_0^3 \frac{1}{72}xy^2dy=\frac{1}{8}x$$

$$f_Y(y)=\int_{-\infty}^{\infty} f_{XY}(x,y)dx=\int_0^4 \frac{1}{72}xy^2dx=\frac{1}{9}y^2$$

Tenemos entonces:

$$f_X(x)=\begin{cases}\frac{1}{8}x & 0<x<4\\ 0 & \forall\ \ otro\ \ x\end{cases}\qquad f_Y(y)=\begin{cases}\frac{1}{9}y^2 & 0<y<3\\ 0 & \forall\ \ otro\ \ x\end{cases}$$

Multiplicándolas se obtiene que el valor es:

$$\frac{1}{8}x \cdot \frac{1}{9}y^2 = \frac{1}{72}xy^2$$

Y el dominio es $0 < x < 4 \cap 0 < y < 3$. Se ve claramente que los valores y el dominio coinciden con los de la función original. Luego, **X e Y son independientes.**

4. 2 Variables discretas:

Ejemplo 8

Analizar para cada una de las siguientes distribuciones si las variables X e Y son independientes o no:

a)

X		Y		
	P_{XY}	**1**	**2**	**3**
	1	0.12	0.1	0.08
	2	0.28	0.2	0.22

b)

X		Y		
	P_{XY}	**1**	**2**	**3**
	1	0.08	0.12	0.2
	2	0.12	0.18	0.3

4.3 Consideraciones acerca del dominio

Hay una manera que en algunos casos permite determinar en forma inmediata y sin hacer cuentas que dos variables no son independientes.

Observemos el dominio de la función del **ejemplo 7a)**. Si supiéramos que X vale 1, entonces Y puede asumir cualquier valor entre 0 y 1. Si supiéramos que X vale 2, entonces Y puede asumir cualquier valor entre 0 y 2.

Vemos entonces que el hecho de conocer el valor que arrojó X nos afecta cuáles son los valores posibles de Y. Entonces es evidente que X e Y **no son independientes**.

Miremos en cambio el dominio de la función del **ejemplo 7b)**. La variable Y puede asumir cualquier valor entre 0 y 3, sin importar el valor que haya arrojado X. Análogamente, saber cuánto vale Y tampoco condiciona los valores posibles de X. ¿Esto significa que son independientes? **NO**.

Solamente significa que el conjunto de valores posibles de cada variable no es afectado por el conocimiento del valor que arrojó la otra. Pero lo que sí puede cambiar es cómo se distribuye la probabilidad entre los valores posibles. Entonces puede que no sean independientes.

¿Qué característica del dominio del **ejemplo 7b)** es la que hace que el conjunto de valores posibles de cada variable no sea afectado por el conocimiento del valor que arrojó la otra variable?

Que tiene forma **rectangular**. Entonces por lo que dijimos antes, que el dominio sea rectangular es **condición necesaria** para que las variables sean independientes. Pero no suficiente.

Entonces, en el **ejemplo 7a)**, con solo mirar el dominio podríamos haber contestado que las variables no son independientes, sin hacer ninguna cuenta. En el **ejemplo 7b)**, vemos que pueden ser independientes porque el dominio lo permite, pero también podrían no serlo, por lo cual hay que hacer la cuenta para determinarlo.

Dijimos que el hecho de que el dominio tenga forma rectangular es condición necesaria (pero no suficiente) para que las variables sean independientes.

4.4 Esperanza condicional

Para definir la esperanza condicional, vamos a combinar dos conceptos que ya hemos estudiado: la esperanza de una distribución, y la distribución condicional.

Dada una distribución, su media o esperanza nos da una idea de cuál es el valor que podemos esperar obtener al hacer el experimento. A su vez, la distribución condicional es un modelo que, dado el valor arrojado por una variable, nos permite tener una distribución de probabilidades para la otra variable.

La función de densidad condicional, por ejemplo de X dado Y, depende de "x" y de "y", y nos permite obtener una **distribución para X**, al conocer el valor de Y. Podemos pensar que el "y" que aparece en $f_{X/Y}(x,y)$ es simplemente un número, un parámetro, ya que para cualquier valor válido de "y", $f_{X/Y}(x,y)$ es una distribución perfectamente válida para x. Recordemos que la diferencia entre $f_{X/A}(x,a)$ y una $f_X(x,a)$ es si consideramos a A una variable aleatoria o simplemente un parámetro.

Por lo tanto, podemos calcularle la esperanza a $f_{X/Y}(x,y)$ asumiendo que "y" es simplemente un número. Luego, obtendremos una esperanza para X que dependerá de "y".

Esta herramienta sirve ver cómo los valores de Y afectan al valor esperado de X.

Así como la esperanza de la distribución $f_X(x)$ es $E(X)=\mu_X = \int_{-\infty}^{\infty} x \, . \; f_X(x)\,dx$

Análogamente la esperanza de la distribución $f_{X/Y}(x)$ es $E(X/Y)=\mu_{X/Y} = \int_{-\infty}^{\infty} x \, . \; f_{X/Y}(x,y)\,dx$

Ejemplo 9

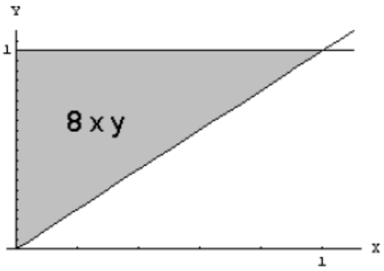
Obtener la esperanza de ambas distribuciones condicionadas.

Se tienen las variables aleatorias X e Y, cuya distribución conjunta es:

$$f_{XY}(x,y)=\begin{cases} 8xy & 0<x<1, \quad x<y<1 \\ 0 & \forall \; otro \; x,y \end{cases}$$

Podemos obtener las distribuciones marginales:

$$f_X(x)=\begin{cases} 4x(1-x^2) & 0<x<1 \\ 0 & \forall \; otro \; x \end{cases} \qquad f_Y(y)=\begin{cases} 4y^3 & 0<y<1 \\ 0 & \forall \; otro \; y \end{cases}$$



Y las condicionales:

$$f_{X/Y}(x,y)=\begin{cases}\dfrac{2x}{y^2} & 0<x<y,\quad 0<y<1\\ 0 & \forall \; otro \; x,y\end{cases};$$

$$f_{Y/X}(x,y)=\begin{cases}\dfrac{2y}{1-x^2} & x<y<1,\quad 0<x<1\\ 0 & \forall \; otro \; x,y\end{cases}$$

Calculemos la esperanza condicional de X dado Y:

$$\mu_{X/Y}=\int_{-\infty}^{\infty} x\, f_{X/Y}(x,y)\, dx=\int_0^y \frac{2x^2}{y^2}\, dx=\frac{2}{3}y$$

Eso quiere decir, que si por ejemplo la variable Y arroja el valor 1/2, el valor esperado para X será 1/3. Es decir, si conocemos el valor que arrojó Y, el valor esperado de X es 2/3 de ese valor.

También podemos calcular la esperanza condicional de Y dado X:

$$\mu_{Y/X}=\int_{-\infty}^{\infty} y\, f_{Y/X}(x,y)\, dx=\int_x^1 \frac{2y^2}{1-x^2}\, dy=\frac{2(x^2+x+1)}{3(x+1)}$$