Distribución Conjunta y Marginal

Índice

1 Introducción 3
2 Conceptos Básicos (Rincón, 2010) 3
2.1 Variables Aleatorias 3
2.2 Variable Aleatoria Discreta 3
2.3 Variable Aleatoria Continua 4
2.4 Función de Probabilidad para una Variable Discreta (Rincón, 2010) 4
2.5 Función de Densidad para una Variable Continua 4
3 Distribución Conjunta 4
3.1 Función de Densidad para una Variable Continua 5
3.2 Densidad Conjunta 5
3.3 Variables Aleatorias Bidimensionales Discretas 5
3.4 Función Masa o de Densidad Probabilidad Conjunta 5
3.5 Variables Aleatorias Bidimensionales Continuas 6
4 Distribución Marginal 7
4.1 Función de Distribución Marginal 8
4.2 Función de Densidad Marginal 8
5 Resumen 9
6 Referencias Bibliográficas 10

Objetivos

- Objetivo 1: Conocer los conceptos básicos de la teoría de variables aleatorias.
- Objetivo 2: Estudiar las variables aleatorias bidimensionales.

1 Introducción

Iniciamos esta unidad de aprendizaje con un repaso por los conceptos básicos de la teoría de variables aleatorias, con el fin de enlazar los contenidos de la última unidad de aprendizaje vista en el curso de Estadística I.

Siguiendo con la orientación objeto del curso, nuestro interés ahora se centrará en determinar la forma en cómo se estudian variables aleatorias bidimensionales y la interacción que consigo lleva su medición. También es de interés final medir el grado, intensidad y el sentido de asociación entre las características de interés.

2 Conceptos Básicos (Rincón, 2010)

2.1 Variables Aleatorias

Definido un experimento aleatorio de interés, por ejemplo el lanzamiento de un dado, los posibles resultados, los números del uno al seis, son en consecuencia valores numéricos. Sin embargo, se podrían asignar a los valores pares un criterio de éxito denotado por un número, digamos uno y en caso contrario un cero.

"En una variable aleatoria se puede ver una aplicación X del espacio de resultados Ω al conjunto de números reales"

Este tipo de asignación de los valores numéricos a los sucesos de un experimento aleatorio determina la base sobre la cual se define una variable aleatoria. Así una *variable aleatoria* se puede ver una aplicación $\boldsymbol{X}$ del espacio de resultados $\boldsymbol{\Omega}$ al conjunto de números reales, esto es,

$$X: \Omega \to \mathbb{R}$$

A menudo se escribe simplemente $v.a.$ en lugar del término variable aleatoria para efectos de notación.

En general, las variables aleatorias se denotan usando las últimas letras del alfabeto en mayúsculas, U, V, W, X, Y, Z y para un valor cualquiera de ellas se usa la misma letra pero en minúscula.

2.2 Variable Aleatoria Discreta

"Una variable aleatoria es discreta cuando oma un valor particular y este resulta ser un valor entero"

Decimos que una variable aleatoria es *discreta*, cuando toma un valor particular y este resulta de ser un valor entero.

Ejemplo:

Una variable aleatoria discreta puede ser la edad, debido a que toma valores en el conjunto {0, 1, 2,....., n} que es un conjunto discreto porque es finito.

2.3 Variable Aleatoria Continua

"Una variable aleatoria es continua cuando toma todo los valores dentro de un intervalo"

Decimos que una variable aleatoria es *continua* cuando toma todos los valores dentro de un intervalo $(\mathrm{a,b}) \subseteq \mathbb{R}$. Esta clasificación de variables aleatorias es limitada ya que existen variables aleatorias que no son absolutamente continuas o absolutamente discretas, es decir, pueden tomar valores de ambos conjuntos numéricos.

2.4 Función de Probabilidad para una Variable Discreta (Rincón, 2010)

Dada una variable aleatoria X discreta que toma los valores en un conjunto finito o numerable y con probabilidades no nulas. La *función de probabilidad* de la variable X denotada por $f(x): R \rightarrow [0,\infty)$ se define como sigue:

$$f(x) = \begin{cases} P(X = x) & si, x = x_1, x_2, \ldots \\ 0 & c.o.c \end{cases}$$

Acorde con la definición, tenemos ahora las siguientes propiedades de la función de densidad de probabilidad:

- $0 \leq f(x) \leq 1$
- $\sum_{i=1}^{\infty} f(x_i) = 1$

2.5 Función de Densidad para una Variable Continua

Dada una variable aleatoria X continua. Decimos que la función integrable y no negativa $f(x): \mathbb{R} \rightarrow [0,\infty)$, es la *función de densidad* de X si para cualquier intervalo (a,b) de $\mathbb{R}$ se cumple la igualdad

$$P\big(X \in (a,b)\big) = \int_a^b f(x)\, dx$$

Acorde con la definición, tenemos ahora las siguientes propiedades de la función de densidad de probabilidad:

- $f(x) \geq 0$, para toda $x \in \mathbb{R}$.
- $\int_{-\infty}^{\infty} f(x)\, dx = 1$.

3 Distribución Conjunta

Como extensión natural del caso univariado, todo par de variables aleatorias induce una medida de probabilidad. Esta medida de probabilidad puede analizarse, de igual forma que en el caso de estudio de una sola variable, mediante la función de distribución conjunta definida como sigue:

3.1 Función de Densidad para una Variable Continua

"Propiedadesde la función de densidad para una variable continua"

La función de distribución conjunta de un par de variables aleatorias dado (X,Y), denotada por $\boldsymbol{F}(x,y)$, se define como sigue

$$\boldsymbol{F}(x,y) = \boldsymbol{P}(\boldsymbol{X} \leq x, \boldsymbol{Y} \leq y).$$

El número $\boldsymbol{F(x,y)}$ no es más que la probabilidad de que la variable bidimensional tome valores en los intervalos cruzados $(-\infty,\boldsymbol{x}] \times (-\infty,\boldsymbol{y}]$. A $\boldsymbol{F(x,y)}$ se le suele conocer también como función de distribución *bidimensional* de $\boldsymbol{X}$ e $\boldsymbol{Y}$.

Las funciones de distribución bidimensionales cumplen propiedades similares a los escenarios presentados en la teoría univariada, se muestran a continuación algunas de estas propiedades.

- $\mathbf{lim}_{x,y\to\infty} \boldsymbol{F(x,y)} = \mathbf{1}$, ambas variables
- $\mathbf{lim}_{x,y\to-\infty} \boldsymbol{F(x,y)} = \mathbf{0}$, alguna de las variables.
- $\boldsymbol{F(x,y)}$ es no decreciente en cada variable.
- $\boldsymbol{F(x,y)}$ es continua por la derecha en cada variable.
- Si $\boldsymbol{a_1 < b_1 \ y \ a_2 < b_2}$, entonces $\boldsymbol{F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2)} \geq \mathbf{0}$.

3.2 Densidad Conjunta

Como en el caso univariado, existen variables aleatorias bidimensionales asociadas a otra función llamada, función de densidad o masa de probabilidad, y que consigo lleva una interpretación de la medición de dos características a un mismo individuo ya sea de manera separada o como las componentes de un vector en este caso bivariado.

3.3 Variables Aleatorias Bidimensionales Discretas

Un par de variables aleatorias dado $(\boldsymbol{X},\boldsymbol{Y})$ se le dice absolutamente discreta si sus componentes $\boldsymbol{X}$ y $\boldsymbol{Y}$ son variables aleatorias discretas. Entonces al suponer que $\boldsymbol{X}$ y $\boldsymbol{Y}$ tomen valores $\boldsymbol{x_i}$ y $\boldsymbol{y_j}$ con $(\boldsymbol{i,j} = \mathbf{1,2,3,\ldots})$ con sus respectivas probabilidades inducidas por naturaleza $\boldsymbol{P(X = x_i)}$ y $\boldsymbol{P(Y = y_j)}$, se podrá definir su función asociada, función masa o de densidad conjunta para esta variable.

3.4 Función Masa o de Densidad Probabilidad Conjunta

La función masa o de densidad de probabilidad conjunta de un par de variables aleatorias dado $(\boldsymbol{X},\boldsymbol{Y})$, denotada por $\boldsymbol{f(x,y)}$, viene dada por:

$$f(x,y) = P(X = x, Y = y)$$

Es de notar que al ser función de probabilidad para una variable aleatoria bidimensional de tipo discreto debe cumplir las siguientes propiedades:

- $f(x, y) \geq 0$.
- $\sum_{x,y} f(x, y) = 1$.

Ejemplo 1: Considere el par de variables aleatorias dado $(\boldsymbol{X}, \boldsymbol{Y})$, con función de densidad dada por la siguiente tabla.

x/y	0	1
-1	0.2	0.1
1	0.5	0.2

Tabla 1: *ejemplo 1*

De la tabla se deduce que la variable X toma valores del conjunto $\{-1,1\}$, mientras que Y toma valores en $\{0,1\}$. Además las probabilidades conjuntas están dadas por las entradas de la tabla.

Ejemplo: $P(X = -1, Y = 0) = 0.2$, esto es, la probabilidad de que X tome el valor de -1 y al mismo tiempo Y tome el valor 0 es 0.2. El resto de la información puede escribirse de la siguiente manera.

$$f(x, y) = P(X = x, Y = y) = \begin{cases} 0.2 & Si\ x = -1, y = 0, \\ 0.1 & Si\ x = -1, y = 1, \\ 0.5 & Si\ \ x = 1, y = 0, \\ 0.2 & Si\ x = 1, y = 1, \\ 0 & otro\ caso \end{cases}$$

Como todos estos valores son probabilidades, naturalmente son no negativos, y todos ellos suman uno.

3.5 Variables Aleatorias Bidimensionales Continuas

Un par de variables aleatorias dadas $(\boldsymbol{X}, \boldsymbol{Y})$ se le dice absolutamente continua si existe una función no negativa e integrable $\boldsymbol{f(x, y)}$, tal que, para todo $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^2$ la función de distribución conjunta del par de variables aleatorias $(\boldsymbol{X}, \boldsymbol{Y})$ puede expresarse de la siguiente manera

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\, du\, dv$$

En consecuencia la función $\boldsymbol{f(x, y)}$, se le llama función masa o de densidad conjunta de $\boldsymbol{X}$ y $\boldsymbol{Y}$.

Luego al determinar la función masa o de densidad de probabilidad del par de variables aleatorias $(\boldsymbol{X}, \boldsymbol{Y})$, esta debe cumplir las siguientes propiedades:

- $f(x, y) \geq 0$.
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$.

Ejemplo 2: Dado el par de variables aleatorias $(\boldsymbol{X}, \boldsymbol{Y})$, con función de densidad conjunta dada por:

$$f(x, y) = \begin{cases} k\left(\frac{xy}{2} + 1\right) & Si\ 0 < x < 1, -1 < y < 1. \\ 0 & Cualquier\ caso \end{cases}$$

Calculemos $\boldsymbol{k}$ de tal forma que $\boldsymbol{f(x, y)}$ sea función de masa o de densidad de probabilidad, entonces tenemos que verificar que:

- $f(x, y) \geq 0$, **si y sólo si** $k \geq 0$.
- Como $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\ dx\, dy = 1$.

Así veamos que:

$$k \int_{0}^{1} \int_{-1}^{1} \left(\frac{xy}{2} + 1\right) dy$$

$$= k \int_{0}^{1} \left[x \int_{-1}^{1} \frac{y}{2} dy + \int_{-1}^{1} dy\right] dx$$

$$= k \int_{0}^{1} \left[x \frac{y^2}{4} |_{-1}^{1} + y|_{-1}^{1}\right] dx$$

$$= k \int_{0}^{1} \left[\frac{x}{4}(1^2 - (-1)^2) + \left(1 - (-1)\right)\right] dx = k \int_{0}^{1} \left[\frac{x}{4} \cdot (0) + (2)\right] dx = k \int_{0}^{1} 2dx$$

$$= k \left[2 \int_{0}^{1} dx\right] = k[2x|_{0}^{1}] = k[2(1 - 0)] = 2k = 1$$

Por tanto $\boldsymbol{k} = \mathbf{1/2}$ y la función de masa o de densidad conjunta viene dada por:

$$f(x, y) = \begin{cases} \frac{xy + 2}{4} & Si\ 0 < x < 1, -1 < y < 1. \\ 0 & Cualquier\ caso \end{cases}$$

4 Distribución Marginal

"Si la función de distribución F(x, y) es dada, es posible obtener la función de distribución de cada una de las variables aleatorias independientes"

En los escenarios donde la función de distribución F(x,y) del par de variables aleatorias (X,Y), es dada, se hace posible obtener la función de distribución de cada una de las variables aleatorias independientes, de la siguiente manera.

4.1 Función de Distribución Marginal

Dado el par de variables aleatorias $(\boldsymbol{X}, \boldsymbol{Y})$, con función de distribución $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y})$. Entonces la función:

$$F(x) = \lim_{y \to \infty} F(x, y).$$

Se le llama función de distribución marginal de $\boldsymbol{X}$. De igual forma es posible definir la función de distribución marginal de $\boldsymbol{Y}$ como sigue:

$$F(y) = \lim_{x \to \infty} F(x, y).$$

4.2 Función de Densidad Marginal

Dado el par de variables aleatorias $(\boldsymbol{X}, \boldsymbol{Y})$, absolutamente continuas con función masa o de densidad de probabilidad $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$. Entonces la función:

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

Se llama función de densidad marginal de $\boldsymbol{X}$. De igual forma es posible definir la función de densidad marginal de $\boldsymbol{Y}$ como sigue:

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

Nota: Si el vector es de tipo discreto solo basta con reemplazar las integrales con términos de sumatorias para que las definiciones cobren sentido para tal caso.

Ejemplo 3: Para los datos de la tabla del ejemplo 1, tenemos que la marginal de $\boldsymbol{X}$ vendría dada por:

x	1	1
$P(X = x)$	0.3	0.7

Tabla 2: *Marginal de* $\boldsymbol{x}$

De manera similar, la marginal de $\boldsymbol{Y}$ vendría dada por:

y	0	1
$P(Y = y)$	0.7	0.3

Tabla 3: *Marginal de* $\boldsymbol{y}$

Ejemplo 4: En concordancia con los datos del ejemplo 2 y a partir de su función masa o de densidad de probabilidad se tiene que la marginal de $\boldsymbol{X}$ vendría dada por:

$$f_1(x) = \int_{-1}^{1}\left(\frac{xy}{4}+\frac{1}{2}\right)\,dy = \frac{x}{4}\int_{-1}^{1} y dy + \frac{1}{2}\int_{-1}^{1} dy = \frac{x}{8}y^2|_{-1}^{1} + \frac{1}{2}y|_{-1}^{1}$$

$$= \frac{x}{8}(1^2-(-1)^2) + \frac{1}{2}\big(1-(-1)\big) = \frac{x}{8}\cdot(0) + \frac{1}{2}\cdot(2) = 1$$

Para valores de $\boldsymbol{X}$ dentro de su campo de definición. Análogamente para la marginal de $\boldsymbol{Y}$ tenemos que:

$$f_2(y)$$

$$= \int_{0}^{1}\left(\frac{xy}{4}+\frac{1}{2}\right)dx$$

$$= \frac{y}{4}\int_{0}^{1} x dx + \frac{1}{2}\int_{0}^{1} dx = \frac{y}{8}x^2|_{0}^{1} + \frac{1}{2}x|_{0}^{1} = \frac{y}{8}(1-0) + \frac{1}{2}(1-0) = \frac{y}{8}+\frac{1}{2} = \frac{y}{8}+\frac{4}{8}$$

$$= \frac{y+4}{8}$$

Para valores de $\boldsymbol{Y}$ dentro de su campo de definición.

5 Resumen

- La asignación de los valores numéricos a los sucesos de un experimento aleatorios determina la base sobre la cual se define una variable aleatoria.
- Decimos que una $\boldsymbol{v.a.}$ es *discreta* cuando toma un valor particular y este resulta ser un valor entero.
- Decimos que una variable aleatoria es continua cuando toma todos los valores dentro de un intervalo $(\mathrm{a},\mathrm{b}) \subseteq \mathbb{R}$.
- La función de probabilidad de la variable X denotada por $f(x): R \to [0,\infty)$ se define como sigue

$$f(x) = \begin{cases} P(X = x) & si, x = x_1, x_2, \dots \\ 0 & c.o.c \end{cases}$$

- Decimos que la función integrable y no negativa $f(x): \mathbb{R} \to [0,\infty)$, es la *función de densidad* de X si para cualquier intervalo (a,b) de $\mathbb{R}$ se cumple la igualdad

$$P\big(X \in (a,b)\big) = \int_a^b f(x)\,dx$$

6 Referencias Bibliográficas

- Miranda, I. E., Palacín, F., Sánchez, M. L., Márquez, M., Chía, A. R., Navas, A. S., y otros. (3ra. Edición 2006). Estadística Descriptiva y Probabilidad. Cádiz: Servicio de Publicaciones de la Universidad de Cádiz.
- Montgomery, D., & R., R. (2da. Edición 2008). Probabilidad y Estadística Aplicada a la Ingeniería. México: Limusa Wiley.
- Rincón, L. (2010). Curso elemental de Probabilidad y Estadística. México: Circuito Exterior de CU.
- Walpole, R., Myers, R., & Myers, S. y. (2007). Probabilidad y Estadística para Ingeniería y Ciencias. México: Pearson.