SEP
SECRETARÍA DE EDUCACIÓN PÚBLICA

Instituto Politécnico Nacional

Escuela Superior de Física y Matemáticas

CARTA CESIÓN DE DERECHOS

En la Ciudad de México el día 12 del mes de Mayo del año 2017, el que suscribe Gabriel Steeven Cruz Méndez alumno del Programa Académico de Licenciatura en Física y Matemáticas con número de boleta 2013330031 adscrito a la Escuela Superior de Física y Matemáticas, manifiesta que es autor intelectual del presente trabajo de Tesis bajo la dirección del Dr. José María Rocha Martínez y cede los derechos del trabajo titulado "Distribuciones Bivariadas Discretas", al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección gcruzm1202@gmail.com Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

Gabriel Steeven Cruz Méndez
Nombre y firma del alumno

INSTITUTO POLITÉCNICO NACIONAL

Escuela Superior de Física y Matemáticas

Distribuciones bivariadas discretas

TESIS

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN FÍSICA Y MATEMÁTICAS

PRESENTA
GABRIEL STEEVEN CRUZ MÉNDEZ

Director de Tesis
Dr. José María Rocha Martínez

Ciudad de México Mayo de 2017

Agradecimientos

A *Gabriela Mireya Fuentes Corona*, porque aunque en un sentido puramente biológico seas mi abuela, en la práctica siempre has sido mi madre. No hay suficientes palabras, ni páginas para agradecer todo lo que has hecho y dado por mi. Gracias por estar ahí siempre, en todos esos momentos afortunados y desafortunados que te hice pasar, todas esas desveladas en que me esperabas mientras hacía mi trabajo, por estar a mi lado incluso en aquellos días sin noche. Con magna alegría recuerdo siempre tus palabras y enseñanzas. Gracias por vestirme, alimentarme, educarme y cuidarme, por darme toda tu atención, por enderezarme y llevarme siempre por buen camino, nada de todo esto tendría sentido de otra forma. Todo lo que soy y puedo llegar a ser te lo debo a ti. ¡Te quiero mucho abue, gracias!

A *Gilberto Badillo Balderas*, mi padre. Por tus consejos y enseñanzas para enfrentar la vida. Gracias por todos esos momentos tan divertidos, por ese gran sentido del humor que te caracteriza, tus palabras han marcado mi vida de alguna forma y me han hecho ver claramente las cosas en muchos sentidos. Te agradezco también por todo el apoyo, atención y cuidados que me has dado siempre.

A *Jessica Mireya Méndez Fuentes*, por darme la vida, porque eres mi mejor amiga, te quiero. A *Jorge Eduardo Betancourt Méndez*, mi hermano. Gracias por todo su apoyo, por estar ahí en los momentos más significativos.

A mis profesores, particularmente al Dr. *José María Rocha Martínez*, por su tiempo, atención y dedicación en la dirección de este trabajo, por sus enseñanzas y por esas maravillosas cátedras de análisis matemático. ¡Muchas gracias!

A los integrantes del jurado, por su valiosa contribución y observaciones.

A mis amigos y compañeros, por todo el apoyo.

Índice general

Introducción

Las distribuciones discretas univariadas son usadas frecuentemente en diversas aplicaciones tanto teóricas como prácticas en probabilidad y en otras áreas de estudio. En cualquier texto introductorio de probabilidad o estadística es seguro encontrarse con las siguientes distribuciones: de Bernoulli, binomial, de Poisson, hipergeométrica, geométrica y binomial negativa. La distribución de Bernoulli modela experimentos aleatorios en los que sólo se observa uno de dos resultados posibles, identificados como "éxito" o "fracaso". Un ejemplo clásico es el experimento que consiste en lanzar una moneda, el resultado observable es "águila" o "sol". Las restantes distribuciones pueden ser obtenidas a partir de la distribución de Bernoulli bajo ciertas operaciones o procesos al límite sobre muestreos aleatorios (finitos o infinitos) en los cuales cada observación corresponde al resultado de un ensayo de Bernoulli.

En diversas situaciones se puede estar interesado, en cambio, en experimentos en los que el número de resultados posibles de cada prueba sea más de dos, en estos casos se usa, entre otras, la distribución multinomial para el análisis de los problemas. Sin embargo, es más natural y útil trabajar con vectores aleatorios, digamos k-dimensionales, en los que cada una de sus componentes sea una variable aleatoria con distribución de Bernoulli, aportando así un total de 2^k resultados posibles en cada prueba. En este trabajo, por cierto, nos restringiremos al caso $k = 2$, es decir, experimentos en los que cada prueba arroja uno de cuatro posibles resultados. Un vector aleatorio con estas características se dirá que tiene distribución de Bernoulli bivariada. Una extensión de las distribuciones discretas clásicas antes mencionadas a vectores aleatorios bidimensional resultará conveniente, ya que así será posible analizar casos análogos pero con mayor alcance o cobertura de atributos. Piense, por ejemplo, en extraer aleatoriamente una muestra de una urna con una cantidad finita de bolas blancas y negras, en el caso bivariado podrá extender este experimento a uno en el que además cada una de las bolas está numerada, de modo que en cada prueba será observada una bola blanca o negra y de número par o impar. En general, esta extensión al caso bivariado permitirá clasificar a los elementos de una población mediante cuatro atributos distinguibles en lugar de solamente dos, como se acostumbra en la literatura y/o en cursos introductorios de probabilidad. Desde luego que lo anterior podría ser ampliado aún más al caso de vectores aleatorios k-dimensionales, pero como la idea básica desarrollada en el caso bivariado sería esencialmente la misma para el caso general multivariado (aunque esto aumentaría substancialmente la complejidad y longitud de las fórmulas involucradas) este caso será omitido.

En esta tesis se demuestra que es posible generar de manera natural una familia de distribuciones bivariadas discretas a partir de experimentos de Bernoulli bivariados, procediendo de manera análoga al caso univariado. Más precisamente, de la teoría de probabilidad se sabe que, en el caso univariado, la distribución binomial surge de un muestreo aleatorio finito en el que cada observación corresponde a un experimento de Bernoulli; la hipergeométrica surge de forma similar a la binomial salvo que se trata de un muestreo aleatorio sin reemplazo extraído desde una población finita; las distribuciones geométrica y binomial negativa surgen de una sucesión infinita de experimentos de Bernoulli; la distribución de Poisson es obtenida de un experimento binomial en el que el muestreo es muy grande (tiende a infinito). Para el caso bivariado, se muestra que las versiones análogas a las distribuciones binomial, de Poisson, hipergeométrica, geométrica y binomial negativa son originadas de manera similar a partir de la distribución de Bernoulli bivariada. En el sentido de convergencia, como es bien sabido, la distribución de Poisson es obtenida como límite desde las distribuciones binomial y binomial negativa, así mismo, la distribución binomial es límite de la distribución hipergeométrica y, finalmente, la distribución normal es límite de las distribuciones binomial y de Poisson.

Los objetivos que se persiguen en esta tesis son los siguientes: obtener distribuciones bivariadas discretas análogas al caso univariado, que de hecho compartirán el mismo nombre, partiendo de la distribución de Bernoulli bivariada; obtener algunas de las cantidades estadísticas de interés significativo, tales como el coeficiente de correlación y los valores esperados condicionales; demostrar que los resultados de convergencia entre ciertas distribuciones se cumplen también entre sus análogos en el caso bivariado; y por último, proporcionar ejemplos en diversas situaciones que ilustren y muestren el uso potencial de los modelos presentados. Los resultados de esta tesis se basan esencialmente en el trabajo desarrollado por Marshall y Olkin [9].

La tesis se divide en tres capítulos. El primero corresponde a una exposición sobre todos los resultados fundamentales de la teoría de probabilidad que son utilizados en el desarrollo de las demostraciones que componen los capítulos restantes, la sección titulada "Distribuciones discretas" es de interés especial porque ahí se muestran todas las distribuciones discretas univariadas que serán extendidas, el desarrollo es análogo a la exposición de la tesis e incluye las aproximaciones que serán discutidas posteriormente.

En el segundo capítulo se muestra la derivación correspondiente de las distribuciones bivariadas de Bernoulli, binomial y de Poisson, para las dos primeras se discute un criterio de independencia entre las variables aleatorias que caracterizan al vector aleatorio mediante el que se definen las distribuciones. En el tercer capítulo, se muestran las distribuciones hipergeométrica, geométrica y binomial negativa bivariadas. Para la mayoría de las distribuciones se determinan la función generadora de momentos, coeficiente de correlación y valores esperados condicionales. Dichas distribuciones bivariadas son ilustradas mediante algunos ejemplos de aplicación en situaciones comunes, por ejemplo, un estudio estadístico sobre accidentes automovilísticos y sobre diagnósticos encontrados en un hospital, por mencionar algunos. Los resultados de mayor relevancia en el trabajo son aquellos que corresponden a las aproximaciones, todos son discutidos y demostrados en un apartado especial dentro de cada sección.

Los cálculos numéricos que se exponen fueron realizados con apoyo del siguiente software: *"Wolfram Mathematica"*. Al final de la tesis se incluye un apartado que ilustra una forma en la que habrá de escribirse el código en *Mathematica* para realizar los cálculos correspondientes a cada una de las distribuciones bivariadas.

Fueron de gran utilidad los trabajos de Hamdan y Jensen [13] y de Wicksell en [11] y [12]. La obra de Subrahmaniam y Kathleen Kocherlakota [8], a diferencia de esta tesis, es una colección de trabajos que muestran las distribuciones que aquí se discuten en formas alternativas, el enfoque está orientado a la obtención de cantidades y problemas estadísticos (por ejemplo, prueba de hipótesis) que aquí no son tratados, se recomienda porque ahí es posible encontrar una fuente extensa de referencias que puede ser de provecho para profundizar en diversos temas concernientes a las distribuciones bivariadas discretas.

Capítulo 1

Preliminares de probabilidad

En este capítulo se presentan los principales resultados de la teoría de probabilidad que son utilizados en el desarrollo del capítulo 2.

1.1. Fundamentos

Asumiremos que para el lector son conocidos los fundamentos de la teoría de la probabilidad y que posee conocimientos elementales de análisis combinatorio. Las demostraciones de los resultados que serán enunciados en toda esta sección pueden ser consultadas en [1], salvo en algunos casos en los que se indicará la cita específica para consultar la demostración correspondiente.

Un **espacio de probabilidad** es una tripleta $(\Omega, \mathfrak{F}, P)$, donde Ω es un conjunto, al que llamaremos **espacio muestral** y a cuyos elementos $\omega \in \Omega$ se les llama **eventos simples** o **puntos muestrales** (los cuales corresponden usualmente a todos los resultados posibles derivados de algún experimento aleatorio que se esté analizando); $\mathfrak{F}$ es una sigma-álgebra de subconjuntos de Ω, cuyos elementos $A \in \mathfrak{F}$ se llaman **eventos**; y P es una medida de probabilidad sobre $\mathfrak{F}$.

Definición 1.1 *Se dice que un espacio de probabilidad* $(\Omega, \mathfrak{F}, P)$ *es* **discreto** *si* Ω *contiene una cantidad a lo sumo numerable de puntos muestrales y que es* **finito** *si, de hecho,* Ω *es un conjunto finito.*

En el caso de espacios de probabilidad discretos, se suele tomar a $\mathfrak{F}$ como el conjunto potencia $\mathcal{P}(\Omega)$ de Ω.

Un espacio de probabilidad finito donde todos los puntos muestrales tienen una misma probabilidad de ocurrir se dice que es **equiprobable** o **uniforme**. En este caso, la probabilidad de un evento se calcula contando el número de puntos muestrales que conforman a dicho evento y dividiéndolo entre el total de puntos muestrales del espacio. A esta forma de calcular la probabilidad de un evento se le conoce como **método del punto muestral**.

En este trabajo aparecerán frecuentemente espacios de probabilidad finitos equiprobables. Para el cálculo de la probabilidad de eventos en tales espacios se requieren pues de algunos resultados básicos de análisis combinatorio. El siguiente teorema establece un principio básico de conteo.

Teorema 1.2 (Regla del producto.) *Dados* n_1 *elementos* $a_1, a_2, \ldots, a_{n_1}$*; y* n_2 *elementos* $b_1, b_2, \ldots, b_{n_2}$*; y …; y* n_r *elementos* $c_1, c_2, \ldots, c_{n_r}$*, es posible formar* $n_1 \cdot n_2 \cdots n_r$ *r-eadas* $(a_{i_1}, b_{i_2}, \ldots, c_{i_r})$ *ordenadas que* *contengan un elemento de cada tipo.*

Un objeto de uso obligado en la teoría de probabilidad es el que corresponde a la siguiente definición, se utiliza para contar el número de *combinaciones* de n objetos tomados r a la vez.

Definición 1.3 (Coeficiente binomial.) *Si $r, n \in \mathbb{Z}$ positivos, entonces se define el coeficiente binomial*

$$\binom{n}{r} = \frac{n!}{(n-r)!\,r!}.$$

El siguiente resultado es utilizado para determinar el número de subconjuntos que se pueden formar al dividir un conjunto de n objetos distintos en k grupos que no se traslapen.

Teorema 1.4 (Coeficiente multinomial.) *Sean $n_1, n_2, \ldots, n_k \in \mathbb{Z}$ tales que $n_1 + n_2 + \cdots + n_k = n$ ($n_i \geq 0$). El número de formas en las que se puede dividir una población de tamaño n en k grupos distintos que contienen $n_1, n_2, \ldots, n_k$ objetos, respectivamente, donde cada objeto aparece en exactamente un grupo, es*

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!\,n_2!\,\cdots\,n_k!}.$$

A veces se cuenta con información parcial sobre cierto fenómeno y se requiere saber cómo esto afecta la probabilidad de los eventos en los que se esté interesado. Es decir, la probabilidad de un evento en ocasiones depende de si sabemos que han ocurrido otros eventos.

Definición 1.5 (Probabilidad condicional.) *Sean $A, B \subset \Omega$ dos eventos y $P[B] > 0$. Se define la* **probabilidad condicional** *de A dado que B ha ocurrido o simplemente* **la probabilidad de A dado B**, *como*

$$P[A \mid B] = \frac{P[A \cap B]}{P[B]}.$$

Teorema 1.6 *Si B es un evento tal que $P[B] > 0$, entonces la función $P[\,\cdot \mid B]$ es una medida de probabilidad sobre la sigma-álgebra de los eventos.*

Si la probabilidad de que ocurra un evento no es afectada porque ocurra o no otro evento se dirá entonces que estos son independientes. Formalmente se establece dicha propiedad mediante la siguiente definición.

Definición 1.7 *Se dice que dos eventos A y B son* **independientes** *si $P[A \cap B] = P[A]P[B]$.*

1.1.1. Variables aleatorias discretas y funciones de probabilidad

Una variable aleatoria es una función que asigna valores numéricos a los resultados de un experimento aleatorio. Similarmente, un vector aleatorio es una función que asigna valores vectoriales a los resultados de un experimento aleatorio. Frecuentemente, sobre todo en las definiciones y resultados mostrados, abreviaremos “variable aleatoria” y “vector aleatorio” como “v.a.” y “ve.a.”, respectivamente, y como “v.as.” y “ve.as.” cuando se trate en plurales.

Definición 1.8 *Una* **variable aleatoria** *(***v.a.***) X es una función medible del espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (es decir, para todo $\omega \in \Omega, X(\omega) \in \mathbb{R}$), donde $\mathcal{B}(\mathbb{R})$ es la sigma-álgebra de Borel de $\mathbb{R}$. Un* **vector aleatorio bidimensional** *(X, Y) es una función medible del espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ en $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ (es decir, para todo $\omega \in \Omega, (X, Y)(\omega) = (X(\omega), Y(\omega)) \in \mathbb{R}^2$), donde $\mathcal{B}(\mathbb{R}^2)$ es la sigma-álgebra de Borel de $\mathbb{R}^2$. En este caso se sabe que ambas funciones X y Y deben ser variables aleatorias sobre $(\Omega, \mathfrak{F}, P)$.*

Definición 1.9 *Se dice que una variable aleatoria X es* **discreta** *si puede tomar sólo una cantidad a lo sumo numerable de valores distintos. Si X y Y son variables aleatorias discretas, se dice que el vector aleatorio bidimensional (X, Y) es* **discreto***.*

Por ser una variable aleatoria X una función medible, los conjuntos siguientes deben ser eventos

$$\{\omega \in \Omega \mid X(\omega) = c\}, \quad \{\omega \in \Omega \mid X(\omega) \leq b\}, \quad \{\omega \in \Omega \mid X(\omega) > a\}, \quad \{\omega \in \Omega \mid a < X(\omega) \leq b\}.$$

Definición 1.10 *Si X es una variable aleatoria, se define su* **función de distribución** *(o* **función de distribución acumulativa***) como*

$$F_X(x) = P[X \leq x] = P[\{\omega \in \Omega \mid X(\omega) \leq x\}], \quad \forall x \in \mathbb{R}.$$

Definición 1.11 *La* **función de probabilidad** *(o* **distribución de probabilidad** *o* **función de densidad***) de una v.a. discreta X que asume los valores $x_1, x_2, \ldots$ es*

$$p(x_i) = P[X = x_i] = P[\{\omega \in \Omega \mid X(\omega) = x_i\}], \quad i = 1, 2, \ldots.$$

Teorema 1.12 *Una función $p(x)$ es la función de probabilidad (o de densidad) de alguna v.a. discreta X si y sólo si $p(x) \geq 0$, $\forall x$,*

$$\sum_{i=1}^{\infty} p(x_i) = 1, \qquad y \qquad p(x) = 0 \quad si \quad x \notin \{x_1, x_2, \ldots\}.$$

Las definiciones que se mostrarán a continuación pueden ser encontradas en textos comunes de probabilidad y estadística, por ejemplo [2] o [3], mientras que las versiones multivariadas pueden ser consultadas en [1]. Por simplicidad, en este trabajo se mencionará *"vector aleatorio"* en lugar de *"vector aleatorio bidimensional"*.

Definición 1.13 *Si (X, Y) es un vector aleatorio, se define su* **función de distribución conjunta (bivariada)** *como*

$$F_{X,Y}(x, y) = P[X \leq x, Y \leq y] = P[\{\omega \in \Omega \mid X(\omega) \leq x\} \cap \{\omega \in \Omega \mid Y(\omega) \leq y\}], \quad \forall x, y.$$

Definición 1.14 *Si X y Y son v.as. discretas, se define la* **función de probabilidad conjunta** *(o* **distribución de probabilidad conjunta** *o* **función de densidad conjunta***) para el vector aleatorio (X, Y) como*

$$p(x, y) = P[X = x, Y = y] = P[\{\omega \in \Omega \mid X(\omega) = x\} \cap \{\omega \in \Omega \mid Y(\omega) = y\}], \quad \forall x, y.$$

Definición 1.15 *Si X y Y son v.as. discretas con función de probabilidad conjunta $p(x, y)$, se define la* **función de probabilidad marginal** *(o* **distribución de probabilidad marginal** *o* **función de densidad marginal***) de X y Y, respectivamente, como*

$$p(x) = P[X = x] = \sum_{y} p(x, y), \quad \forall y,$$
$$p(y) = P[Y = y] = \sum_{x} p(x, y), \quad \forall x.$$

Definición 1.16 *Si X y Y son v.as. discretas con función de probabilidad conjunta $p(x, y)$ y funciones de probabilidad marginales $p(x)$ y $p(y)$, respectivamente, se define la* **función de probabilidad condicional** *de X dada $Y = y$ como*

$$P[X = x \mid Y = y] = \frac{P[X = x, Y = y]}{P[Y = y]},$$

siempre que $P[Y = y] > 0$.

Así mismo, la función de probabilidad condicional de Y dada $X = x$ se define como

$$P[Y = y \mid X = x] = \frac{P[Y = y, X = x]}{P[X = x]},$$

siempre que $P[X = x] > 0$.

Definición 1.17 *Si X y Y son v.as. discretas con función de probabilidad conjunta $p(x, y)$ y funciones de probabilidad marginales $p(x)$ y $p(y)$, respectivamente. Se dice que X y Y son* **independientes** *si y sólo si para todo par (x, y)*

$$F_{X,Y}(x, y) = F_X(x)F_Y(y).$$

En la práctica es más "cómodo" utilizar el siguiente resultado para establecer independencia entre variables aleatorias en términos de sus distribuciones de probabilidad.

Teorema 1.18 *Si X y Y son v.as. discretas con función de probabilidad conjunta $p(x, y)$ y funciones de probabilidad marginales $p(x)$ y $p(y)$, respectivamente. Entonces X y Y son independientes si y sólo si para todo par (x, y)*

$$p(x, y) = p(x)p(y).$$

Es posible apreciar que el resultado anterior coincide con la caracterización de independencia entre eventos dada por la Definición 1.7, simplemente nótese que $p(x, y) = p(x)p(y)$ es exactamente lo mismo que $P[A \cap B] = P[A]P[B]$ para los eventos $A = \{\omega \in \Omega \mid X(\omega) = x\}$ y $B = \{\omega \in \Omega \mid Y(\omega) = y\}$.

1.1.2. Momentos

Las variables aleatorias pueden ser tratadas y caracterizadas con eficacia para diversos propósitos mediante la consideración de sus "momentos". Los momentos que se mostrarán en este apartado corresponden a las cantidades más relevantes en el desarrollo de las distribuciones mostradas en el Capítulo 2, a saber: valor esperado, media, varianza, covarianza, coeficiente de correlación parcial, valor esperado condicional y por último uno de los instrumentos más importantes, las funciones generadoras de momentos.

Observación 1.19 En este apartado las demostraciones de los Teoremas 1.21 y 1.27 pueden ser consultadas en [2]. ◁

Definición 1.20 *Sea X una v.a. discreta con función de probabilidad $p(x)$. Se define el* **valor esperado** *de X como*

$$E[X] = \sum_x xp(x)$$

donde la suma considera todos los posibles valores que asume X. Se dice que $E[X]$ existe siempre que la suma sea absolutamente convergente.

Teorema 1.21 *Sea X una v.a. discreta con función de probabilidad $p(x)$ y sea $g : \mathbb{R} \to \mathbb{R}$ una función. Entonces $g(X)$ es una variable aleatoria discreta tal que*

$$E[g(X)] = \sum_x g(x)p(x).$$

Definición 1.22 *Sea X una v.a. con función de probabilidad $p(x)$. Se definen*

i. *El* n-**ésimo momento (no central)** *de X como $\mu_n = E[X^n]$ en caso de existir.*

ii. *El* n-**ésimo momento central** *de X como $\mu_n^* = E\left[(X - E[X])^n\right]$ en caso de existir.*

Definición 1.23 *Si X es una v.a. con valor esperado finito se define la* **varianza** *de X, $Var[X]$, como el segundo momento central de X. Esto es*

$$Var[X] = E\left[(X - E[X])^2\right].$$

Teorema 1.24 *Sea X una v.a. discreta con función de probabilidad $p(x)$ y sean $g_1, \ldots, g_k$ funciones reales. Entonces $g_1(X), \ldots, g_k(X)$ son v.as. discretas y*

$$E\left[\sum_{i=1}^k g_i(X)\right] = \sum_{i=1}^k E[g_i(x)].$$

Proposición 1.25 *Se cumple* $\quad Var[X] = E[X^2] - E[X]^2.$

Teorema 1.26 *Sean $(X_1, \ldots, X_n)$ un vector aleatorio discreto con función de probabilidad conjunta $p(x_1, \ldots, x_n)$ y $g : \mathbb{R}^n \to \mathbb{R}$ una función. Se define el valor esperado de $g(X_1, \ldots, X_n)$ como*

$$E[g(X_1, \ldots, X_n)] = \sum_{x_1} \cdots \sum_{x_n} g(X_1, \ldots, X_n)p(x_1, \ldots, x_n)$$

donde las sumas consideran todos los posibles valores que asumen $X_1, \ldots, X_n$. Se dice que $E[g(X_1, \ldots, X_n)]$ existe siempre que la suma sea absolutamente convergente.

Teorema 1.27 *Sean X y Y dos v.as. con valores esperados finitos. Entonces*

i. *El valor esperado de $X + Y$ existe y $E[X + Y] = E[X] + E[Y]$.*

ii. *Si además X y Y son independientes, entonces el valor esperado de XY existe y $E[XY] = E[X]E[Y]$.*

Teorema 1.28 *Sean X y Y dos v.as. independientes y sean h, g dos funciones reales. Entonces $h(Y), g(X)$ son v.as. discretas y*

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

siempre que las esperanzas de la derecha sean finitas.

Definición 1.29 *Si X y Y son v.as. con valores esperados finitos. Se definen*

i. La **covarianza** *de X y Y como*

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])]. \tag{1.1}$$

ii. El **coeficiente de correlación** *de X y Y como*

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var[X]\,Var[Y]}}. \tag{1.2}$$

Proposición 1.30 *Se cumple* $Cov(X,Y) = E[XY] - E[X]E[Y]$.

Teorema 1.31 *Se cumple*

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{i<j} Cov(X_i, X_j).$$

Teorema 1.32 *Si X y Y son v.as. independientes, entonces $Cov(X,Y) = 0$.*

Teorema 1.33 *Si X y Y son v.as. con valores esperados finitos, entonces $-1 \leq Corr(X,Y) \leq 1$.*

Definición 1.34 *Sean X y Y dos v.as. discretas arbitrarias. El* **valor esperado condicional** *de X dado que $Y = y$ es, considerando todos los posibles valores x que asume X,*

$$E[X \mid Y = y] = \sum_{x} xP[X = x \mid Y = y].$$

Así mismo el valor esperado condicional de Y dado que $X = x$ es, considerando todos los posibles valores y que asume Y,

$$E[Y \mid X = x] = \sum_{y} yP[Y = y \mid X = x].$$

A continuación se enuncia una definición alternativa y más general del valor esperado condicional, está dada en términos de variables aleatorias y eventos en Ω.

Definición 1.35 *Sea X una v.a. y $B \subset \Omega$ un evento. Se define el* **valor esperado condicional** *de X dado B, considerando todos los posibles valores x que asume X, como*

$$E[X \mid B] = \frac{E[X \cdot I_B]}{P[B]} = \sum_{x} x\frac{P[X = x, I_B = 1]}{P[B]},$$

siempre que $P[B] > 0$ y donde

$$I_B(\omega) = \begin{cases} 1 & si\ \omega \in B, \\ 0 & en\ otro\ caso. \end{cases}$$

es la función indicadora del evento B.

Teorema 1.36 *Sean X_1 y X_2 dos v.as. y $\mathbf{Y} = (Y_1, \ldots, Y_k)$, $\mathbf{Z} = (Z_1, \ldots, Z_k)$ dos vectores aleatorios k-dimensionales. Se tienen*

i. $E[X_1 + X_2 \mid \mathbf{Y}] = E[X_1 \mid \mathbf{Y}] + E[X_2 \mid \mathbf{Y}]$.

ii. $E\left[E[X_1 \mid \mathbf{Y}, \mathbf{Z}] \,\middle|\, \mathbf{Y}\right] = E[X_1 \mid \mathbf{Y}]$.

Definición 1.37 *La* **función generadora de momentos (*f.g.m.*)** *de una v.a. X se define (en caso de existir), para todo $t \in \mathbb{R}$ tal que $|t| < \epsilon$, para algún $\epsilon > 0$, como $\psi_X(t) = E\left[e^{tX}\right]$.*

Teorema 1.38 *Si $X_1, \ldots, X_n$ son v.as. independientes y $Z = \sum_{i=1}^n X_i$, entonces*

$$\psi_Z(t) = \prod_{i=1}^n \psi_{X_i}(t).$$

Además, si dichas v.as. son idénticamente distribuidas, entonces

$$\psi_Z(t) = (\psi_{X_1}(t))^n .$$

Definición 1.39 *La* ***f.g.m.* conjunta** *de un vector aleatorio $(X_1, \ldots, X_n)$ n-dimensional se define (en caso de existir), para todo $t_1, \ldots, t_n \in \mathbb{R}$ tales que $|t_i| < \epsilon$, para cada $i = 1, \ldots, n$ y para algún $\epsilon > 0$, como*

$$\psi_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = E\left[e^{t_1 X_1 + \cdots + t_n X_n}\right].$$

Teorema 1.40 *Suponga que la f.g.m. del vector aleatorio (X, Y) existe. Entonces X y Y son v.as. independientes si y sólo si $\psi_{X,Y}(s, t) = \psi_X(s)\psi_Y(t)$ para todo $s, t \in \mathbb{R}$ tales que $|s| < \epsilon$ y $|t| < \epsilon$ $(\epsilon > 0)$.*

Definición 1.41 *La* **función generadora de momentos factoriales (*f.g.m.f.*)** *de una v.a. X se define (en caso de existir), para todo $t \in \mathbb{R}$ tal que $|t| < \epsilon$, para algún $\epsilon > 0$, como $m_X(t) = E\left[t^X\right]$.*

Lema 1.42 (Relación entre funciones generadoras de momentos) *Si la f.g.m. y la f.g.m.f. de una v.a. X existen, entonces se tiene que*

$$\psi_X(\ln(t)) = m_X(t).$$

Definición 1.43 *La* ***f.g.m.f.* conjunta** *de un vector aleatorio $(X_1, \ldots, X_n)$ n-dimensional se define (en caso de existir), para todo $t_1, \ldots, t_n \in \mathbb{R}$ tales que $|t_i| < \epsilon$, para cada $i = 1, \ldots, n$ y para algún $\epsilon > 0$, como*

$$m_{X_1,\ldots,X_n}(t_1, \ldots, t_n) = E\left[t_1^{X_1} \cdots t_n^{X_n}\right].$$

En lo que resta de este apartado se muestran algunas extensiones de interés significativo en el desarrollo del Capítulo 2.

Una extensión natural de (*i.*) del Teorema 1.27 y del Teorema 1.31 es el siguiente resultado.

Teorema 1.44 *Si $X_1, \ldots, X_n$ son v.as. independientes con valores esperados finitos. Entonces*

$$E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i],$$

y

$$Var\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n Var[X_i].$$

El siguiente resultado es una extensión de (*i.*) del Teorema 1.36.

Teorema 1.45 *Sean X_1 y X_2 dos v.as. y $\mathbf{Y} = (Y_1, \ldots, Y_k)$ un vector aleatorio k-dimensional. Se tiene*

$$E\left[\sum_{i=1}^{n} X_i \mid \mathbf{Y}\right] = \sum_{i=1}^{n} E[X_i \mid \mathbf{Y}].$$

Teorema 1.46 *Sea (X, Y) un vector aleatorio. Denotamos (suponiendo que son finitas) a las medias y varianzas de X y Y por μ_1, μ_2 y σ_1^2, σ_2^2, respectivamente, y sea $\rho_{x,y}$ el coeficiente de correlación entre X y Y. Si $E[X \mid Y]$ es lineal respecto de Y, entonces*

$$E[X \mid Y] = \mu_1 + \rho_{x,y}\,\frac{\sigma_1}{\sigma_2}(Y - \mu_2),$$

y

$$E[Y \mid X] = \mu_2 + \rho_{x,y}\,\frac{\sigma_2}{\sigma_1}(X - \mu_1).$$

Observación 1.47 Las demostraciones de los Teoremas 1.44 y 1.45 pueden ser consultadas en [4], mientras que la demostración del Teorema 1.46 puede ser consultada en [5].$\triangleleft$

El siguiente resultado corresponde a una extensión natural del Lema 1.42, su uso implica inmediatamente los resultados equivalentes a los Teoremas 1.38 y 1.40 para el caso de las *f.g.m.f.*

Lema 1.48 (Relación entre funciones generadoras de momentos conjuntas) *Si (X, Y) es un vector aleatorio para el que la f.g.m. y la f.g.m.f. conjuntas existen, entonces se tiene que*

$$\psi_{X,Y}\left(\ln(s), \ln(t)\right) = m_{X,Y}(s, t).$$

Teorema 1.49 *Si $X_1, \ldots, X_n$ son v.as. independientes y $Z = \sum_{i=1}^{n} X_i$, entonces*

$$m_Z(t) = \prod_{i=1}^{n} m_{X_i}(t).$$

Además, si dichas v.as. son idénticamente distribuidas, entonces

$$m_Z(t) = \left(m_{X_1}(t)\right)^n.$$

Teorema 1.50 *Suponga que la f.g.m.f. del vector aleatorio (X, Y) existe. Entonces X y Y son v.as. independientes si y sólo si $m_{X,Y}(s, t) = m_X(s)m_Y(t)$ para todo $s, t \in \mathbb{R}$ tales que $|s| < \epsilon$ y $|t| < \epsilon$ $(\epsilon > 0)$.*

Teorema 1.51 *Suponga que la f.g.m.f. del vector aleatorio (X, Y) existe. Entonces*

$$E[XY] = \left.\frac{\partial^2 m_{X,Y}(s, t)}{\partial s\, \partial t}\right|_{s=1, t=1}.$$

1.1.3. La distribución normal y el teorema de límite central

Definición 1.52 *Se dice que una v.a. X tiene la* **distribución normal** $N(\mu,\sigma^2)$ *si para $\sigma > 0$ y $-\infty < \mu < \infty$, la función de densidad de X es*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty.$$

Se escribe $X \sim N(\mu,\sigma^2)$.

La distribución $N(0,1)$ es conocida como la **distribución normal estándar**.

Teorema 1.53 *Si una v.a. $X \sim N(\mu,\sigma^2)$, entonces $E[X] = \mu$ y $Var[X] = \sigma^2$.*

Definición 1.54 *Se dice que un vector aleatorio $\mathbf{X} = (X_1,\ldots,X_n)$ n-dimensional tiene la* **distribución normal multivariada de orden n**, *abreviado $N_n(\overrightarrow{\mu},\Sigma)$, si su función de densidad conjunta es*

$$f(x_1,\ldots,x_n) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}}\, e^{-\frac{1}{2}\left(\overrightarrow{X}-\overrightarrow{\mu}\right)'\Sigma^{-1}\left(\overrightarrow{X}-\overrightarrow{\mu}\right)}, \quad \forall\, (x_1,\ldots,x_n) \in \mathbb{R}^n,$$

donde

$$\overrightarrow{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad y \quad \overrightarrow{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$$

es un vector de tamaño n donde $\mu_1,\ldots,\mu_n \in \mathbb{R}$ y

$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,n} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n,1} & \sigma_{n,2} & \cdots & \sigma_{n,n} \end{bmatrix}$$

es cualquier matriz $n \times n$ simétrica y definida positiva donde $\sigma_{i,j} \in \mathbb{R}$ para todo i,j. Se escribe $\mathbf{X} \sim N_n(\overrightarrow{\mu},\Sigma)$.

La distribución $N_n\left(\overrightarrow{0}, I\right)$, donde I es la matriz identidad, es conocida como la **distribución normal multivariada estándar de orden n**.

Teorema 1.55 *Si un vector aleatorio $\mathbf{X} \sim N_n(\overrightarrow{\mu},\Sigma)$, entonces $E\left[\mathbf{X}\right] = \overrightarrow{\mu}$ y*

$$\Sigma = \begin{bmatrix} Cov(X_1,X_1) & Cov(X_1,X_2) & \cdots & Cov(X_1,X_n) \\ Cov(X_2,X_1) & Cov(X_2,X_2) & \cdots & Cov(X_2,X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n,X_1) & Cov(X_n,X_2) & \cdots & Cov(X_n,X_n) \end{bmatrix}$$

es la **matriz de covarianzas** *del vector aleatorio* $\mathbf{X}$.

Definición 1.56 *Sea $\{Z_n\}_{n=1}^{\infty}$ una sucesión de v.as. definidas sobre un espacio de probabilidad $(\Omega,\mathfrak{F},P)$ y sea Z otra v.a. definida en ese mismo espacio de probabilidad. Sean F_{Z_n} y F_Z las funciones de distribución de Z_n y Z, respectivamente. Se dice que* $\mathbf{Z_n}$ *converge a* $\mathbf{Z}$ **en distribución**, *se escribe $\mathbf{Z_n} \xrightarrow{\mathbf{d}} \mathbf{Z}$ si*

$$\lim_{n\to\infty} F_{Z_n}(t) = F_Z(t)$$

en cada punto t donde F_Z es continua.

Teorema 1.57 *Sea $\{Z_n\}_{n=1}^{\infty}$ una sucesión de v.as. con f.g.m. $\psi_{Z_n}(t)$ que existe para $|t| < \epsilon$ ($\epsilon > 0$) y para todo n. Sea Z una v.a. con f.g.m. $\psi_Z(t)$ que existe para $|t| \leq \epsilon_1 \leq \epsilon$. Si $\lim_{n\to\infty} \psi_{Z_n}(t) = \psi_Z(t)$, entonces $Z_n \xrightarrow{d} Z$.*

Definición 1.58 *Sea $\{\mathbf{Z}_n\}_{n=1}^{\infty}$ una sucesión de vectores aleatorios con función de distribución $F_{\mathbf{Z}_n}$ y sea $\mathbf{Z}$ otro vector aleatorio con función de distribución $F_{\mathbf{Z}}$. Se dice que $\mathbf{Z_n}$* **converge a Z en distribución***, se escribe $\mathbf{Z_n} \xrightarrow{\mathbf{d}} \mathbf{Z}$ si*

$$\lim_{n\to\infty} F_{\mathbf{Z}_n}(\mathbf{t}) = F_{\mathbf{Z}}(\mathbf{t})$$

en cada punto $\mathbf{t}$ donde $F_{\mathbf{Z}}$ es continua.

Teorema 1.59 *Sea $\{\mathbf{Z}_n\}_{n=1}^{\infty}$ una sucesión de vectores aleatorios con f.g.m. conjunta $\psi_{\mathbf{Z}_n}(\mathbf{t})$. Sea $\mathbf{Z}$ un vector aleatorio con f.g.m. conjunta $\psi_{\mathbf{Z}}(\mathbf{t})$. Entonces $\mathbf{Z}_n \xrightarrow{d} \mathbf{Z}$ si y sólo si para algún $\epsilon > 0$*

$$\lim_{n\to\infty} \psi_{\mathbf{Z}_n}(\mathbf{t}) = \psi_{\mathbf{Z}}(\mathbf{t})$$

para todo $\mathbf{t}$ tal que $||\mathbf{t}|| < \epsilon$.

Teorema 1.60 (Teorema de límite central) *Sea $\{X_i\}_{i=1}^{\infty}$ una sucesión de v.as. independientes e idénticamente distribuidas, cada una con valores esperados finitos $E[X_1] = \mu$ y varianzas $Var[X_1] = \sigma^2$ no nulas, entonces*

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n\sigma^2}} \quad \xrightarrow{\quad d \quad} \quad N(0,1) \quad \text{cuando } n \to \infty.$$

Teorema 1.61 (Teorema de límite central multivariado) *Sea $\{\mathbf{X}_i\}_{i=1}^{\infty}$ una sucesión de vectores aleatorios p-dimensionales independientes e idénticamente distribuidos, cada uno con $E[\mathbf{X}_1] = \overrightarrow{\mu}$ y matriz de covarianzas Σ, entonces*

$$\frac{\sum_{i=1}^{n} \mathbf{X}_i - n\overrightarrow{\mu}}{\sqrt{n}} \quad \xrightarrow{\quad d \quad} \quad N_p\left(\overrightarrow{0}, \Sigma\right) \quad \text{cuando } n \to \infty.$$

Teorema 1.62 *Si un vector aleatorio (X, Y) tiene la distribución normal bivariada $N_2(\overrightarrow{\mu}, \Sigma)$, entonces su f.g.m. está dada por*

$$\psi_{X,Y}(s,t) = \exp\left[\mathbf{T}\overrightarrow{\mu} \ + \ \frac{1}{2}\mathbf{T}\Sigma\mathbf{T}'\right]$$

donde $\mathbf{T} = [s\ t]$ es una matriz 1×2.

Teorema 1.63 *Si un vector aleatorio $\mathbf{X} \sim N_n(\overrightarrow{\mu}, \Sigma)$ y $\mathbf{B}$ es una matriz $p \times n$ ($p \leq n$) de rango p, entonces el vector $\mathbf{Z} = \mathbf{BX}$ tiene la distribución normal multivariada de orden p con $E[\mathbf{Z}] = \mathbf{B}\overrightarrow{\mu}$ y matriz de covarianzas $\mathbf{B}\Sigma\mathbf{B}'$.*

Observación 1.64 Las demostraciones del Teorema 1.60, así como su versión multivariada, el Teorema 1.61, pueden ser consultadas en [5], la demostración del Teorema 1.62 puede ser consultada en [7] y la del Teorema 1.63 se encuentra en [6]. $\lhd$

1.2. Distribuciones discretas

En esta sección se presentan las distribuciones discretas clásicas. El orden en que se exponen corresponde al mismo en que se muestran las distribuciones análogas del Capítulo 2, los teoremas enunciados indican las cantidades típicas de interés; a saber, valor esperado, varianza, *f.g.m.* y *f.g.m.f.* Se recomienda que consulte [7] para ver las demostraciones correspondientes, tómese en cuenta que ahí no se exponen las *f.g.m.f.*, pero tales cantidades se siguen inmediatamente de las *f.g.m.* con una simple aplicación del Lema 1.42.

Los ensayos repetidos e independientes se llaman de *Bernoulli* cuando, en cada ensayo, sólo hay dos posibles resultados y sus probabilidades son las mismas en cada ensayo. A uno de tales resultados se le llama "éxito" y se denota por p a la probabilidad de dicho resultado, el otro resultado por supuesto que tendrá probabilidad $(1-p)$ y se le conoce como "fracaso".

Definición 1.65 *Se dice que una v.a. X tiene la* **distribución de Bernoulli** *si, para algún $0 \leq p \leq 1$, llamado el* **parámetro** *de la distribución, se tiene que*

$$P[X=x] = \begin{cases} p^x(1-p)^{1-x} & si \quad x = 0, 1, \\ 0 & en\ otro\ caso. \end{cases}$$

Teorema 1.66 *Si una v.a. X tiene la distribución de Bernoulli, entonces*

$$E[X] = p, \quad Var[X] = p(1-p), \quad \psi_X(t) = pe^t + q \quad y \quad m_X(t) = pt + q,$$

donde $q = 1 - p$.

En ocasiones nos interesa solamente el número total de éxitos obtenidos en una sucesión de n ensayos de Bernoulli, al margen del orden en que se presenten. El número de éxitos puede ser $0, 1, \ldots, n$. La siguiente definición trata con n ensayos que resultan en x éxitos y $n - x$ fracasos.

Definición 1.67 *Se dice que una v.a. X tiene la* **distribución binomial** *si, para algún entero positivo n y para algún $0 \leq p \leq 1$, se tiene que*

$$P[X=x] = \begin{cases} \binom{n}{x} p^x(1-p)^{n-x} & si \quad x = 0, 1, \ldots, n, \\ 0 & en\ otro\ caso. \end{cases}$$

Se escribe $X \sim Binomial(n,p)$, donde n y p son llamados los **parámetros** *de la distribución.*

Teorema 1.68 *Si $X \sim Binomial(n,p)$, entonces*

$$E[X] = np, \quad Var[X] = np(1-p), \quad \psi_X(t) = (pe^t + q)^n \quad y \quad m_X(t) = (pt+q)^n,$$

donde $q = 1 - p$.

Observación 1.69

i. Una variable aleatoria binomial puede ser considerada como una suma de n variables aleatorias Bernoulli, es decir, como el número de éxitos en n ensayos Bernoulli.

ii. Cuando un muestreo se realiza desde una población finita, la distribución binomial surge sólo cuando el muestreo se realiza *con reemplazo.* $\triangleleft$

Definición 1.70 *Se dice que una v.a.* X *tiene la* **distribución de Poisson** *si, para algún* $\lambda > 0$, *llamado el* **parámetro** *de la distribución, se tiene que*

$$P[X = x] = \begin{cases} \dfrac{\lambda^x}{x!} e^{-\lambda} & si \quad x = 0, 1, \ldots, \\ 0 & en\ otro\ caso. \end{cases}$$

Se escribe $X \sim Poisson(\lambda)$.

Teorema 1.71 *Si* $X \sim Poisson(\lambda)$, *entonces*

$$E[X] = Var[X] = \lambda, \quad \psi_X(t) = e^{\exp(t)-1} \quad y \quad m_X(t) = e^{t-1}.$$

Suponga que una población contiene un número finito N de elementos que poseen dos características distintivas digamos A y B; a son del tipo A y el resto $N - a$ del tipo B. Una muestra aleatoria de tamaño $n \leq N$ se selecciona de la población *sin reemplazo.* Si X es la variable aleatoria que corresponde al número de elementos con la característica A en la muestra, entonces X tiene la distribución siguiente.

Definición 1.72 *Se dice que una v.a.* X *tiene la* **distribución hipergeométrica** *si, para* $n, a, N \in \mathbb{Z}$, *llamados los* **parámetros** *de la distribución, con* $1 \leq n \leq N$ *y* $0 \leq a \leq N$, *se tiene que*

$$P[X = x] = \begin{cases} \dfrac{\dbinom{a}{x}\dbinom{N-a}{n-x}}{\dbinom{N}{n}} & si \quad x = 0, 1, \ldots, n, \\ 0 & en\ otro\ caso. \end{cases}$$

Las distribuciones que se muestran en la siguientes dos definiciones comparten características similares a las de un experimento binomial (número de "éxitos" observados en n pruebas). La primera corresponde a un experimento en el que la variable aleatoria cuenta el número, digamos x, de "fracasos" observados antes del primer "éxito" (de modo que el número de pruebas es $x + 1$). La segunda corresponde a un experimento en el que la variable aleatoria cuenta el número de "fracasos" observados antes del r-ésimo "éxito" ($x + r$ pruebas).

Definición 1.73 *Se dice que una v.a.* X *tiene la* **distribución geométrica** *si, para algún* $0 \leq p \leq 1$, *llamado el* **parámetro** *de la distribución, se tiene que*

$$P[X = x] = \begin{cases} p(1-p)^x & si \quad x = 0, 1, \ldots, \\ 0 & en\ otro\ caso. \end{cases}$$

Teorema 1.74 *Si una v.a.* X *tiene la distribución geométrica, entonces*

$$E[X] = \frac{1-p}{p}, \quad Var[X] = \frac{1-p}{p^2} \quad y \quad \psi_X(t) = \frac{p}{1-qe^t},$$

donde $q = 1 - p$.

Definición 1.75 *Se dice que una v.a. X tiene la* **distribución binomial negativa** *si, para algún entero $r \geq 1$, y para algún $0 \leq p \leq 1$, llamados los* **parámetros** *de la distribución, se tiene que*

$$P[X = x] = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x & \textit{si} \quad x = 0, 1, \ldots, \\ 0 & \textit{en otro caso.} \end{cases}$$

Observación 1.76 La distribución geométrica es un caso particular de la binomial negativa. Para comprobarlo simplemente habrá que tomar $r = 1$ en la Definición 1.75. $\triangleleft$

Es posible proporcionar dos definiciones alternativas para estas dos últimas distribuciones, todo depende del punto de vista o el objeto de la investigación. La geométrica puede ser dada en términos de una variable aleatoria que cuenta el número de pruebas hasta observar el primer "éxito", en cuyo caso se tendría que

$$P[X = x] = \begin{cases} p(1-p)^{x-1} & \text{si} \quad x = 1, 2, \ldots, \\ 0 & \text{en otro caso.} \end{cases}$$

Por otro lado, la binomial negativa puede ser dada en términos de una variable aleatoria que cuenta el número de pruebas hasta que se observa el r-ésimo "éxito", en cuyo caso se tendría que

$$P[X = x] = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{x-r} & \text{si} \quad x = r, r+1, r+2, \ldots, \\ 0 & \text{en otro caso.} \end{cases}$$

Una variable aleatoria con la distribución binomial surge de un experimento caracterizado porque en cada prueba hay dos posibles resultados, a los que llamamos "éxito" o "fracaso". En diversas ocasiones se encuentran situaciones similares pero en las que el número de posibles resultados por cada prueba son más de dos. Una distribución discreta multivariada bien conocida y tratada en la literatura clásica es la *multinomial*. Un experimento multinomial consta de n pruebas, cada una con k posibles resultados, caracterizado en términos de k variables aleatorias que cuentan el número de ocasiones en las que se observa el k-ésimo resultado respectivamente.

Definición 1.77 *Se dice que un vector aleatorio k-dimensional $(X_1, \ldots, X_k)$ tiene la* **distribución multinomial** *si, para $k, n \in \mathbb{Z}$ $(k \geq 2,\ n \geq 1)$ y $0 \leq p_i \leq 1$ $(i = 1, \ldots, k)$ tales que $\sum_{i=1}^{k} p_i = 1$, se tiene que*

$$P[X_1 = x_1, \ldots, X_k = x_k] = \begin{cases} \dfrac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} & \textit{si} \quad x_i = 0, 1, \ldots, n,\ i = 1, \ldots, k, \\ & \qquad\qquad x_1 + \cdots + x_k = n, \\ 0 & \textit{en otro caso.} \end{cases}$$

Tomando $k = 2$ es posible observar que la distribución binomial es un caso particular de la distribución multinomial. En [5] se trata el caso $k = 3$, es decir, la distribución *trinomial*.

1.2.1. Aproximaciones

Algunas distribuciones pueden ser aproximadas mediante otras bajo ciertas condiciones en los parámetros de sus funciones de probabilidad, en otros casos las aproximaciones son establecidas mediante una

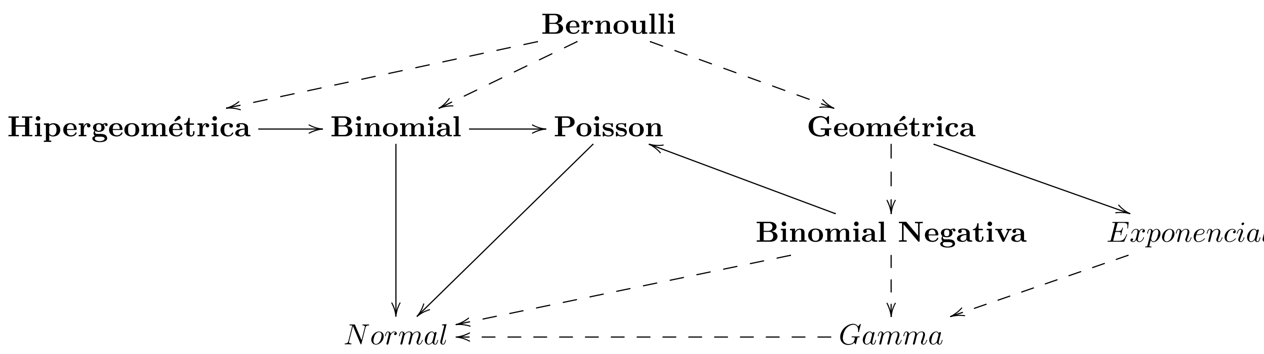

Figura 1.1: *Relación entre distribuciones.* Las distribuciones hipergeométrica, binomial y geométrica surgen directamente de la distribución de Bernoulli; la distribución binomial puede ser obtenida de la hipergeométrica; la distribución binomial negativa es obtenida de la geométrica; la distribución de Poisson se puede obtener de las distribuciones binomial y binomial negativa; la distribución exponencial puede ser obtenida de la geométrica; la distribución gamma es obtenida de la exponencial y puede ser obtenida de la binomial negativa; la distribución normal puede ser obtenida de las distribuciones binomial negativa, binomial y de Poisson.

convergencia en distribución. La Figura 1.1 muestra la relación que existe entre las distribuciones discretas, además incluye su relación con algunas distribuciones continuas bien conocidas; a saber, exponencial, gamma y normal. Las aproximaciones que serán discutidas en este apartado están indicadas mediante una linea continua, los resultados análogos para el caso bivariado se exponen en el Capítulo 2.

La *aproximación normal a la distibución binomial* es un resultado de discusión obligada en los textos de probabilidad y estadística, se puede encontrar como un ejemplo de aplicación del teorema de límite central en cualquiera de las siguientes referencias: [1], [2], [5]; en [3] se dedica un apartado especial para este tema que incluye ejemplos numéricos sumamente ilustrativos; en [7] se enuncia justamente como el teorema de límite de DeMoivre-Laplace. Los resultados restantes de aproximación pueden ser consultados en [1].

Teorema 1.78 (Aproximación normal a la distribución binomial) *Sea $X \sim Binomial(n,p)$. Se cumple, cuando $n \to \infty$,*

$$\frac{X - np}{\sqrt{np(1-p)}} \xrightarrow{\quad d \quad} N(0,1).$$

Teorema 1.79 (Aproximación de Poisson a la distribución binomial) *Sea $\{X_i\}_{i=1}^{\infty}$ una sucesión de v.as. independientes tal que X_i tiene la distribución de Bernoulli con parámetro p_n. Suponga que para algún $\lambda > 0$ se tiene que*

$$\lim_{n\to\infty} np_n = \lambda.$$

Sea

$$Z_n = \sum_{i=1}^{n} X_i, \quad \forall\, n \in \mathbb{N}.$$

Si $Z \sim Poisson(\lambda)$, entonces para $k = 0, 1, \ldots,$

$$\lim_{n\to\infty} P[Z_n = k] = P[Z = k].$$

Teorema 1.80 (Aproximación normal a la distribución de Poisson) *Sea $\{X_i\}_{i=1}^{\infty}$ una sucesión de v.as. independientes tal que $X_i \sim Poisson(\lambda)$ y sea*

$$Y_n = \frac{\left(\sum_{i=1}^{n} X_i - n\lambda\right)}{\sqrt{n\lambda}}.$$

Se tiene, cuando $n \to \infty$,

$$Y_n \quad \xrightarrow{\quad d \quad} \quad N(0,1).$$

Teorema 1.81 (Aproximación binomial a la distribución hipergeométrica) *Sea Z una v.a. con distribución hipergeométrica. Entonces, cuando $N \to \infty$, $a \to \infty$ y $N - a \to \infty$ de modo que*

$$\lim_{N\to\infty} \frac{a}{N} = p,$$

se tiene que

$$\lim_{N\to\infty} P[Z = k] = P[Z^* = k],$$

donde $Z^* \sim Binomial(n, p)$.

Los dos resultados siguientes no se encuentran en alguna de las referencias que se otorgan en este trabajo por lo que serán incluidas las demostraciones correspondientes.

Teorema 1.82 (Aproximación exponencial a la distribución geométrica) *Sea $p_n = \lambda/n$ para todo $n \in \mathbb{N}$ y $\lambda > 0$ arbitrario. Si Z es una v.a. con la distribución geométrica de parámetro p_n y Z^* es una v.a. con la distribución exponencial de parámetro λ y $n \to \infty$, se cumple $Z/n \xrightarrow{\ d\ } Z^*$.*

Demostración. Considere la $f.g.m.$ de Z/n, $\psi_{Z/n}(t)$, y sea $t_n = t/n$ para todo $n \in \mathbb{N}$, entonces

$$\psi_{Z/n}(t) = \psi_Z(t_n) = \frac{p_n}{1-(1-p_n)e^{t_n}} = \frac{\frac{\lambda}{n}}{1-\left(1-\frac{\lambda}{n}\right)e^{\frac{t}{n}}} = \frac{\frac{\lambda}{n}}{1-e^{\frac{t}{n}}+\frac{\lambda}{n}e^{\frac{t}{n}}},$$

donde

$$1-e^{\frac{t}{n}} = 1-\sum_{i=0}^{\infty}\frac{1}{i!}\left(\frac{t}{n}\right)^i = -\left(\frac{t}{n}+\left(\frac{t}{n}\right)^2+\left(\frac{t}{n}\right)^3+\cdots\right),$$

por lo que

$$\begin{aligned}
\psi_Z(t_n) &= \frac{\frac{\lambda}{n}}{\frac{\lambda}{n}e^{\frac{t}{n}}-\left(\frac{t}{n}+\left(\frac{t}{n}\right)^2+\left(\frac{t}{n}\right)^3+\cdots\right)} \\
&= \frac{\frac{1}{n}(\lambda)}{\frac{1}{n}\left(\lambda e^{\frac{t}{n}}-t-\frac{t^2}{n}-\frac{t^3}{n^2}-\cdots\right)} \\
&= \frac{\lambda}{\lambda e^{\frac{t}{n}}-t-\frac{t^2}{n}-\frac{t^3}{n^2}-\cdots}.
\end{aligned}$$

Puesto que

$$\lim_{n\to\infty} e^{\frac{t}{n}} = 1,$$

se tiene

$$\lim_{n\to\infty} \psi_Z(t_n) = \frac{\lambda}{\lambda - t},$$

donde el último cociente es la *f.g.m.* $\psi_{Z^*}(t)$ de la distribución exponencial de parámetro λ. ∎

Teorema 1.83 (Aproximación de Poisson a la distribución binomial negativa) *Sea $\lambda > 0$ arbitrario. Para $r \geq 1$, se define*

$$p = \frac{r}{r+\lambda}.$$

Si Z es una v.a. con distribución binomial negativa de parámetros r y p, entonces

$$\lim_{r\to\infty} P[Z = k] = P[Z^* = k],$$

donde $Z^* \sim Poisson(\lambda)$.

Demostración. Se tiene que

$$\begin{aligned} P[Z = k] &= \binom{k+r-1}{r-1} p^r (1-p)^k = \frac{(k+r-1)!}{k!(r-1)!} p^r (1-p)^k \\ &= \frac{(k+r-1)(k+r-2)\cdots(r+1)r(r-1)!}{k!(r-1)!} p^r (1-p)^k \\ &= \frac{(k+r-1)(k+r-2)\cdots(r+1)r}{k!} \left(\frac{r}{r+\lambda}\right)^r \left(1 - \frac{r}{r+\lambda}\right)^k, \end{aligned}$$

donde $(k+r-1)(k+r-2)\cdots(r+1)r$ es un producto de $(k+r-1)-(r-1) = k$ factores, así que

$$(k+r-1)(k+r-2)\cdots(r+1)r = r^k \left(1 + \frac{k-1}{r}\right)\left(1 + \frac{k-2}{r}\right)\cdots\left(1 + \frac{1}{r}\right).$$

Luego

$$\begin{aligned} P[Z = k] &= \frac{r^k \left(1 + \frac{k-1}{r}\right)\left(1 + \frac{k-2}{r}\right)\cdots\left(1 + \frac{1}{r}\right)}{k!} \left(\frac{r}{r+\lambda}\right)^r \left(\frac{\lambda}{r+\lambda}\right)^k \\ &= \frac{r^k \left(1 + \frac{k-1}{r}\right)\left(1 + \frac{k-2}{r}\right)\cdots\left(1 + \frac{1}{r}\right)}{k!} \left(\frac{1}{1+\frac{\lambda}{r}}\right)^r \frac{\lambda^k}{r^k} \left(\frac{1}{1+\frac{\lambda}{r}}\right)^k \\ &= \frac{\lambda^k}{k!}\left(1 + \frac{k-1}{r}\right)\left(1 + \frac{k-2}{r}\right)\cdots\left(1 + \frac{1}{r}\right)\left(\frac{1}{1+\frac{\lambda}{r}}\right)^r \left(\frac{1}{1+\frac{\lambda}{r}}\right)^k \end{aligned}$$

y finalmente, dado que

$$\lim_{r\to\infty} \left(\frac{1}{1+\frac{\lambda}{r}}\right)^r = e^{-\lambda}$$

y que los factores restantes convergen a 1, se tiene que

$$\lim_{r\to\infty} P[Z = k] = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Por lo tanto

$$\lim_{r\to\infty} P[Z = k] = P[Z^* = k],$$

donde $Z^* \sim Poisson(\lambda)$. ∎

1.3. Algunos resultados útiles

La expansión multinomial. Sean $x_1, x_2, \ldots, x_k \in \mathbb{R}$ arbitrarios, entonces para todo $n \in \mathbb{N}$, si $b_1, b_2, \ldots, b_k \in \mathbb{Z}$ *no negativos* tales que $\sum_{i=1}^{k} b_i = n$ se cumple

$$(x_1 + x_2 + \cdots + x_k)^n = \sum_{b_1} \cdots \sum_{b_k} \binom{n}{b_1, b_2, \ldots, b_k} \prod_{j=1}^{k} x_j^{b_j} .$$

El teorema del binomio. Sean $a, b \in \mathbb{R}$ arbitrarios, entonces para todo $n \in \mathbb{N}$,

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^{n-i} b^i .$$

La expansión en serie de Taylor de $\mathbf{e^x}$. Para todo $x \in \mathbb{R}$ se tiene que

$$e^x = \sum_{i=1}^{\infty} \frac{x^i}{i!} .$$

La serie geométrica. Sea $r \in \mathbb{R}$ tal que $|r| < 1$, entonces

$$\sum_{i=0}^{\infty} r^i = \frac{1}{1-r},$$

cuya *suma parcial* está dada por

$$\sum_{i=0}^{m} r^i = \frac{1 - r^{m+1}}{1-r}, \quad \forall\, m \in \mathbb{Z}.$$

Aproximación a la función exponencial. Sea $a \in \mathbb{R}$ arbitrario, entonces

$$\lim_{x \to \infty} \left(1 + \frac{a}{x}\right)^x = e^a .$$

Capítulo 2

Distribuciones de Bernoulli, binomial y de Poisson bivariadas

2.1. La distribución de Bernoulli bivariada

Definición 2.1 *Se dice que un vector aleatorio* (X,Y) *con valores en* $\mathbb{R}^2$ *tiene* **distribución de Bernoulli bivariada** *si* (X,Y) *toma solamente los cuatro valores* $(1,1)$, $(1,0)$, $(0,1)$, $(0,0)$ *con probabilidades* $p(1,1)$, $p(1,0)$, $p(0,1)$, $p(0,0)$ *que serán denotadas simplemente como* p_{11}, p_{10}, p_{01}, p_{00}*, respectivamente.*

Observe que las distribuciones marginales de X y Y están dadas por

$$\begin{aligned} P[X=1] &= P[X=1,Y=0]+P[X=1,Y=1]=p_{10}+p_{11}=p_{1+}=1-p_{0+}=1-P[X=0], \\ P[Y=1] &= P[X=0,Y=1]+P[X=1,Y=1]=p_{01}+p_{11}=p_{+1}=1-p_{+0}=1-P[Y=0]. \end{aligned} \tag{2.1}$$

Proposición 2.2 *Si* (X,Y) *es un vector aleatorio con distribución de Bernoulli bivariada, entonces su covarianza está dada por*

$$Cov(X,Y)=p_{11}-p_{1+}p_{+1}=p_{11}p_{00}-p_{10}p_{01}=\tau \tag{2.2}$$

y su coeficiente de correlación, por

$$\rho=\frac{\tau}{\sqrt{p_{1+}\,p_{0+}\,p_{+1}\,p_{+0}}}. \tag{2.3}$$

Demostración. Sabemos, por definición, que

$$Cov(X,Y)=E[XY]-E[X]E[Y].$$

Calculemos pues los valores esperados $E[X]$, $E[Y]$ y $E[XY]$. Se tiene,

$$\begin{aligned} E[X] &= \sum_x x\,P(X=x)=0\cdot P(X=0)+1\cdot P(X=1)=p_{1+}, \\ E[Y] &= 0\cdot P(Y=0)+1\cdot P(Y=1)=p_{+1}, \\ E[XY] &= \sum_x\sum_y xy\,p(x,y)=p(1,1)=p_{11}, \end{aligned}$$

donde la última igualdad para $E[X]$ y $E[Y]$ queda justificada por (2.1). Así, concluimos que

$$Cov(X,Y) = p_{11} - p_{1+}p_{+1}.$$

Ahora, en la igualdad anterior, sustituimos p_{1+}, p_{+1} y utilizamos el hecho de que $p_{11}+p_{01}+p_{10}+p_{00} = 1$ para obtener

$$\begin{aligned}
p_{11} - p_{1+}p_{+1} &= p_{11} - (p_{11} + p_{10})(p_{11} + p_{01}) \\
&= p_{11} - p_{11}p_{11} - p_{11}p_{01} - p_{11}p_{10} - p_{10}p_{01} \\
&= p_{11}(1 - p_{11} - p_{01} - p_{10}) - p_{10}p_{01} \\
&= p_{11}[1 - (p_{11} + p_{01} + p_{10})] - p_{10}p_{01} \\
&= p_{11}p_{00} - p_{10}p_{01}.
\end{aligned}$$

Esto muestra la igualdad (2.2).

Para demostrar (2.3) vamos a recurrir a la definición de correlación,

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var[X]Var[Y]}}.$$

Puesto que $Var[X] = E[X^2] - E[X]^2$ y X es una variable aleatoria de Bernoulli, se tiene

$$E[X^2] = E[X] = p_{1+}.$$

Bajo el mismo argumento se tiene que $E[Y^2] = E[Y] = p_{+1}$. Por lo tanto,

$$Var[X] = p_{1+} - p_{1+}^2 = p_{1+}(1 - p_{1+}) = p_{1+}\, p_{0+} \qquad \text{y} \qquad Var[Y] = p_{+1}\, p_{+0},$$

de donde

$$Var[X]Var[Y] = p_{1+}\, p_{0+}\, p_{+1}\, p_{+0},$$

probando así (2.3). ∎

Observación 2.3 El coeficiente de correlación ρ toma el valor -1 cuando $p_{11} = p_{00} = 0$ y toma el valor 1 cuando $p_{10} = p_{01} = 0$.

En efecto, si $p_{11} = p_{00} = 0$, entonces $p_{10} + p_{01} = 1$. Recordando (2.1) y (2.2), se tendría que

$$\tau = -p_{10}\, p_{01}, \qquad p_{1+} = p_{10} \qquad \text{y} \qquad p_{+1} = p_{01},$$

luego

$$p_{0+} = 1 - p_{1+} \qquad \text{y} \qquad p_{+1} = 1 - p_{+0},$$

llegando a que

$$p_{0+} = 1 - p_{10} = p_{01} \qquad \text{y} \qquad p_{+0} = 1 - p_{01} = p_{10}.$$

Por lo tanto,

$$\sqrt{p_{1+}\, p_{0+}\, p_{+1}\, p_{+0}} = \sqrt{p_{10}\, p_{01}\, p_{01}\, p_{10}} = p_{10}\, p_{01},$$

lo cual implica que $\rho = -1$.

De forma similar, si $p_{10} = p_{01} = 0$, entonces $p_{11} + p_{00} = 1$, luego

$$\tau = p_{11}\, p_{00}, \qquad p_{1+} = p_{11} \qquad \text{y} \qquad p_{+1} = p_{11}.$$

Ya que

$$p_{0+} = 1 - p_{1+} \qquad \text{y} \qquad p_{+1} = 1 - p_{+0},$$

se concluye que

$$p_{0+} = 1 - p_{11} = p_{00} \qquad \text{y} \qquad p_{+0} = 1 - p_{11} = p_{00}.$$

Por lo tanto,

$$\sqrt{p_{1+}\, p_{0+}\, p_{+1}\, p_{+0}} = \sqrt{p_{11}\, p_{00}\, p_{11}\, p_{00}} = p_{11}\, p_{00}$$

lo cual implica que $\rho = 1$. $\triangleleft$

2.1.1. Independencia

Sea (X, Y) un vector aleatorio con distribución de Bernoulli bivariada. Recuerde que se está usando la notación $p_{xy} = p(x, y)$ para la distribución conjunta de (X, Y) y p_{x+} y p_{+y} para las distribuciones marginales de X y Y, respectivamente. Ya se sabe que si X y Y son independientes, entonces $\rho = 0$. Se demostrará la recíproca, es decir que si $\rho = 0$, entonces X y Y son independientes, es decir, que $p_{xy} = p_{x+}p_{+y}$, $\forall x, y$.

Proposición 2.4 *Si (X, Y) es un vector aleatorio con distribución de Bernoulli bivariada y $\rho = 0$ entonces X y Y son independientes.*

Demostración. Supongamos que $\rho = 0$, luego $\tau = 0$. De acuerdo a (2.3), se debe tener $p_{11} - p_{1+}p_{+1} = 0$ y por lo tanto $p_{11} = p_{1+}p_{+1}$.

Ahora, utilizando $p_{11} = p_{1+}p_{+1}$, probaremos que $p_{00} = p_{0+}p_{+0}$. En efecto, se tiene que

$$p_{11} + p_{01} + p_{10} + p_{00} = 1$$

implica

$$p_{11} + p_{01} + p_{10} + p_{00} + p_{11} = 1 + p_{11},$$

despejando a p_{00} se obtiene

$$p_{00} = 1 - p_{11} - p_{10} - p_{11} - p_{01} + p_{11}.$$

Así pues,

$$\begin{aligned} p_{00} &= 1 - (p_{11} + p_{10}) - (p_{11} + p_{01}) + p_{11} \\ &= 1 - p_{1+} - p_{+1} + p_{11} \\ &= 1 - p_{1+} - p_{+1} + p_{1+}p_{+1} \\ &= (1 - p_{1+})(1 - p_{+1}) = p_{0+}p_{+0}. \end{aligned}$$

Se concluye que

$$p_{00} = p_{0+}p_{+0}.$$

Volvemos a utilizar la identidad $p_{11} = p_{1+}p_{+1}$ en lo siguiente. Por (2.1), tenemos que

$$p_{10} = p_{1+} - p_{11} \qquad \text{y} \qquad p_{01} = p_{+1} - p_{11}$$

implica

$$\begin{aligned} p_{10} &= p_{1+} - p_{1+}p_{+1} = p_{1+}(1 - p_{+1}) = p_{1+}p_{+0}, \\ p_{01} &= p_{+1} - p_{1+}p_{+1} = (1 - p_{1+})p_{+1} = p_{0+}p_{+1}. \end{aligned}$$

Se ha probado pues que $p_{xy} = p_{x+}p_{+y}$, $\forall x, y$. Por lo tanto, X y Y son independientes. ∎

2.1.2. Ejemplos

Ejemplo 2.5 Se realizó un estudio de accidentes automovilísticos en los que un niño, menor de cinco años de edad, estaba en el auto y hubo al menos una persona muerta. El estudio se concentró, específicamente, en si el niño sobrevivió y si utilizaba el cinturón de seguridad. Los resultados obtenidos son:

i. Sobrevivió y no usaba el cinturón en el 32 % de los casos.

ii. No sobrevivió y no usaba el cinturón, 17 % de los casos.

iii. Sobrevivió y usaba el cinturón, 44 % de los casos.

iv. No sobrevivió y usaba el cinturón, 7 % de los casos.

Considere las siguientes variables aleatorias

$$X = \begin{cases} 1 & \text{si el niño sobrevivió,} \\ 0 & \text{si no,} \end{cases} \qquad \text{y} \quad Y = \begin{cases} 1 & \text{si el niño usaba cinturón,} \\ 0 & \text{si no.} \end{cases}$$

Entonces el vector aleatorio (X, Y) tiene distribución de Bernoulli bivariada. De acuerdo con los resultados del estudio es posible determinar los parámetros de la distribución (vea la Tabla 2.1, note que $p_{11} + p_{10} + p_{01} + p_{00} = 1$).

El grado de dependencia entre las v.as. X y Y es medido por medio del coeficiente de correlación, en este caso, dado por (2.3). Se obtuvo, sustituyendo las probabilidades dadas en la Tabla 2.1, $\rho = 0.2454$, lo cual indica que existe poca dependencia positiva entre las v.as. ◄

Un vector aleatorio con distribución de Bernoulli bivariada está caracterizado porque cada componente es una v.a. con distribución de Bernoulli, así es posible tratar algunos otros ejemplos aplicados en diversas áreas, como la industria, la economía o la medicina.

	$X = 1$	$X = 0$	
$Y = 1$	$p_{11} = 0.44$	$p_{01} = 0.07$	$p_{+1} = 0.51$
$Y = 0$	$p_{10} = 0.32$	$p_{00} = 0.17$	$p_{+0} = 0.49$
	$p_{1+} = 0.76$	$p_{0+} = 0.24$	1.00

Tabla 2.1: Accidentes automovilísticos. Parámetros de la distribución.

	$X=1$	$X=0$	
$Y=1$	$p_{11}=0.75$	$p_{01}=0.10$	$p_{+1}=0.85$
$Y=0$	$p_{10}=0.07$	$p_{00}=0.08$	$p_{+0}=0.15$
	$p_{1+}=0.82$	$p_{0+}=0.18$	1.00

Tabla 2.2: Casos clínicos. Parámetros de la distribución.

Ejemplo 2.6 Un producto industrial es sometido a dos pruebas para detectar fallas, las cuales se clasifican en dos tipos, digamos del tipo I y II. Una prueba detecta las fallas tipo I y otra las del tipo II. Si X y Y son v.as. con distribución de Bernoulli tales que se observa un "éxito" cuando el producto presenta fallas del tipo I y del tipo II, respectivamente, entonces (X, Y) es un vector aleatorio con distribución de Bernoulli bivariada. ◀

Ejemplo 2.7 Se prevé que en el mediano plazo sea imposible que el gobierno cubra los montos en pensiones para las personas que cuentan con este beneficio. Por ello es que se requiere un estudio de los individuos pertenecientes a la población económicamente activa, particularmente sobre si están afiliados al sistema de pensiones y si se encuentran próximos al retiro. Como en el ejemplo anterior es posible definir un par de variables aleatorias con distribución de Bernoulli, para las que un "éxito" es alcanzable cuando se observa que el individuo está afiliado al sistema de pensiones y próximo al retiro, respectivamente. De modo tal que dicha pareja de v.as. forma un vector aleatorio con distribución de Bernoulli bivariada. ◀

Ejemplo 2.8 En medicina un paciente que acude a consulta en un hospital por dolor en la región lumbar y/o dificultad al orinar es diagnosticado con litasis renal o urolitiasis (enfermedad caracterizada por la aparición de cálculos en el aparato urinario superior) por dos métodos, clínica o ultrasonido abdominal. El hospital realizó un estudio en el que se registraron todos los casos de pacientes que acudieron con los síntomas indicados para saber si fueron o no diagnosticados con litiasis renal en clínica y por un ultrasonido abdominal. Los resultados son:

i. El paciente fue diagnosticado por clínica y por un ultrasonido en el 75 % de los casos.

ii. Fue diagnosticado por clínica y no presentó signos de litiasis en un ultrasonido, 7 % de los casos.

iii. No manifestó los síntomas clínicos correspondientes y sí presentó signos de litiasis en un ultrasonido, 10 % de los casos.

iv. No manifestó los síntomas clínicos y tampoco presentó signos de litiasis en un ultrasonido, 8 % de los casos.

Considerando las variables aleatorias

$$X = \begin{cases} 1 & \text{si es diagnosticado por clínica,} \\ 0 & \text{si no presenta los síntomas clínicos correspondientes,} \end{cases}$$

y

$$Y = \begin{cases} 1 & \text{si es diagnosticado por un ultrasonido abdominal,} \\ 0 & \text{si no presenta signos de litiasis en el ultrasonido,} \end{cases}$$

se tendría que (X, Y) es un vector aleatorio con distribución de Bernoulli bivariada cuyos parámetros se muestran en la Tabla 2.2. La correlación entre X y Y, de acuerdo con (2.3), es $\rho = 0.3863$, por lo cual X y Y son poco dependientes. ◀

2.2. La distribución binomial bivariada

Una variable aleatoria univariada con distribución binomial se define como el número de éxitos en cierto número fijo de ensayos de Bernoulli. Siguiendo esta idea se define a continuación una posible versión de la distribución binomial en el caso bivariado.

2.2.1. Distribución binomial bivariada

Definición 2.9 *Sean $(X_1, Y_1), \ldots, (X_n, Y_n)$ vectores aleatorios independientes con distribución de Bernoulli bivariada. Se define el vector aleatorio*

$$(U, V) = \sum_{i=1}^{n} (X_i, Y_i),$$

donde

$$U = \sum_{i=1}^{n} X_i \qquad y \qquad V = \sum_{i=1}^{n} Y_i. \tag{2.4}$$

A la distribución conjunta de (U, V) dada, para $u, v = 0, 1, \ldots, n$, por

$$P[U = u, V = v] = \sum_{a=0}^{\min(u,v)} \binom{n}{a, u-a, v-a, n-u-v+a} p_{11}^{a}\, p_{10}^{u-a}\, p_{01}^{v-a}\, p_{00}^{n-u-v+a}, \tag{2.5}$$

se le llama **distribución binomial bivariada**.

Ejemplo 2.10 Una urna contiene M bolas. Se pueden encontrar bolas blancas o negras que han sido marcadas con un número que puede ser par o impar, a las bolas marcadas con un número par les llamaremos "*bolas pares*" y a aquellas marcadas con un número impar les llamaremos "*bolas impares*". El experimento consiste en seleccionar aleatoriamente de esta urna una bola y anotar sus atributos: blanca o negra, par o impar; la bola es devuelta en la urna para ser mezclada de nuevo junto con las demás. Tal experimento es repetido n veces, y nos preguntamos por la probabilidad de hallar el atributo blanca en u ocasiones y par en v ocasiones.

Definimos pues las variables aleatorias $X_1, \ldots, X_n$ y $Y_1, \ldots, Y_n$ como sigue:

$$X_i = \begin{cases} 1 & \text{si la bola seleccionada es blanca,} \\ 0 & \text{si la bola seleccionada es negra,} \end{cases} \qquad i \in \{1, \ldots, n\}$$

y

$$Y_i = \begin{cases} 1 & \text{si la bola seleccionada es par,} \\ 0 & \text{si la bola seleccionada es impar,} \end{cases} \qquad i \in \{1, \ldots, n\}.$$

Los vectores aleatorios $(X_1, Y_1), \ldots, (X_n, Y_n)$ modelan adecuadamente el experimento descrito si en la i-ésima selección, X_i indica el color de la bola y Y_i indica si la bola es par o impar. Para todo $i \in$

$\{1, \ldots, n\}$, $(X_1, Y_1), \ldots, (X_n, Y_n)$ son vectores aleatorios independientes con distribución de Bernoulli bivariada, donde X_i y Y_i son variables aleatorias de Bernoulli.

Para determinar la probabilidad por la que nos preguntamos tomamos la definición (2.4) donde U es el número de veces que se observó el atributo "blanca" y V es el número de veces que se observó el atributo "par". El número de bolas que muestran las distintas combinaciones de atributos en la urna es conocido, supongamos que estos números son:

Mp_{1+} bolas blancas y Mp_{0+} bolas negras;

Mp_{+1} bolas pares y Mp_{+0} bolas impares;

Mp_{11} bolas blancas y pares;

Mp_{01} bolas negras y pares;

Mp_{10} bolas blancas e impares;

Mp_{00} bolas negras e impares,

donde p_{11}, p_{01}, p_{10}, p_{00} son las probabilidades de seleccionar una bola blanca y par, una bola negra y par, una bola blanca e impar, una bola negra e impar, respectivamente.

Cada punto muestral de Ω estará caracterizado por un arreglo de n parejas, donde cada pareja aparece con cierta probabilidad. Denotaremos por BP, BI, NP, NI a los atributos anotados en cada selección, bola blanca y par, bola blanca e impar, bola negra y par, bola negra e impar, respectivamente. Un punto muestral típico será de la siguiente forma

$$\underbrace{BI, BP, BP, NP, BP, NI, NI, NP, NI, BP, BI, \ldots, BP, NI}_{n \text{ posiciones}}$$

Hacemos un reordenamiento apropiado de los resultados, considerando que cada punto muestral en el experimento estará caracterizado porque la bola seleccionada tiene el atributo blanca en u ocasiones y par en v ocasiones, así

$$\underbrace{BP, BP, \ldots, BP}_{a} \quad \underbrace{BI, BI, \ldots, BI}_{u-a} \quad \underbrace{NP, NP, \ldots, NP}_{v-a} \quad \underbrace{NI, NI, \ldots, NI}_{n-u-v+a}$$

este punto muestral representa la intersección de n pruebas independientes agrupadas en cuatro grupos distintos y por lo tanto la probabilidad de la ocurrencia de este punto muestral es

$$\underbrace{p_{11}\, p_{11} \cdots p_{11}}_{a \text{ términos}} \quad \underbrace{p_{10}\, p_{10} \cdots p_{10}}_{u-a \text{ términos}} \quad \underbrace{p_{01}\, p_{01} \cdots p_{01}}_{v-a \text{ términos}} \quad \underbrace{p_{00}\, p_{00} \cdots p_{00}}_{n-u-v+a \text{ términos}} = p_{11}^{a}\, p_{10}^{u-a}\, p_{01}^{v-a}\, p_{00}^{n-u-v+a}$$

donde $p_{11}^{a}\, p_{10}^{u-a}$ significa que en a ocasiones se seleccionó una bola blanca y par y que en $u-a$ ocasiones se seleccionó una bola blanca e impar, es decir que en u ocasiones se obtuvo una bola blanca. Y $p_{11}^{a}\, p_{01}^{v-a}$ significa que en a ocasiones se seleccionó una bola blanca y par y que en $v-a$ ocasiones se seleccionó una bola negra y par, es decir que en v ocasiones se obtuvo una bola par.

Cada uno de los puntos muestrales puede representarse por un reordenamiento como el mostrado anteriormente y tendrán la probabilidad indicada. Por otro lado, el número N de formas en que puede suceder el evento es el número de formas en que se pueden dividir las n pruebas en cuatro grupos de tamaños $a, u-a, v-a$ y $n-u-v+a$, entonces

$$N = \frac{n!}{a!(u-a)!(v-a)!(n-u-v+a)!} = \binom{n}{a, u-a, v-a, n-u-v+a}$$

donde cada uno de los N puntos ocurre con probabilidad

$$p_{11}^{a}\, p_{10}^{u-a}\, p_{01}^{v-a}\, p_{00}^{n-u-v+a},$$

cumpliéndose así (2.5). Así pues, (U, V) tiene la distribución binomial bivariada o también se dirá simplemente que tiene distribución (2.5). ◄

Más generalmente, la distribución binomial bivariada puede ser asociada a un experimento aleatorio que consiste en tomar una muestra de n individuos con reemplazo provenientes de alguna población cuyos elementos son similares y en el que cada prueba indica la presencia o ausencia de dos atributos, digamos A y B. En la Tabla 2.3 se muestra la tabla de contingencia de tamaño 2×2 donde cada celda representa una probabilidad conjunta; por ejemplo, p_{01} representa la probabilidad de que el individuo seleccionado de la muestra presente el atributo B y no el atributo A.

Consideramos también las siguientes cantidades:

$$U = \text{número de individuos con el atributo } A.$$

$$V = \text{número de individuos con el atributo } B.$$

Se afirma que las celdas en la Tabla 2.3 tienen una distribución multinomial y las sumas resultantes de los elementos de la primer fila y la segunda columna tienen una distribución binomial bivariada. Así pues, la distribución binomial bivariada puede ser obtenida a partir de una distribución multinomial.

Vamos a probar la primera afirmación. Definimos las siguientes variables, Z_{00}, Z_{01}, Z_{10} y Z_{11} como el número de individuos cuyos atributos son $A^c \cap B^c$, $A^c \cap B$, $A \cap B^c$ y $A \cap B$, con probabilidades p_{00}, p_{01}, p_{10} y p_{11}, respectivamente. Entonces, en efecto, $(Z_{00}, Z_{01}, Z_{10}, Z_{11})$ tienen una distribución multinomial de cuatro celdas, es decir,

$$P[Z_{00} = r, Z_{01} = s, Z_{10} = t, Z_{11} = w] = \frac{n!}{r!\, s!\, t!\, w!}\, p_{00}^{r}\, p_{01}^{s}\, p_{10}^{t}\, p_{11}^{w}.$$

Para probar la segunda afirmación pensemos en el ejemplo de la urna con M bolas descrita en el Ejemplo 2.10, donde A representa el atributo "bola blanca" y B el atributo "bola par", en otras palabras A es el conjunto cuyos elementos son bolas blancas y B es el conjunto cuyos elementos son bolas pares. El experimento es tal que se puede obtener alguno de los cuatro resultados siguientes, bola negra e impar;

	A^c	A	
B	p_{01}	p_{11}	p_{+1}
B^c	p_{00}	p_{10}	p_{+0}
	p_{0+}	p_{1+}	

Tabla 2.3: Tabla de contingencia 2×2

bola negra y par; bola blanca e impar; bola blanca y par, es decir, $A^c \cap B^c$; $A^c \cap B$; $A \cap B^c$; $A \cap B$, respectivamente. Tomando la definición del último párrafo de las variables Z_{00}, Z_{01}, Z_{10} y Z_{11}, tenemos que $U = Z_{10} + Z_{11}$ es el número de bolas seleccionadas que presentaron el atributo "blanca" y $V = Z_{01} + Z_{11}$ es el número de bolas seleccionadas que presentaron el atributo "par". Por el Ejemplo 2.10, (U, V) tiene la distribución (2.5). En esta forma la distribución binomial bivariada puede ser obtenida a partir de una distribución multinomial.

Por otro lado, la distribución multinomial es un caso muy especial de una distribución binomial en cuatro dimensiones en la que todos los parámetros son cero excepto $p_{1000}, p_{0100}, p_{0010}$ y p_{0001}. Esta conexión íntima entre la distribución multinomial y binomial multivariada persiste en dimensiones superiores.

2.2.2. Una generalización de la distribución binomial bivariada

Para que un vector aleatorio (U, V) tenga la distribución binomial bivariada dada por (2.5) es necesario que ambas variables aleatorias U y V tomen valores sobre un mismo conjunto $\{0, 1, \ldots, n\}$. Sin embargo, existen situaciones prácticas donde se desearía disponer de alguna versión de esta distribución donde U tome valores en $\{0, 1, \ldots, n\}$ y V tome valores en $\{0, 1, \ldots, m\}$ con n y m posiblemente distintos. A continuación se propone una versión de la distribución binomial bivariada con estas características.

Teorema y Definición 2.11 *Suponga que* (U^*, V^*) *tiene la distribución binomial bivariada* (2.5) *para una muestra de tamaño* k*, es decir, con* k *en lugar de* n*, que* S *tenga distribución binomial*$(n-k, p_{1+})$*, que* T *tenga distribución binomial*$(m-k, p_{+1})$ *y que* (U^*, V^*)*,* S *y* T *sean independientes. Defina el vector aleatorio* (U, V) *como*

$$U = U^* + S \qquad \text{y} \qquad V = V^* + T. \tag{2.6}$$

Entonces (U, V) *tiene distribución conjunta dada por*

$$\begin{aligned} P[U = u, V = v] &= \sum_{i,j} P[U^* = u - i, S = i, V^* = v - j, T = j] \\ &= \sum_{i,j} \sum_{a} \binom{k}{a, u-i-a, v-j-a, k-u-v+i+j+a} \\ &\qquad \times p_{11}^{a}\, p_{10}^{u-i-a}\, p_{01}^{v-j-a}\, p_{00}^{k-u-v+i+j+a} \\ &\qquad\qquad \times \binom{n-k}{i} p_{1+}^{i} p_{0+}^{n-k-i} \binom{m-k}{j} p_{+1}^{j} p_{+0}^{m-k-j}. \end{aligned} \tag{2.7}$$

A esta distribución de (U, V) *se le llama* **distribución binomial bivariada generalizada***, abreviada BBG.*

Demostración. Para demostrar (2.7) vamos a requerir de la función generadora de momentos factoriales, abreviada *f.g.m.f.*, para la distribución binomial bivariada (2.5) correspondiente a una muestra de tamaño k.

Por definición sabemos que la *f.g.m.f.* para alguna v.a. X es $m_X(t) = E\left[t^X\right]$ y la *f.g.m.f.* conjunta para una pareja de v.as. (X, Y) es $m_{X,Y}(s, t) = E\left[s^X t^Y\right]$. Entonces, si (X, Y) tiene distribución Bernoulli bivariada, su *f.g.m.f.* debe ser

$$m_{X,Y}(s, t) = E\left[s^X t^Y\right] = \sum_x \sum_y s^x t^y p(x, y) = p(0,0) + p(1,0)s + p(0,1)t + p(1,1)st,$$

es decir,

$$m_{X,Y}(s,t) = p_{00} + p_{10}s + p_{01}t + p_{11}st. \tag{2.8}$$

Sean ahora $(X_1, Y_1), \ldots, (X_k, Y_k)$ vectores aleatorios independientes de Bernoulli bivariados y sean

$$U^* = \sum_{i=1}^{k} X_i \qquad \text{y} \qquad V^* = \sum_{i=1}^{k} Y_i.$$

Entonces

$$\begin{aligned} m_{U^*,V^*}(s,t) = E\left[s^{U^*} t^{V^*}\right] &= E\left[s^{\sum_{i=1}^k X_i}\, t^{\sum_{i=1}^k Y_i}\right] \\ &= E\left[s^{X_1}t^{Y_1}s^{X_2}t^{Y_2}\cdots s^{X_k}t^{Y_k}\right] = E\left[s^{X_1}t^{Y_1}\right]E\left[s^{X_2}t^{Y_2}\right]\cdots E\left[s^{X_k}t^{Y_k}\right], \end{aligned}$$

donde la última igualdad está justificada por independencia de las parejas (X_i, Y_i). Finalmente, como tales parejas son idénticamente distribuidas, se tiene que

$$E\left[s^{X_i}t^{Y_i}\right] = E\left[s^{X_j}t^{Y_j}\right], \quad \forall i, j = 1, 2, \ldots, k,$$

o sea, todas las parejas tienen la misma $f.g.m.f.$. Así es que,

$$E\left[s^{U^*}t^{V^*}\right] = E\left[s^{X_i}t^{Y_i}\right]^k$$

donde $E\left[s^{X_i}t^{Y_i}\right]$ está dada por (2.8) para cualquier $i = 1, \ldots, k$. Por lo tanto,

$$m_{U^*,V^*}(s,t) = (p_{00} + p_{10}s + p_{01}t + p_{11}st)^k \tag{2.9}$$

es la $f.g.m.f.$ para una pareja (U^*, V^*) con distribución Binomial bivariada (2.5) para una muestra de tamaño k.

Consideremos ahora el siguiente esquema de muestreo. Son realizados tres muestreos independientes con reemplazo provenientes de una población en la que cada individuo puede ser clasificado en uno de dos atributos A y B. En el primer muestreo k individuos son clasificados acorde a ambos atributos, sean U^* el número de veces que se observó el atributo A y V^* el número de veces que ocurrió el atributo B. En el segundo muestreo de tamaño $n-k$ sólo se estudia el atributo A, sea S el número de veces que se observó A en la muestra. En el tercer muestreo de tamaño $m-k$ sólo B es estudiado y ocurre en T ocasiones. Entonces el número total de veces que ocurrió A es $U = U^* + S$, mientras que el número total de veces que ocurrió B es $V = V^* + T$. Nos preguntamos por la distribución de (U, V). Supongamos, como se ilustra en la Tabla 2.3, que las proporciones de A, B y $A \cap B$ en la población son p_{1+}, p_{+1} y p_{11}, respectivamente. Entonces tenemos, efectivamente, que

$$\begin{aligned} S &\sim Binomial(n-k, p_{1+}), \\ T &\sim Binomial(m-k, p_{+1}), \end{aligned}$$

mientras que (U^*, V^*) tiene la distribución Binomial bivariada (2.5), con (U^*, V^*), S y T independientes debido al esquema de muestreo descrito.

Si $X \sim Binomial(n,p)$, entonces la $f.g.m.f.$ para X es, tomando $q = 1-p$,

$$m_X(t) = E\left[t^X\right] = \sum_{x=0}^{n} \left(\begin{array}{c} n \\ x \end{array} \right) t^x p^x q^{n-x} = \sum_{x=0}^{n} \left(\begin{array}{c} n \\ x \end{array} \right) (tp)^x q^{n-x} = (tp+q)^n.$$

Por lo que las $f.g.m.f.$ para S y T son

$$m_S(s) = (p_{0+} + p_{1+}s)^{n-k} \qquad \text{y} \qquad m_T(t) = (p_{+0} + p_{+1}t)^{m-k}. \tag{2.10}$$

Buscamos entonces la $f.g.m.f.$ para (U,V) con la cual obtendremos la distribución buscada. Primero vamos a dar de forma explícita las variables que están involucradas en el esquema de muestreo descrito. Definimos las variables aleatorias $X_1,\ldots,X_k$, $Y_1,\ldots,Y_k$, $X_1',\ldots,X_{n-k}'$ y $Y_1',\ldots,Y_{m-k}'$ como sigue:

$$X_i = \begin{cases} 1 & \text{si el individuo seleccionado presenta el atributo} A, \\ 0 & \text{si el individuo seleccionado presenta el atributo} A^c, \end{cases} \qquad \forall\, i \in \{1,\ldots,k\},$$

$$Y_i = \begin{cases} 1 & \text{si el individuo seleccionado presenta el atributo} B, \\ 0 & \text{si el individuo seleccionado presenta el atributo} B^c, \end{cases} \qquad \forall\, i \in \{1,\ldots,k\},$$

$$X_i' = \begin{cases} 1 & \text{si el individuo seleccionado presenta el atributo} A, \\ 0 & \text{si el individuo seleccionado presenta el atributo} A^c, \end{cases} \qquad \forall\, i \in \{1,\ldots,n-k\},$$

$$Y_i' = \begin{cases} 1 & \text{si el individuo seleccionado presenta el atributo} B, \\ 0 & \text{si el individuo seleccionado presenta el atributo} B^c, \end{cases} \qquad \forall\, i \in \{1,\ldots,m-k\}.$$

Las variables aleatorias X_i y Y_i están ligadas en el sentido de que corresponden a la observación (X_i,Y_i) del i-ésimo individuo, mientras que X_i' y Y_i' no están ligadas. Entonces, podemos decir que, de acuerdo al esquema de muestreo, $(X_1,Y_1),\ldots,(X_k,Y_k)$, $X_1',\ldots,X_{n-k}'$, $Y_1',\ldots,Y_{m-k}'$ son independientes, (X_i,Y_i) tiene la distribución Bernoulli bivariada para $i=1,\ldots,k$, $X_i' \sim Bernoulli(p_{1+})$ para $i=1,\ldots,n-k$ y $Y_i' \sim Bernoulli(p_{+1})$ para $i=1,\ldots,m-k$. Por lo tanto

$$U = U^* + S = \sum_{i=1}^{k} X_i + \sum_{i=1}^{n-k} X_i', \quad V = V^* + T = \sum_{i=1}^{k} Y_i + \sum_{i=1}^{m-k} Y_i',$$

luego

$$\begin{aligned} m_{U,V}(s,t) =& E\left[s^U\, t^V\right] \\ =& E\left[s^{\sum_{i=1}^{k} X_i + \sum_{i=1}^{n-k} X_i'}\, t^{\sum_{i=1}^{k} Y_i + \sum_{i=1}^{m-k} Y_i'}\right] \\ =& E\left[s^{\sum_{i=1}^{k} X_i}\, s^{\sum_{i=1}^{n-k} X_i'}\, t^{\sum_{i=1}^{k} Y_i}\, t^{\sum_{i=1}^{m-k} Y_i'}\right] \\ =& E\left[s^{\sum_{i=1}^{k} X_i}\, t^{\sum_{i=1}^{k} Y_i}\, s^{\sum_{i=1}^{n-k} X_i'}\, t^{\sum_{i=1}^{m-k} Y_i'}\right] \\ =& E\left[s^{U^*} t^{V^*}\, s^{\sum_{i=1}^{n-k} X_i'}\, t^{\sum_{i=1}^{m-k} Y_i'}\right] \\ =& E\left[s^{U^*} t^{V^*}\right]\, E\left[s^{\sum_{i=1}^{n-k} X_i'}\right]\, E\left[t^{\sum_{i=1}^{m-k} Y_i'}\right] \\ =& m_{U^*,V^*}(s,t)\, m_S(s)\, m_T(t), \end{aligned}$$

donde la penúltima igualdad está justificada por independencia. Requerimos de (2.9) y (2.10) para obtener así que

$$m_{U,V}(s,t) = (p_{00} + p_{10}s + p_{01}t + p_{11}st)^k\, (p_{0+} + p_{1+}s)^{n-k}\, (p_{+0} + p_{+1}t)^{m-k} \tag{2.11}$$

es la $f.g.m.f.$ para la distribución BBG (2.7).

Podemos deducir de (2.11) la distribución de (U, V) expandiendo los tres factores que aparecen y determinando los coeficientes de tales desarrollos. Se tiene,

$$(p_{0+} + p_{1+}s)^{n-k} = \sum_{i=0}^{n-k} a_i\, s^i,$$
$$(p_{+0} + p_{+1}t)^{m-k} = \sum_{j=0}^{m-k} b_j\, t^j,$$
$$(p_{00} + p_{10}s + p_{01}t + p_{11}st)^k = \sum_{u^*}\sum_{v^*} c_{u^*,v^*}\, s^{u^*} t^{v^*},$$

donde

$$a_i = \binom{n-k}{i} p_{1+}^{i}\, p_{0+}^{n-k-i}, \tag{2.12}$$

$$b_j = \binom{m-k}{j} p_{+1}^{j}\, p_{+0}^{m-k-j}, \tag{2.13}$$

$$c_{u^*,v^*} = \sum_a \binom{k}{a, u^*-a, v^*-a, k-u^*-v^*+a} p_{11}^{a}\, p_{10}^{u^*-a}\, p_{01}^{v^*-a}\, p_{00}^{k-u^*-v^*+a}. \tag{2.14}$$

Por otra parte, por definición tenemos que

$$m_{U,V}(s,t) = \sum_u \sum_v s^u\, t^v P[U = u, V = v]. \tag{2.15}$$

Sustituyendo en (2.11) los desarrollos dados, se tiene

$$\begin{aligned} m_{U,V}(s,t) &= \sum_{u^*}\sum_{v^*} c_{u^*,v^*}\, s^{u^*} t^{v^*} \sum_{i=0}^{n-k} a_i\, s^i \sum_{j=0}^{m-k} b_j\, t^j \\ &= \sum_{u^*}\sum_{v^*}\sum_{i=0}^{n-k}\sum_{j=0}^{m-k} c_{u^*,v^*}\, a_i\, b_j\, s^{u^*} t^{v^*} s^i t^j \\ &= \sum_{u^*}\sum_{v^*}\sum_{i=0}^{n-k}\sum_{j=0}^{m-k} c_{u^*,v^*}\, a_i\, b_j\, s^{u^*+i}\, t^{v^*+j}. \end{aligned}$$

Reemplazando $u = u^* + i$ y $v = v^* + j$ llegamos a que

$$m_{U,V}(s,t) = \sum_u \sum_v \sum_{i=0}^{n-k}\sum_{j=0}^{m-k} c_{u-i,v-j}\, a_i\, b_j\, s^u\, t^v. \tag{2.16}$$

Recuerde que, en el caso univariado, si dos series de potencias centradas en c, $\sum a_n(x-c)^n$ y $\sum b_n(x-c)^n$, representan a una misma función $f(x)$ entonces $a_n = b_n$ para cada n. El resultado es similar en el caso bivariado, de modo que, comparando (2.15) y (2.16) obtenemos la distribución buscada

$$P[U = u, V = v] = \sum_{i=0}^{n-k}\sum_{j=0}^{m-k} c_{u-i,v-j}\, a_i\, b_j, \tag{2.17}$$

donde, por (2.14),

$$c_{u-i,v-j} = \sum_a \binom{k}{a, u-i-a, v-j-a, k-u-v+i+j+a} p_{11}^a p_{10}^{u-i-a} p_{01}^{v-j-a} p_{00}^{k-u-v+i+j+a} \tag{2.18}$$

y entonces sustituyendo (2.12), (2.13) y (2.18) en (2.17) se concluye (2.7). ∎

Claramente la distribución binomial bivariada (2.5) es un caso particular de la distribución BBG (2.7), pues al suponer $k = m = n$, es decir, $S = 0$ y $T = 0$, se obtiene por (2.6) que $U = U^*$ y $V = V^*$. Por lo tanto (U, V) tiene distribución (2.5).

Es posible obtener el coeficiente de correlación de (U, V) a partir de la $f.g.m.f.$ dada por (2.11) como a continuación se indica.

Proposición 2.12 *Se tiene*

$$Corr(U, V) = \left(\frac{k}{\sqrt{mn}}\right)\rho, \tag{2.19}$$

donde ρ es el coeficiente de correlación de la distribución de Bernoulli bivariada subyacente dado por (2.3).

Demostración. Vamos a recurrir a la definición de correlación donde se requiere de $E[UV]$, $E[U]$, $E[V]$, $Var[U]$ y $Var[V]$.

Para encontrar $E[UV]$ vamos a usar (2.11), que denotaremos simplemente como $m(s,t)$, de la siguiente forma. Sabemos que (Teorema 1.51)

$$E[UV] = \left.\frac{\partial^2 m(s,t)}{\partial s\, \partial t}\right|_{s=1,t=1}.$$

Derivamos (2.11) primero parcialmente con respecto de t para obtener

$$\begin{aligned}\frac{\partial m(s,t)}{\partial t} &= k\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k-1}\left(p_{01} + p_{11}s\right)\left(p_{0+} + p_{1+}s\right)^{n-k}\left(p_{+0} + p_{+1}t\right)^{m-k} \\ &\quad + (m-k)\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k}\left(p_{0+} + p_{1+}s\right)^{n-k}\left(p_{+0} + p_{+1}t\right)^{m-k-1} p_{+1}.\end{aligned}$$

Volvemos a derivar parcialmente ahora con respecto de s para obtener

$$\begin{aligned}\frac{\partial^2 m(s,t)}{\partial s\,\partial t} &= \Big\{k(k-1)\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k-2}\left(p_{10} + p_{11}t\right)\left(p_{01} + p_{11}s\right) \\ &\quad + k\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k-1} p_{11}\Big\} \times \left(p_{0+} + p_{1+}s\right)^{n-k}\left(p_{+0} + p_{+1}t\right)^{m-k} \\ &\quad + (n-k)k\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k-1}\left(p_{01} + p_{11}s\right)\left(p_{0+} + p_{1+}s\right)^{n-k-1}\left(p_{+0} + p_{+1}t\right)^{m-k} p_{1+} \\ &\quad + (m-k)k\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k-1}\left(p_{10} + p_{11}t\right)\left(p_{0+} + p_{1+}s\right)^{n-k}\left(p_{+0} + p_{+1}t\right)^{m-k-1} p_{+1} \\ &\quad + (m-k)(n-k)\left(p_{00} + p_{10}s + p_{01}t + p_{11}st\right)^{k}\left(p_{0+} + p_{1+}s\right)^{n-k-1}\left(p_{+0} + p_{+1}t\right)^{m-k-1} p_{1+}\, p_{+1}\end{aligned}$$

y finalmente evaluamos en $(s = 1, t = 1)$ para obtener el $E[UV]$. Considerando que $p_{11} + p_{01} + p_{10} + p_{00} = 1$

y (2.1) obtenemos

$$\begin{aligned}
&\left.\frac{\partial^2 m(s,t)}{\partial s\,\partial t}\right|_{s=1,t=1} \\
&\quad= k(k-1)p_{1+}\,p_{+1} + k\,p_{11} + (n-k)k\,p_{1+}\,p_{+1} + (m-k)k\,p_{1+}\,p_{+1} + (m-k)(n-k)p_{1+}\,p_{+1} \\
&\quad= k\,p_{11} + k\left((n-k)+(k-1)\right)p_{1+}\,p_{+1} + (m-k)\left(k+(n-k)\right)p_{1+}\,p_{+1} \\
&\quad= k\,p_{11} + k(n-1)p_{1+}\,p_{+1} + (m-k)n\,p_{1+}\,p_{+1} \\
&\quad= k\,p_{11} + kn\,p_{1+}\,p_{+1} - k\,p_{1+}\,p_{+1} + mn\,p_{1+}\,p_{+1} - kn\,p_{1+}\,p_{+1} \\
&\quad= k\,p_{11} + (nm-k)p_{1+}\,p_{+1},
\end{aligned}$$

por lo tanto

$$E[UV] = k\,p_{11} + (nm-k)p_{1+}\,p_{+1}. \tag{2.20}$$

Ahora bien, recordando (2.6), las distribuciones de S, T y de las variables $X_1,\ldots,X_k$, $Y_1,\ldots,Y_k$, llegamos a que

$$E[U] = E[U^* + S] = E[U^*] + E[S] = \sum_{i=1}^{k} E[X_i] + (n-k)p_{1+} = k\,p_{1+} + (n-k)p_{1+} = n\,p_{1+} \tag{2.21}$$

$$E[V] = E[V^* + T] = E[V^*] + E[T] = \sum_{i=1}^{k} E[Y_i] + (m-k)p_{+1} = k\,p_{+1} + (m-k)p_{+1} = m\,p_{+1}. \tag{2.22}$$

Sustituyendo (2.20), (2.21) y (2.22) en la definición de covarianza $Cov(U,V) = E[UV] - E[U]E[V]$, obtenemos

$$Cov(U,V) = k(p_{11} - p_{1+}p_{+1}). \tag{2.23}$$

Finalmente, de manera similar a como se calcularon los valores esperados, obtenemos $Var[U]$ y $Var[V]$ como sigue

$$\begin{aligned}
Var[U] &= Var[U^* + S] \\
&= Var[U^*] + Var[S] + 2Cov(U^*, S) \\
&= \sum_{i=1}^{k} Var[X_i] + (n-k)p_{1+}\,p_{0+} \\
&= k\,p_{1+}\,p_{0+} + (n-k)p_{1+}\,p_{0+} \\
&= n\,p_{1+}\,p_{0+}
\end{aligned} \tag{2.24}$$

y

$$\begin{aligned}
Var[V] &= Var[V^* + T] \\
&= Var[V^*] + Var[T] + 2Cov(V^*, T) \\
&= \sum_{i=1}^{k} Var[Y_i] + (m-k)p_{+1}\,p_{+0} \\
&= k\,p_{+1}\,p_{+0} + (m-k)p_{+1}\,p_{+0} \\
&= m\,p_{+1}\,p_{+0}
\end{aligned} \tag{2.25}$$

En las identidades (2.24) y (2.25) se tiene que $Cov(U^*, S) = 0$ y $Cov(V^*, T) = 0$, por independencia.

Utilizando (2.23), (2.24), (2.25) y recordando (2.3), llegamos a (2.19). En efecto,

$$Corr(U,V) = \frac{Cov(U,V)}{\sqrt{Var[U]\,Var[V]}} = \frac{k(p_{11} - p_{1+}p_{+1})}{\sqrt{mn\,p_{1+}\,p_{0+}\,p_{+1}\,p_{+0}}} = \left(\frac{k}{\sqrt{mn}}\right)\left(\frac{p_{11} - p_{1+}p_{+1}}{\sqrt{p_{1+}\,p_{0+}\,p_{+1}\,p_{+0}}}\right).$$

Por lo tanto,

$$Corr(U,V) = \left(\frac{k}{\sqrt{mn}}\right)\rho,$$

completando la prueba. ∎

Como ya se mencionó anteriormente, la distribución binomial bivariada (2.5) es un caso particular de (2.7) cuando $k = m = n$. Al sustituir $k = m = n$ en (2.19) se recobra de inmediato la identidad $Corr(U,V) = \rho$.

2.2.3. Independencia

Comprobaremos para la distribución binomial bivariada que correlación cero implica independencia de manera análoga a la distribución de Bernoulli bivariada. Esto se debe a que si $Corr(U,V) = 0$, entonces $k = 0$ o $p_{11} = p_{1+}\,p_{+1}$ y de cualquiera de estas dos condiciones se implicará la independencia de U y V.

Proposición 2.13 *Si* (U,V) *tiene la distribución BBG dada por* (2.7) *y* $Corr(U,V) = 0$ *entonces* U *y* V *son independientes.*

Demostración. Supongamos que $Corr(U,V) = 0$. Entonces, por (2.19) y (2.2), tenemos alguno de los dos casos siguientes.

Caso $k = 0$. Por (2.6), necesariamente $U = S$ y $V = T$, pero como ya sabemos que S y T son independientes se tiene el resultado deseado.

Caso $p_{11} = p_{1+}\,p_{+1}$. Sea (X,Y) un vector aleatorio con distribución de Bernoulli bivariada con esos parámetros. Se sabe de la Proposición 2.4 que entonces X y Y son independientes. Recordemos que, para $i = 1,\ldots,k$, las parejas (X_i, Y_i) son independientes e idénticamente distribuidas con distribución Bernoulli bivariada común con esos mismos parámetros. Por lo anterior, X_i y Y_i son independientes para todo i. Por un resultado conocido de independencia, la variable aleatoria $\sum_{i=1}^{k} X_i$ debe ser independiente a $\sum_{i=1}^{k} Y_i$, es decir, U^* y V^* son independientes. Ya que U^*, V^*, S y T son independientes, por ese mismo resultado se concluye que las variables aleatorias

$$U = U^* + S \qquad \text{y} \qquad V = V^* + T$$

son también independientes. ∎

2.2.4. Valor esperado condicional

En esta sección vamos a calcular los valores esperados condicionales cuando el vector aleatorio (U,V) tiene tanto la distribución binomial bivariada dada por (2.5) como la distribución binomial bivariada generalizada dada por (2.7).

Teorema 2.14

i. *Si* (U, V) *es un vector aleatorio con distribución binomial bivariada dada por* (2.5), *entonces*

$$E[U \mid V = v] = n\,\frac{p_{10}}{p_{+0}} + v\left(\frac{\tau}{p_{+1}p_{+0}}\right), \tag{2.26}$$

$$E[V \mid U = u] = n\frac{p_{01}}{p_{0+}} + u\left(\frac{\tau}{p_{1+}p_{0+}}\right). \tag{2.27}$$

ii. *Si* (U, V) *es un vector aleatorio con distribución binomial bivariada generalizada* (2.7), *entonces*

$$E[U \mid V = v] = \left(np_{1+} - \frac{k\tau}{p_{+0}} + \frac{v\tau}{p_{+1}p_{+0}}\right) - \left(\frac{\tau}{p_{+1}p_{+0}}\right)\frac{\displaystyle\sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} t\binom{m-k}{t}p_{+1}^{t}p_{+0}^{m-k-t}}{\displaystyle\sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} \binom{m-k}{t}p_{+1}^{t}p_{+0}^{m-k-t}}, \tag{2.28}$$

$$E[V \mid U = u] = \left(mp_{+1} - \frac{k\tau}{p_{0+}} + \frac{u\tau}{p_{1+}\,p_{0+}}\right) - \left(\frac{\tau}{p_{1+}\,p_{0+}}\right)\frac{\displaystyle\sum_{s=\text{máx}(0,u-k)}^{\text{mín}(n-k,u)} s\binom{n-k}{s}p_{1+}^{s}\,p_{0+}^{n-k-s}}{\displaystyle\sum_{s=\text{máx}(0,u-k)}^{\text{mín}(n-k,u)} \binom{n-k}{s}p_{1+}^{s}p_{0+}^{n-k-s}}. \tag{2.29}$$

Demostración. **(i)** Suponga que (U, V) tiene la distribución binomial bivariada (2.5). Se calculará primero la esperanza condicional de U dado que $Y_1 = y_1, \ldots, Y_n = y_n$ con $y_1, \ldots, y_n$ iguales a 0 o 1 y $\sum_{i=1}^{n} y_i = v$. Se tiene

$$\begin{aligned} E[U \mid Y_1 = y_1, \ldots, Y_n = y_n] &= E\left[\sum_{i=1}^{n} X_i \middle| Y_1 = y_1, \ldots, Y_n = y_n\right] \\ &= \sum_{i=1}^{n} E[X_i \,|\, Y_1 = y_1, \ldots, Y_n = y_n] \\ &= \sum_{i=1}^{n} E[X_i \,|\, Y_i = y_i], \end{aligned} \tag{2.30}$$

donde la última igualdad está justificada por la independencia de los vectores aleatorios $(X_1, Y_1), \ldots, (X_n, Y_n)$, es decir, si tomamos, por ejemplo, $i = 1$, entonces

$$E[X_1 \,|\, Y_1 = y_1, \ldots, Y_n = y_n] = E[X_1 \,|\, Y_1 = y_1]$$

porque (X_1, Y_1) es independiente a $(X_2, Y_2), \ldots, (X_n, Y_n)$ en particular X_1 es independiente a Y_i siempre que $i \neq 1$. Luego tenemos que

$$\sum_{i=1}^{n} E[X_i \mid Y_i = y_i] = (n - v)E[X_1 \mid Y_1 = 0] + vE[X_1 \mid Y_1 = 1], \tag{2.31}$$

ya que

$$E[X_i \mid Y_i = y_i] = E[X_j \mid Y_j = y_j], \quad i, j \in \{1, \ldots, n\}$$

y para cualquier $y_i, y_j \in \{0, 1\}$, por ser los vectores aleatorios $(X_1, Y_1), \ldots, (X_n, Y_n)$ idénticamente distribuidos.

Vemos ahora, recordando las probabilidades marginales dadas en la Sección 2.1, que

$$E[X_1 \mid Y_1 = 0] = \sum_{x_1} x_1 P[X_1 = x_1 \mid Y_1 = 0] = \sum_{x_1} x_1 \frac{P[X_1 = x_1, Y_1 = 0]}{P[Y_1 = 0]} = \frac{P[X_1 = 1, Y_1 = 0]}{p_{+0}} = \frac{p_{10}}{p_{+0}}$$

y

$$E[X_1 \mid Y_1 = 1] = \sum_{x_1} x_1 P[X_1 = x_1 \mid Y_1 = 1] = \sum_{x_1} x_1 \frac{P[X_1 = x_1, Y_1 = 1]}{P[Y_1 = 1]} = \frac{P[X_1 = 1, Y_1 = 1]}{p_{+1}} = \frac{p_{11}}{p_{+1}}.$$

Sustituyendo en (2.31) obtenemos

$$\begin{aligned} \sum_{i=1}^{n} E[X_i \mid Y_i = y_i] &= (n - v)\frac{p_{10}}{p_{+0}} + v\,\frac{p_{11}}{p_{+1}} \\ &= n\,\frac{p_{10}}{p_{+0}} + v\left(\frac{p_{11}}{p_{+1}} - \frac{p_{10}}{p_{+0}}\right) \\ &= n\,\frac{p_{10}}{p_{+0}} + v\left(\frac{p_{11}\,p_{+0} - p_{10}\,p_{+1}}{p_{+1}\,p_{+0}}\right) \\ &= n\,\frac{p_{10}}{p_{+0}} + v\left(\frac{p_{11} - p_{1+}\,p_{+1}}{p_{+1}\,p_{+0}}\right) \\ &= n\,\frac{p_{10}}{p_{+0}} + v\left(\frac{\tau}{p_{+1}\,p_{+0}}\right), \end{aligned} \tag{2.32}$$

donde en la última igualdad recurrimos a (2.2). Finalmente, sustituyendo (2.32) en (2.30) obtenemos

$$E[U \mid Y_1 = y_1, \ldots, Y_n = y_n] = n\frac{p_{10}}{p_{+0}} + v\left(\frac{\tau}{p_{+1}p_{+0}}\right)$$

Se afirma ahora que como $E[U \mid Y_1 = y_1, \ldots, Y_n = y_n]$ es constante con respecto a $y_1, \ldots, y_n$, entonces

$$E[U \mid V = v] = n\frac{p_{10}}{p_{+0}} + v\left(\frac{\tau}{p_{+1}p_{+0}}\right).$$

En efecto, se considera el siguiente evento

$$J = \left\{ (y_1, \ldots, y_n) \;\middle|\; \sum_{i=1}^{n} y_i = v \right\}$$

con el que es posible caracterizar por componentes a la variable aleatoria V. Recordando la definición de valor esperado condicional, se tiene

$$E[U \mid V = v] = E[U \mid J] = \sum_{u} u \frac{P[U = u, I_J(y_1, \ldots, y_n) = 1]}{P[J]},$$

donde

$$I_J(y_1, \ldots, y_n) = \begin{cases} 1 & \text{si } (y_1, \ldots, y_n) \in J, \\ 0 & \text{en otro caso} \end{cases}$$

es la función indicadora del evento J. Dado que $V = \sum_{i=1}^{n} Y_i$, se debe tener $P[J] = P[V = v]$. También se tiene que

$$P[U = u, I_J(y_1, \ldots, y_n) = 1] = P[U = u, Y_1 = y_1, \ldots, Y_n = y_n]\, I_J(y_1, \ldots, y_n)$$

y así es como

$$E[U \mid V = v] = \frac{\sum_u uP[U = u, Y_1 = y_1, \ldots, Y_n = y_n]\, I_J(y_1, \ldots, y_n)}{P[V = v]}.$$

Obsérvese que el numerador de esta última expresión sólo es significativo (no se anula) toda vez que $(y_1, \ldots, y_n) \in J$, es decir, cada vez que la función indicadora toma el valor 1, esto es posible describirlo de la siguiente forma

$$\begin{aligned}\sum_u & uP[U = u, Y_1 = y_1, \ldots, Y_n = y_n]\, I_J(y_1, \ldots, y_n) \\ &= \sum_{(y_1,\ldots,y_n)\in J} \sum_u uP[U = u, Y_1 = y_1, \ldots, Y_n = y_n],\end{aligned}$$

por lo que se obtiene

$$E[U \mid V = v] = \frac{\sum_{(y_1,\ldots,y_n)\in J} \sum_u uP[U = u, Y_1 = y_1, \ldots, Y_n = y_n]}{P[V = v]}.$$

Por otro lado,

$$E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n] = \sum_u u\frac{P[U = u, Y_1 = y_1, \ldots, Y_n = y_n]}{P[Y_1 = y_1, \ldots, Y_n = y_n]},$$

de donde

$$\sum_u uP[U = u, Y_1 = y_1, \ldots, Y_n = y_n] = E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n]P[Y_1 = y_1, \ldots, Y_n = y_n]$$

y así es que

$$E[U \mid V = v] = \frac{\sum_{(y_1,\ldots,y_n)\in J} E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n]P[Y_1 = y_1, \ldots, Y_n = y_n]}{P[V = v]},$$

donde

$$\begin{aligned}\sum_{(y_1,\ldots,y_n)\in J} & E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n]P[Y_1 = y_1, \ldots, Y_n = y_n] \\ &= E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n] \sum_{(y_1,\ldots,y_n)\in J} P[Y_1 = y_1, \ldots, Y_n = y_n] \\ &= E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n]P[Y_1 = y_1, \ldots, Y_n = y_n]\, I_J(y_1, \ldots, y_n) \\ &= E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n]P[V = v].\end{aligned}$$

Por lo tanto,

$$E[U \mid V = v] = E[U = u \mid Y_1 = y_1, \ldots, Y_n = y_n]$$

lo cual demuestra la afirmación y completa la prueba de (2.26).

Intercambiando los papeles de U y V se concluye que también se cumple

$$E[V \mid U = u] = n\frac{p_{01}}{p_{0+}} + u\left(\frac{\tau}{p_{1+}p_{0+}}\right),$$

que es (2.27).

(ii) Suponga ahora que (U,V) tiene distribución BBG dada por (2.7). Recuerde que, por (2.6), $U = U^* + S$ y $V = V^* + T$, donde (U^*, V^*) tiene distribución binomial bivariada con parámetro k, S tiene distribución binomial con parámetros $n-k$ y p_{1+}, T tiene distribución binomial con parámetros $m-k$ y p_{+1} y que (U^*, V^*), S y T son independientes. Para calcular $E[U \mid V = v]$ utilizamos (2.26) de la siguiente manera. Por una propiedad general de la esperanza condicional, se tiene

$$E[U \mid V] = E\left[E[U \mid V, T] \,\middle|\, V\right].$$

Por otras propiedades generales, se tiene entonces

$$E[U \mid V = v] = E\left[E[U \mid V, T] \,\middle|\, V = v\right] = E\left[E[U \mid V = v, T] \,\middle|\, V = v\right], \quad v = 0, \ldots, m.$$

Recuerde que, por definición, la variable aleatoria T toma los valores $0, \ldots, m-k$, V toma los valores $0, \ldots, m$, V^* toma los valores $0, \ldots, k$ y $V = V^* + T$. Entonces, dado que $V = v$ con $v \in \{0, \ldots, m\}$, la variable aleatoria condicionada $[T \mid V = v]$ sólo puede tomar valores t tales que

$$t \in \{v-k, \ldots, v\} \cap \{0, \ldots, m-k\} = \{\text{máx}(0, v-k), \ldots, \text{mín}(v, m-k)\}.$$

Así pues,

$$\begin{aligned} E\left[E[U \mid V = v, T] \,\middle|\, V = v\right] &= \sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} E[U \mid V = v, T = t] P[T = t \mid V = v] \\ &= \sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} E[U \mid V^* = v-t, T = t] P[T = t \mid V = v]. \end{aligned}$$

Sustituyendo en la penúltima ecuación se obtiene

$$\begin{aligned} E[U \mid V = v] &= \sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} E[U \mid V^* = v-t, T = t] P[T = t \mid V = v] \\ &= \sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} E[U^* + S \mid V^* = v-t, T = t] P[T = t \mid V = v], \quad v = 0, \ldots, m, \end{aligned} \tag{2.33}$$

donde la última igualdad se cumple porque $U = U^* + S$.

Ahora bien, tenemos

$$E[U^* + S \,|\, V^* = v-t, T = t] = E[U^* \,|\, V^* = v-t, T = t] + E[S \,|\, V^* = v-t, T = t]$$

donde

$$E[U^* \,|\, V^* = v-t, T = t] + E[S \,|\, V^* = v-t, T = t] = E[U^* \,|\, V^* = v-t] + E[S]$$

porque U^* es independientes de T y S es independiente de V^* y T. Ya sabemos que

$$E[S] = (n-k)p_{1+} \qquad \text{y} \qquad E[U^* \,|\, V^* = v-t] = k\,\frac{p_{10}}{p_{+0}} + (v-t)\frac{\tau}{p_{+1}\,p_{+0}},$$

por (2.26), pues (U^*, V^*) tiene la distribución Binomial bivariada (2.5). Así que

$$\begin{aligned} E[U^* \mid V^* = v - t] + E[S] &= k\,\frac{p_{10}}{p_{+0}} + (v-t)\frac{\tau}{p_{+1}\,p_{+0}} + (n-k)p_{1+} \\ &= n\,p_{1+} + k\left(\frac{p_{10}}{p_{+0}} - p_{1+}\right) + \frac{v\,\tau}{p_{+1}\,p_{+0}} - \frac{t\,\tau}{p_{+1}\,p_{+0}} \\ &= n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}} - \frac{t\,\tau}{p_{+1}\,p_{+0}}. \end{aligned}$$

Por lo tanto,

$$E[U^* + S \mid V^* = v - t, T = t] = n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}} - \frac{t\,\tau}{p_{+1}\,p_{+0}}$$

que al sustituir en (2.33) se obtiene

$$E[U \mid V = v] = \sum_{t=\max(0,v-k)}^{\min(m-k,v)} \left(n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}} - \frac{t\,\tau}{p_{+1}\,p_{+0}}\right) P[T = t \mid V = v] \tag{2.34}$$

Tenemos además que

$$P[T = t \mid V = v] = \frac{P[T = t, V = v]}{P[V = v]} = \frac{\binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}{\sum_t \binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}.$$

Sustituyendo esto último en (2.34), resulta

$$E[U \mid V = v] = \sum_t \left(n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}} - \frac{t\,\tau}{p_{+1}\,p_{+0}}\right) \frac{\binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}{\sum_t \binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}. \tag{2.35}$$

Desarrollando la suma en (2.35) se obtiene

$$\begin{aligned} &\sum_t \left(n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}} - \frac{t\,\tau}{p_{+1}\,p_{+0}}\right) \frac{\binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}{\sum_t \binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}} \\ &\quad= \sum_t \left(n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}}\right) \frac{\binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}{\sum_t \binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}} \\ &\qquad\qquad - \sum_t \left(\frac{t\,\tau}{p_{+1}\,p_{+0}}\right) \frac{\binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}{\sum_t \binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}} \\ &\quad= \left(n\,p_{1+} - \frac{k\,\tau}{p_{+0}} + \frac{v\,\tau}{p_{+1}\,p_{+0}}\right) \\ &\qquad\qquad - \left(\frac{\tau}{p_{+1}\,p_{+0}}\right) \frac{\sum_t t\binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}{\sum_t \binom{m-k}{t} p_{+1}^t\, p_{+0}^{m-k-t}}. \end{aligned}$$

Por lo tanto

$$E[U\,|\,V=v]=\left(n\,p_{1+}-\frac{k\,\tau}{p_{+0}}+\frac{v\,\tau}{p_{+1}\,p_{+0}}\right)-\left(\frac{\tau}{p_{+1}\,p_{+0}}\right)\frac{\displaystyle\sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)} t\binom{m-k}{t}p_{+1}^{t}p_{+0}^{m-k-t}}{\displaystyle\sum_{t=\text{máx}(0,v-k)}^{\text{mín}(m-k,v)}\binom{m-k}{t}p_{+1}^{t}\,p_{+0}^{m-k-t}},$$

que es (2.28).

Finalmente, intercambiando los papeles de U^* y V^* se concluye también que

$$E[V\,|\,U=u]=\left(m\,p_{+1}-\frac{k\,\tau}{p_{0+}}+\frac{u\,\tau}{p_{1+}\,p_{0+}}\right)-\left(\frac{\tau}{p_{1+}\,p_{0+}}\right)\frac{\displaystyle\sum_{s=\text{máx}(0,u-k)}^{\text{mín}(n-k,u)} s\binom{n-k}{s}p_{1+}^{s}\,p_{0+}^{n-k-s}}{\displaystyle\sum_{s=\text{máx}(0,u-k)}^{\text{mín}(n-k,u)}\binom{n-k}{s}p_{1+}^{s}\,p_{0+}^{n-k-s}},$$

que es (2.29). ∎

2.2.5. Convergencia a la distribución normal bivariada

El resultado siguiente establece que la distribución BBG para el caso particular $k=m=n$, es decir, la distribución binomial bivariada, dada por (2.5), estandarizada, converge en distribución a la normal bivariada con media $\overrightarrow{\mu}=\overrightarrow{0}$ y matriz de covarianzas $\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}$.

Teorema 2.15 (Convergencia a la distribución normal bivariada.) *Sean $(X_1,Y_1),\ldots,(X_n,Y_n)$ vectores aleatorios independientes con distribución común de Bernoulli bivariada. Si*

$$U=\sum_{i=1}^{n}X_i \qquad y \qquad V=\sum_{i=1}^{n}Y_i,$$

entonces, cuando $n\to\infty$, se tiene

$$\left(\frac{U-E[U]}{\sqrt{Var[U]}}\;,\;\frac{V-E[Y]}{\sqrt{Var[V]}}\right)\xrightarrow{\quad d\quad} N_2\left(\overrightarrow{0}\;,\;\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$$

donde ρ está dado en (2.3).

Demostración. Para cada pareja (X_i,Y_i) se tiene que $E[X_i]=p_{1+}$ y $E[Y_i]=p_{+1}$. Definimos el vector aleatorio

$$\overrightarrow{Z_i}=\begin{pmatrix}X_i\\ Y_i\end{pmatrix},\quad \forall\, i\in\{1,\ldots,n\}$$

y $\overrightarrow{\mu}=E\left[\overrightarrow{Z_i}\right]$, es decir,

$$\overrightarrow{\mu}=\begin{pmatrix}p_{1+}\\ p_{+1}\end{pmatrix}.$$

Así que $\overrightarrow{Z_1}, \ldots, \overrightarrow{Z_n}$ son vectores aleatorios independientes e idénticamente distribuidos, con entradas cuyas medias son finitas y cada uno con matriz de covarianza

$$\Sigma = \begin{bmatrix} Var[X_i] & Cov(X_i, Y_i) \\ Cov(Y_i, X_i) & Var[Y_i] \end{bmatrix} = \begin{bmatrix} p_{1+}p_{0+} & \tau \\ \tau & p_{+1}p_{+0} \end{bmatrix}$$

donde $Var[X_i]$ y $Var[Y_i]$ se determinaron en la Sección 2.1 y τ está dado por (2.2).

Por el Teorema de límite central multivariado, con $p = 2$, se tiene que, cuando $n \to \infty$,

$$\frac{\sum_{i=1}^{n} \overrightarrow{Z_i} \; - \; n\overrightarrow{\mu}}{\sqrt{n}} \quad \xrightarrow{\quad d \quad} \quad N_2\left(\overrightarrow{0}\;, \Sigma\right).$$

Definimos

$$\overrightarrow{W} = \frac{\sum_{i=1}^{n} \overrightarrow{Z_i} \; - \; n\overrightarrow{\mu}}{\sqrt{n}},$$

donde, de acuerdo a (2.4), (2.21) y (2.22),

$$\sum_{i=1}^{n} \overrightarrow{Z_i} = \begin{pmatrix} \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} Y_i \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix} \qquad \text{y} \qquad n\overrightarrow{\mu} = \begin{pmatrix} np_{1+} \\ np_{+1} \end{pmatrix} = \begin{pmatrix} E[U] \\ E[V] \end{pmatrix},$$

por lo que

$$\overrightarrow{W} = \frac{1}{\sqrt{n}} \begin{pmatrix} U - E[U] \\ V - E[V] \end{pmatrix}.$$

Entonces, cuando $n \to \infty$, podemos afirmar que $\overrightarrow{W} \sim N_2\left(\overrightarrow{0}, \Sigma\right)$. Por lo que si consideramos la matriz

$$R = \begin{bmatrix} \frac{1}{\sqrt{p_{1+}\, p_{0+}}} & 0 \\ 0 & \frac{1}{\sqrt{p_{1+}\, p_{0+}}} \end{bmatrix}$$

tenemos, por el Teorema 1.63, que $R\,\overrightarrow{W} \sim N_2\left(\overrightarrow{0}\;, R\,\Sigma\,R'\right)$. Donde, por (2.24) y (2.25),

$$R\,\overrightarrow{W} = \begin{pmatrix} \dfrac{U - E[U]}{\sqrt{n\,p_{1+}\,p_{0+}}} \\ \dfrac{V - E[V]}{\sqrt{n\,p_{1+}\,p_{0+}}} \end{pmatrix} = \begin{pmatrix} \dfrac{U - E[U]}{Var[U]} \\ \dfrac{V - E[V]}{\sqrt{Var[V]}} \end{pmatrix}$$

y

$$R\,\Sigma\,R' = \begin{bmatrix} \frac{1}{\sqrt{p_{1+}\, p_{0+}}} & 0 \\ 0 & \frac{1}{\sqrt{p_{1+}\, p_{0+}}} \end{bmatrix} \begin{bmatrix} p_{1+}p_{0+} & \tau \\ \tau & p_{+1}p_{+0} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{p_{1+}\, p_{0+}}} & 0 \\ 0 & \frac{1}{\sqrt{p_{1+}\, p_{0+}}} \end{bmatrix} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

Por lo tanto, cuando $n \to \infty$,

$$\begin{pmatrix} \dfrac{U - E[U]}{Var[U]} \\ \dfrac{V - E[V]}{\sqrt{Var[V]}} \end{pmatrix} \quad \xrightarrow{\quad d \quad} \quad N_2\left(\overrightarrow{0}\;, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right),$$

como afirmamos. ∎

2.2.6. Ejemplos

Ejemplo 2.16 De la población total en la que se realizó el estudio del Ejemplo 2.5 se tomó una muestra aleatoria con reemplazo de $n = 20$ casos, suponga que las observaciones son independientes y que corresponden a los siguientes vectores aleatorios $(X_1, Y_1), \ldots, (X_n, Y_n)$ con distribución de Bernoulli bivariada cuyos parámetros se muestran en la Tabla 2.1.

Sean U y V dos v.as. tales que U cuenta el número de niños que sobrevivieron y V el número de niños que utilizaban el cinturón, entonces el vector aleatorio (U, V) tiene distribución binomial bivariada (2.5).

La probabilidad de que los veinte niños hayan sobrevivido y utilizado el cinturón es casi nula, el resultado correspondiente es $P[U = 20, V = 20] = 7.3969 \times 10^{-8}$, pero es más improbable que ninguno haya sobrevivido y utilizado el cinturón, pues $P[U = 0, V = 0] = 4.0642 \times 10^{-16}$. En términos prácticos se concluye que los dos eventos son poco probables.

Una situación más razonable es preguntarse por la probabilidad de que 15 niños hayan sobrevivido y 10 utilizado el cinturón, esto es $P[U = 15, V = 10] = 3.65718965 \times 10^{-2}$. Este caso especial se ha mostrado con una cantidad de dígitos mayor que en los anteriores porque más adelante será retomado para ser comparado con la distribución hipergeométrica bivariada.

Se tiene que $P[U \geq 10, V \geq 5] = 0.9927$, significa que muy probablemente al menos la mitad de los niños hayan sobrevivido y al menos la cuarta parte haya utilizado el cinturón. Por otro lado se obtuvo el siguiente resultado $P[U \geq 10, V \geq 10] = 0.6221$, es decir, que en aproximadamente 12 de los 20 casos (aproximadamente el 60 % de la muestra) al menos la mitad de los niños sobrevivieron y utilizaban el cinturón.

La forma en que se realizaron los cálculos para el último párrafo fue utilizando la siguiente expresión general

$$P[U \geq u, V \geq v] = 1 - \left\{ \sum_{i=0}^{u-1} \sum_{j=0}^{n} P[U = i, V = j] + \sum_{i=u}^{n} \sum_{j=0}^{v-1} P[U = i, V = j] \right\}. \quad \blacktriangleleft$$

Ejemplo 2.17 Considere el Ejemplo 2.8, a los pacientes que llegan a una consulta se les diagnostica por clínica y se les hace una orden de ultrasonido para confirmar el diagnóstico o descartar otras posibilidades. El encargado de realizar el estudio encontró tres situaciones que distinguen a los pacientes, aquellos que acuden al hospital y son tratados por ambos métodos, los que sólo fueron diagnosticados por clínica (muchos pacientes prefieren realizar el ultrasonido en otro lugar), y los que llegaron directamente a hacerse un ultrasonido (esto ocurre porque se les hizo una orden de ultrasonido en otro hospital).

De la población total se tomó una muestra aleatoria con reemplazo de $k = 50$ casos para los pacientes que fueron tratados por ambos métodos, otra muestra similar de $n - k = 30$ casos para los pacientes que fueron tratados sólo por clínica, y una última muestra de $m - k = 25$ casos para los que fueron tratados sólo con un ultrasonido. Suponga que las observaciones son independientes. Para la primer muestra cada observación corresponde a los vectores aleatorios $(X_1, Y_1), \ldots, (X_k, Y_k)$ con distribución de Bernoulli bivariada cuyos parámetros se muestran en la Tabla 2.2, a la segunda muestra le corresponden las v.as. $X'_1, \ldots, X'_{n-k}$ y a la tercera $Y'_1, \ldots, Y'_{m-k}$. Donde, para cada $i \in \{1, \ldots, k\}$ y $j \in \{1, \ldots, n-k\}$, las v.as. X_i y X'_j tienen la misma distribución que X en el Ejemplo 2.8, así como, para cada $i \in \{1, \ldots, k\}$ y $l \in \{1, \ldots, m-k\}$, las v.as. Y_i y Y'_l tienen la misma distribución que Y en el mismo ejemplo.

El estudio está enfocado en el número de pacientes diagnosticados por clínica y en el número de pacientes diagnosticados por un ultrasonido.

Si U, V, S y T son v.as. tales que $(U,V) = \sum_{i=i}^{k}(X_i, Y_i)$, $S = \sum_{j=i}^{n-k} X'_j$ y $T = \sum_{l=i}^{m-k} Y'_l$, entonces (U,V) tiene distribución binomial bivariada dada por (2.5), $S \sim Binomial(n-k, p_{1+})$ y $T \sim Binomial(m-k, p_{+1})$, donde p_{1+} y p_{+1} están dados en la Tabla 2.2.

Suponga que (U,V), S y T son independientes. Si U^* y V^* se definen como $U^* = U + S$ y $V^* = V + T$, entonces (U^*, V^*) tiene distribución BBG dada por (2.7), donde U^* cuenta el número de pacientes diagnosticados por clínica y V^* el número de pacientes diagnosticados por un ultrasonido.

La probabilidad de que 20 pacientes hayan sido diagnosticados por clínica y 10 por un ultrasonido es prácticamente nula: $P[U^* = 20, V^* = 10] = 2.7480 \times 10^{-56}$. Es mucho más probable que 70 pacientes hayan sido diagnosticados por clínica y 60 por un ultrasonido: $P[U^* = 70, V^* = 60] = 1.9433 \times 10^{-3}$, que era de esperarse por las probabilidades marginales dadas en la Tabla 2.2.

Observe que cuando $k = m = n$ es indistinto utilizar (2.5) o (2.7), ya que en este caso la distribución binomial bivariada coincide con la distribución BBG. Por ejemplo, si $k = m = n = 50$, utilizando cualquiera de los dos modelos se obtiene $P[U^* = 42, V^* = 44] = 2.2965 \times 10^{-2}$.

El coeficiente de correlación entre U y V es determinado mediante (2.19). Se tiene que $Corr(U,V) = 0.2494$, por lo que U y V son poco dependientes. ◀

2.3. La distribución de Poisson bivariada

Definición 2.18 *Se dice que un vector aleatorio* (U,V) *tiene* **distribución de Poisson bivariada** *(PB) con parámetros* $\lambda_{01} > 0$, $\lambda_{10} > 0$ *y* $\lambda_{11} > 0$ *si, para todo* $u, v = 0, 1, \ldots,$ *se cumple*

$$P[U = u, V = v] = \sum_{a=0}^{\min(u,v)} \frac{\lambda_{11}^{a}\, \lambda_{10}^{u-a}\, \lambda_{01}^{v-a}}{a!\,(u-a)!\,(v-a)!}\, e^{-(\lambda_{01}+\lambda_{10}+\lambda_{11})}. \tag{2.36}$$

Se sabe que existe una importante relación entre las distribuciones binomial y de Poisson en el caso univariado, conocida en la literatura como la *aproximación de Poisson a la distribución Binomial*, la cual afirma, bajo ciertas condiciones, que la distribución binomial converge a la de Poisson. En esta sección se pretende demostrar este resultado en el caso bivariado.

Teorema 2.19 *Sea* $\{(X_n, Y_n)\}_{n=1}^{\infty}$ *una sucesión de vectores aleatorios independientes tal que* (X_n, Y_n) *tiene distribución de Bernoulli bivariada con parámetros* $p_{11,n}$, $p_{10,n}$, $p_{01,n}$ *y* $p_{00,n}$ *en* $]0,1[$. *Se supone que*

$$\lim_{n\to\infty} np_{11,n} = \lambda_{11}, \quad \lim_{n\to\infty} np_{10,n} = \lambda_{10}, \quad \lim_{n\to\infty} np_{01,n} = \lambda_{01}$$

y $p_{00,n} = p_{00}$, $\forall n \in \mathbb{N}$, *donde* $\lambda_{11}, \lambda_{10}, \lambda_{01} > 0$. *Defina* (U_n, V_n) *como*

$$U_n = \sum_{k=1}^{n} X_k \qquad y \qquad V_n = \sum_{k=1}^{n} Y_k, \quad \forall n \in \mathbb{N}.$$

Si (U,V) *es un vector aleatorio con distribución de Poisson bivariada dada por* (2.36), *entonces*

$$\lim_{n\to\infty} P(U_n = u, V_n = v) = P(U = u, V = v), \quad u, v = 0, 1, \ldots,$$

es decir, que la distribución binomial bivariada converge a la de Poisson bivariada o que la distribución de Poisson bivariada "aproxima" a la distribución binomial bivariada.

Demostración. Por los resultados de la Sección 2.2, la pareja (U_n, V_n) tiene la distribución Binomial bivariada dada por (2.5), la cual en este caso sería

$$P(U_n = u, V_n = v) = \sum_a \binom{n}{a, u-a, v-a, n-u-v+a} \times \left(\frac{np_{11,n}}{n}\right)^a \left(\frac{np_{10,n}}{n}\right)^{u-a} \left(\frac{np_{01,n}}{n}\right)^{v-a} \left(1 - \frac{np_{11,n}}{n} - \frac{np_{10,n}}{n} - \frac{np_{01,n}}{n}\right)^{n-u-v+a}.$$

Tomamos el límite cuando $n \to \infty$ para obtener

$$\begin{aligned}
\lim_{n\to\infty} P(U_n = u, V_n = v) &= \lim_{n\to\infty} \Bigg\{ \sum_a \frac{n!}{a!\,(u-a)!\,(v-a)!\,(n-u-v+a)!} \\
&\quad \times \frac{(np_{11,n})^a\,(np_{10,n})^{u-a}\,(np_{01,n})^{v-a}}{n^a\,n^{u-a}\,n^{v-a}} \left(1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right)^{n-u-v+a} \Bigg\} \\
&= \sum_a \frac{\lim_{n\to\infty} \left[(np_{11,n})^a\,(np_{10,n})^{u-a}\,(np_{01,n})^{v-a}\right]}{a!\,(u-a)!\,(v-a)!} \Bigg\{ \lim_{n\to\infty} \left[\frac{n!}{(n-(u+v-a))!}\right] \\
&\quad \times \left[\frac{1}{n^{u+v-a}}\right] \left[1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right]^{n-(u+v-a)} \Bigg\} \\
&= \sum_a \frac{\lambda_{11}^a\,\lambda_{10}^{u-a}\,\lambda_{01}^{v-a}}{a!\,(u-a)!\,(v-a)!} \Bigg\{ \lim_{n\to\infty} \left[\frac{n!}{(n-(u+v-a))!\;n^{u+v-a}}\right] \\
&\quad \times \left[1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right]^{n} \left[1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right]^{-(u+v-a)} \Bigg\}.
\end{aligned}$$

Observe que el límite anterior existe, primero porque es una suma finita, luego porque, para $k = u+v-a$,

$$\lim_{n\to\infty} \frac{n!}{(n-k)!\,n^k} = 1,$$

$$\lim_{n\to\infty} \left[1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right]^{-(u+v-a)} = 1$$

y porque, para $a_n = -(np_{11,n} + np_{10,n} + np_{01,n})$, se cumple

$$\lim_{n\to\infty} a_n = -(\lambda_{11} + \lambda_{10} + \lambda_{01}) = a,$$

luego

$$\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = e^a.$$

Así pues,

$$\lim_{n\to\infty} \left[\frac{n!}{(n-(u+v-a))!\;n^{u+v-a}}\right] \times \left[1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right]^{n} \left[1 - \frac{np_{11,n} + np_{10,n} + np_{01,n}}{n}\right]^{-(u+v-a)} = e^{-(\lambda_{11}+\lambda_{10}+\lambda_{01})}.$$

Por el desarrollo anterior se concluye que

$$\lim_{n\to\infty} P(U_n = u, V_n = v) = \sum_a \frac{\lambda_{11}^a\,\lambda_{10}^{u-a}\,\lambda_{01}^{v-a}}{a!\,(u-a)!\,(v-a)!}\, e^{-(\lambda_{01}+\lambda_{10}+\lambda_{11})} = P(U = u, V = v)$$

que corresponde a la distribución PB dada por (2.36). ∎

El resultado siguiente es una definición alternativa de la distribución de Poisson bivariada propuesta en [14]. Esta representación de la distribución de Poisson bivariada será de gran utilidad en la derivación de la $f.g.m.f.$, covarianza, coeficiente de correlación y valores esperados condicionales asociados a esta distribución.

Teorema 2.20 *Un vector aleatorio* (U, V) *tiene la distribución de Poisson bivariada dada por* (2.36) *si y sólo si existen variables aleatorias independientes* X, Y *y* Z *con distribuciones de Poisson con parámetros* λ_{10}, λ_{01} *y* λ_{11}, *respectivamente, tales que*

$$U = X + Z \qquad y \qquad V = Y + Z. \tag{2.37}$$

Demostración. Supongamos entonces que $U = X + Z$ y $V = Y + Z$ con $X \sim Poisson(\lambda_{10})$, $Y \sim Poisson(\lambda_{01})$ y $Z \sim Poisson(\lambda_{11})$ independientes. Entonces

$$U \sim Poisson(\lambda_{10} + \lambda_{11}) \qquad \text{y} \qquad V \sim Poisson(\lambda_{01} + \lambda_{11}). \tag{2.38}$$

Observe que $P[U = u, V = v] = P[X + Z = u, Y + Z = v]$, entonces como $X + Z = u$ y $Y + Z = v$, se tiene $X = u - Z$ y $Y = v - Z$, de ahí que todos los posibles valores $Z = z$ no pueden superar a u o v, es decir $z \in \{0, 1, \ldots, \text{mín}(u, v)\}$, de otra forma, por ejemplo, si $z = u + 1$ entonces $x = -1$ lo cual es absurdo. Una situación análoga sirve para mostrar que z no puede ser mayor que v. Por lo que, considerando todos los posibles valores para z, se tiene

$$\begin{aligned}
P[X + Z = u, Y + Z = v] &= \sum_{z=0}^{\text{mín}(u,v)} P[X + z = u, Y + z = v, Z = z] \\
&= \sum_{z} P[X = u - z, Y = v - z, Z = z] \\
&= \sum_{z} P[X = u - z] P[Y = v - z] P[Z = z] \\
&= \sum_{z} \frac{\lambda_{10}^{u-z}}{(u-z)!}\, e^{-\lambda_{10}}\, \frac{\lambda_{01}^{v-z}}{(v-z)!}\, e^{-\lambda_{01}}\, \frac{\lambda_{11}^{z}}{z!}\, e^{-\lambda_{11}} \\
&= \sum_{z} \frac{\lambda_{11}^{z}\, \lambda_{10}^{u-z}\, \lambda_{01}^{v-z}}{z!\,(u-z)!\,(v-z)!}\, e^{-(\lambda_{11}+\lambda_{10}+\lambda_{01})},
\end{aligned}$$

donde la tercera desigualdad está justificada por independencia y la última corresponde a la distribución PB dada por (2.36), es decir que, el vector aleatorio $(U, V) = (X + Z, Y + Z)$ satisface la definición de distribución de Poisson bivariada.

Note por cierto que si $\lambda_1 = \lambda_{10} + \lambda_{11}$ y $\lambda_2 = \lambda_{01} + \lambda_{11}$, entonces

$$P[X + Z = u, Y + Z = v] = \sum_{a=0}^{\text{mín}(u,v)} \frac{\lambda_{11}^{a}\, (\lambda_1 - \lambda_{11})^{u-a}\, (\lambda_2 - \lambda_{11})^{v-a}}{a!\,(u-a)!\,(v-a)!}\, e^{-(\lambda_2 - \lambda_{11} + \lambda_1 - \lambda_{11} + \lambda_{11})},$$

por lo que (2.36) puede ser escrita en forma alternativa como

$$P[U = u, V = v] = \sum_{a=0}^{\text{mín}(u,v)} \frac{\lambda_{11}^{a}\, (\lambda_1 - \lambda_{11})^{u-a}\, (\lambda_2 - \lambda_{11})^{v-a}}{a!\,(u-a)!\,(v-a)!}\, e^{-(\lambda_1 + \lambda_2 - \lambda_{11})}. \tag{2.39}$$

Recíprocamente, suponga que (U,V) tienen la distribución (2.36), entonces

$$\begin{aligned}P[U=u,V=v] &= \sum_a \frac{\lambda_{11}^a\,\lambda_{10}^{u-a}\,\lambda_{01}^{v-a}}{a!\,(u-a)!\,(v-a)!}\,e^{-(\lambda_{11}+\lambda_{10}+\lambda_{01})}\\ &= \sum_a \frac{\lambda_{10}^{u-a}}{(u-a)!}\,e^{-\lambda_{10}}\,\frac{\lambda_{01}^{v-a}}{(v-a)!}\,e^{-\lambda_{01}}\,\frac{\lambda_{11}^a}{a!}\,e^{-\lambda_{11}}.\end{aligned}$$

Note que si X, Y y Z son variables aleatorias independientes con distribuciones de Poisson con parámetros λ_{10}, λ_{01} y λ_{11}, se sigue de la última identidad que

$$\begin{aligned}P[U=u,V=v] &= \sum_a P[X=u-a]P[Y=v-a]P[Z=a]\\ &= \sum_a P[X=u-a,Y=v-a,Z=a]\\ &= P[X+Z=u,Y+Z=v],\end{aligned}$$

lo cual significa que los vectores aleatorios (U,V) y $(X+Z,Y+Z)$ son idénticamente distribuidos, luego, U, V aceptan la representación (2.37). ∎

Observación 2.21 El teorema anterior aporta dos resultados importantes. El primero es que si un vector aleatorio (U,V) tiene la distribución PB dada por (2.36) entonces las distribuciones marginales para U y V están dadas por (2.38). El otro resultado es que la distribución de Poisson bivariada está dada indistintamente por (2.36) o (2.39), de hecho esta última será utilizada en la derivación de los valores esperados condicionales que serán discutidos más adelante. ◁

Proposición 2.22 *Si (U,V) es un vector aleatorio con la distribución PB dada por (2.36), entonces su f.g.m.f. está dada por*

$$m_{U,V}(s,t) = e^{\lambda_{10}(s-1)+\lambda_{01}(t-1)+\lambda_{11}(st-1)}. \tag{2.40}$$

Demostración. Por el Teorema 2.20 sabemos que U y V aceptan la representación (2.37), así es que

$$\begin{aligned}m_{U,V}(s,t) = E\left[s^U\,t^V\right] = E\left[s^{X+Z}\,t^{Y+Z}\right] &= E\left[s^X\,s^Z\,t^Y\,t^Z\right]\\ &= E\left[s^X\right]E\left[t^Y\right]E\left[(st)^Z\right] = m_X(s)\,m_Y(t)\,m_Z(st)\end{aligned}$$

y por lo tanto, dado que la *f.g.m.f.* para una variable aleatoria $X\sim Poisson(\lambda)$ es $m_X(s)=e^{\lambda(s-1)}$, se obtiene (2.40). ∎

Otra cantidad de interés es el coeficiente de correlación de (U,V) que será obtenido de la *f.g.m.f.* (2.40) como se muestra a continuación.

Proposición 2.23 *Se tiene*

$$Corr(U,V) = \frac{\lambda_{11}}{\sqrt{(\lambda_{10}+\lambda_{11})(\lambda_{01}+\lambda_{11})}}. \tag{2.41}$$

Demostración. Vamos a proceder de la misma forma que hicimos en la Sección 2.2 recurriendo a la $f.g.m.f$. Se hará uso de la definición de correlación. Calculemos $Cov(U, V)$ lo cual requiere de $E[UV]$, $E[U]$ y $E[V]$. Se tiene que

$$E[UV] = \left.\frac{\partial^2 m(s,t)}{\partial s\, \partial t}\right|_{s=1,t=1},$$

donde

$$\frac{\partial m(s,t)}{\partial t} = (\lambda_{01} + \lambda_{11} s) e^{\lambda_{10}(s-1)+\lambda_{01}(t-1)+\lambda_{11}(st-1)}$$

y

$$\begin{aligned}\frac{\partial^2 m(s,t)}{\partial s \partial t} &= (\lambda_{01} + \lambda_{11} s)(\lambda_{10} + \lambda_{11} t) e^{\lambda_{10}(s-1)+\lambda_{01}(t-1)+\lambda_{11}(st-1)} + \lambda_{11} e^{\lambda_{10}(s-1)+\lambda_{01}(t-1)+\lambda_{11}(st-1)} \\ &= \left[(\lambda_{01} + \lambda_{11} s)(\lambda_{10} + \lambda_{11} t) + \lambda_{11}\right] e^{\lambda_{10}(s-1)+\lambda_{01}(t-1)+\lambda_{11}(st-1)},\end{aligned}$$

de modo que

$$E[UV] = (\lambda_{01} + \lambda_{11})(\lambda_{10} + \lambda_{11}) + \lambda_{11}.$$

Luego, por (2.38), se tiene que $E[U] = \lambda_{10} + \lambda_{11}$ y $E[V] = \lambda_{01} + \lambda_{11}$. Entonces

$$E[UV] - E[U]E[V] = (\lambda_{01} + \lambda_{11})(\lambda_{10} + \lambda_{11}) + \lambda_{11} - (\lambda_{01} + \lambda_{11})(\lambda_{10} + \lambda_{11}) = \lambda_{11}$$

y finalmente

$$Cov(U, V) = \lambda_{11}. \tag{2.42}$$

Para completar la prueba requerimos de $Var[U]$ y $Var[V]$ que se tienen inmediatamente por las distribuciones de U y de V, mostradas en (2.38). Así que

$$Var[U] = \lambda_{10} + \lambda_{11} \quad \text{y} \quad Var[V] = \lambda_{01} + \lambda_{11}$$

por lo tanto (2.41). ∎

Note que $Corr(U, V) > 0$ debido a que los parámetros de la distribución PB, dados en la Definición 2.18, son todos positivos.

2.3.1. Valor esperado condicional

Vamos a calcular los valores esperados condicionales correspondientes a la distribución PB dada por (2.36).

Teorema 2.24 *Sea (U, V) un vector aleatorio con distribución de Poisson bivariada dada por* (2.39). *Entonces las esperanzas condicionales de U dado que $V = v$ y de V dado que $U = u$ están dadas, respectivamente, por*

$$E[U \mid V = v] = \lambda_{10} + v\,\frac{\lambda_{11}}{\lambda_{01} + \lambda_{11}}, \tag{2.43}$$

$$E[V \mid U = u] = \lambda_{01} + u\,\frac{\lambda_{11}}{\lambda_{10} + \lambda_{11}}. \tag{2.44}$$

Demostración. Se tiene,

$$E[U \mid V=v] = E[X+Z \mid Y+Z=v] = E[X \mid Y+Z=v] + E[Z \mid Y+Z=v] = E[X] + E[Z \mid Y+Z=v], \tag{2.45}$$

donde la última igualdad en (2.45) se cumple por la independencia entre las variables aleatorias X, Y, Z de la representación (2.37). Por otro lado,

$$E[Z \mid Y+Z=v] = \sum_z zP[Z=z \mid Y+Z=v] = \sum_z z\,\frac{P[Z=z, Y+Z=v]}{P[Y+Z=v]} = \sum_z z\,\frac{P[Z=z, Y=v-z]}{P[Y+Z=v]},$$

donde

$$P[Z=z, Y=v-z] = P[Z=z]P[Y=v-z] = \left(\frac{\lambda_{11}^z\, e^{-\lambda_{11}}}{z!}\right)\left(\frac{\lambda_{01}^{v-z}\, e^{-\lambda_{01}}}{(v-z)!}\right),$$

por independencia, y

$$P[Y+Z=v] = \frac{(\lambda_{01}+\lambda_{11})^v\, e^{-(\lambda_{01}+\lambda_{11})}}{v!},$$

luego

$$\frac{P[Z=z, Y+Z=v]}{P[Y+Z=v]} = \frac{v!\,\lambda_{11}^z\,\lambda_{01}^{v-z}}{z!\,(v-z)!\,(\lambda_{01}+\lambda_{11})^v} = \binom{v}{z}\frac{\lambda_{11}^z\,\lambda_{01}^{v-z}}{(\lambda_{01}+\lambda_{11})^v} = \binom{v}{z}\left(\frac{\lambda_{11}}{\lambda_{01}+\lambda_{11}}\right)^z\left(\frac{\lambda_{01}}{\lambda_{01}+\lambda_{11}}\right)^{v-z},$$

donde la última igualdad se obtiene al multiplicar por

$$\frac{(\lambda_{01}+\lambda_{11})^z}{(\lambda_{01}+\lambda_{11})^z}.$$

En esta forma,

$$\frac{P[Z=z, Y+Z=v]}{P[Y+Z=v]} = \binom{v}{z}\left(\frac{\lambda_{11}}{\lambda_{01}+\lambda_{11}}\right)^z\left(\frac{\lambda_{01}}{\lambda_{01}+\lambda_{11}}\right)^{v-z},$$

es decir,

$$P[Z=z \mid Y+Z=v] = \binom{v}{z}\left(\frac{\lambda_{11}}{\lambda_{01}+\lambda_{11}}\right)^z\left(\frac{\lambda_{01}}{\lambda_{01}+\lambda_{11}}\right)^{v-z}.$$

Como esta distribución condicional se identifica con una distribución binomial$\left(v, \frac{\lambda_{11}}{\lambda_{01}+\lambda_{11}}\right)$, se debe tener

$$E[Z \mid Y+Z=v] = v\,\frac{\lambda_{11}}{\lambda_{01}+\lambda_{11}} \tag{2.46}$$

Entonces, sustituyendo (2.46) en (2.45) y sabiendo que $E[X]=\lambda_{10}$, se concluye que

$$E[U \mid V=v] = \lambda_{10} + v\,\frac{\lambda_{11}}{\lambda_{01}+\lambda_{11}},$$

que es (2.43).

Intercambiando los papeles de U y V se demuestra similarmente (2.44). ∎

2.3.2. Convergencia a la distribución normal bivariada

En esta sección demostraremos, utilizando (2.37), que cuando los parámetros λ_{10}, λ_{01} y λ_{11} se aproximan al infinito de tal suerte que λ_{10}/λ, λ_{01}/λ y λ_{11}/λ permanezcan constantes, donde $\lambda = \lambda_{10}+\lambda_{01}+\lambda_{11}$, entonces el vector aleatorio estandarizado

$$\left(\frac{U-E[U]}{\sqrt{Var[U]}}\,,\;\frac{V-E[V]}{\sqrt{Var[V]}}\right)$$

converge en distribución a la distribución normal bivariada con media $\overrightarrow{\mu} = \overrightarrow{0}$ y matriz de covarianzas

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}.$$

Sean $x = \lambda_{10}$, $y = \lambda_{01}$ y $z = \lambda_{11}$. Si $f(x,y,z) = \frac{x}{x+y+z}$, se tiene que

$$0 \leq \lim_{(x,y,z)\to\infty} f(x,y,z) \leq 1$$

siempre que el límite exista donde su valor dependerá de la forma en que las componentes se aproximen a infinito. Por ejemplo, si $z = y = x$ entonces $\lim_{(x,y,z)\to\infty} f(x,y,z) = 1/3$. Pero si $x \gg y$ (con lo que nos referimos a que x es mucho mayor que y) y $x \gg z$, se tendrá que dicho límite es igual a uno, También, si $y \gg x$ y $z \gg x$, se tendrá que dicho límite es igual a cero. Algo similar ocurre para $g(x,y,z) = \frac{y}{x+y+z}$ y $h(x,y,z) = \frac{z}{x+y+z}$. Por lo tanto, seleccionaremos un comportamiento de los parámetros λ_{10}, λ_{01} y λ_{11} de (2.37) de tal suerte que éstos se aproximan al infinito de forma tal que

$$\frac{\lambda_{10}}{\lambda} = \frac{\lambda_{01}}{\lambda} = \frac{\lambda_{11}}{\lambda} \to k \tag{2.47}$$

para alguna constante $k \in \mathbb{R}$.

Proposición 2.25 *Sea* (U,V) *un vector aleatorio con distribución* PB *dada por* (2.36). *Si* $\lambda_{10} \to \infty$, $\lambda_{01} \to \infty$ *y* $\lambda_{11} \to \infty$ *de tal suerte que se cumpla* (2.47), *entonces*

$$Corr(U,V) = \frac{1}{2}\,. \tag{2.48}$$

Demostración. Sabemos que, por (2.41)

$$Corr(U,V) = \frac{\lambda_{11}}{\sqrt{(\lambda_{10}+\lambda_{11})(\lambda_{01}+\lambda_{11})}}.$$

Haciendo tender $\lambda_{10} \to \infty$, $\lambda_{01} \to \infty$ y $\lambda_{11} \to \infty$ de acuerdo a (2.47), se tiene que

$$\begin{aligned}
\frac{\lambda_{11}}{\sqrt{(\lambda_{10}+\lambda_{11})(\lambda_{01}+\lambda_{11})}} &= \frac{\frac{\lambda_{11}}{\lambda}}{\frac{\sqrt{(\lambda_{10}+\lambda_{11})(\lambda_{01}+\lambda_{11})}}{\lambda}} \\
&= \frac{\frac{\lambda_{11}}{\lambda}}{\sqrt{\frac{(\lambda_{10}+\lambda_{11})(\lambda_{01}+\lambda_{11})}{\lambda^2}}} \\
&= \frac{\frac{\lambda_{11}}{\lambda}}{\sqrt{\left(\frac{\lambda_{10}+\lambda_{11}}{\lambda}\right)\left(\frac{\lambda_{01}+\lambda_{11}}{\lambda}\right)}} \\
&= \frac{\frac{\lambda_{11}}{\lambda}}{\sqrt{\left(\frac{\lambda_{10}}{\lambda}+\frac{\lambda_{11}}{\lambda}\right)\left(\frac{\lambda_{01}}{\lambda}+\frac{\lambda_{11}}{\lambda}\right)}} \\
&= \frac{k}{\sqrt{(2k)\,(2k)}} \\
&= \frac{1}{2}.
\end{aligned}$$

Por lo tanto se cumple (2.48). ∎

En el caso de las distribuciones bivariadas anteriores se ha hecho uso sólo de la $f.g.m.f.$ Para el caso de la presente distribución bivariada resultará más conveniente utilizar la función generadora de momentos $f.g.m.$ La siguiente proposición será vital para la demostración del Teorema 2.27.

Proposición 2.26 *Sea (U, V) un vector aleatorio con distribución PB dada por* (2.36)*. Entonces la $f.g.m.$ conjunta de (U, V) está dada por*

$$\psi_{U,V}(s,t) = e^{\lambda_{10}(\exp(s)-1)+\lambda_{01}(\exp(t)-1)+\lambda_{11}(\exp(s+t)-1)}. \tag{2.49}$$

Demostración. El resultado se sigue inmediatamente de la relación entre funciones generadoras de momentos conjuntas (Lema 1.48),

$$\psi_{U,V}(s,t) = m_{U,V}(\exp(s), \exp(t)).$$

En este caso, por (2.40), tenemos

$$\psi_{U,V}(s,t) = e^{\lambda_{10}(\exp(s)-1)+\lambda_{01}(\exp(t)-1)+\lambda_{11}(\exp(s)\exp(t)-1)} = e^{\lambda_{10}(\exp(s)-1)+\lambda_{01}(\exp(t)-1)+\lambda_{11}(\exp(s+t)-1)}.$$

Por lo tanto se cumple (2.49). ∎

El siguiente teorema establece la convergencia de la distribución PB dada por (2.36) a la distribución normal bivariada.

Teorema 2.27 (Convergencia a la distribución normal bivariada) *Sea (U, V) un vector aleatorio con distribución PB dada por* (2.36)*. De la representación* (2.37)*, si $\lambda_{10} \to \infty$, $\lambda_{01} \to \infty$ y $\lambda_{11} \to \infty$ de tal suerte que se cumpla* (2.47)*, entonces*

$$\left(\frac{U - E[U]}{\sqrt{Var[U]}}, \frac{V - E[V]}{\sqrt{Var[V]}}\right) \xrightarrow{\quad d \quad} N_2\left(\vec{0}, \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}\right).$$

Demostración. Sean $\lambda_1 = \lambda_{10} + \lambda_{11}$, $\lambda_2 = \lambda_{01} + \lambda_{11}$,

$$W_1 = \frac{U - \lambda_1}{\sqrt{\lambda_1}} \qquad \text{y} \qquad W_2 = \frac{V - \lambda_2}{\sqrt{\lambda_2}}.$$

Vamos a calcular la *f.g.m.* conjunta para W_1 y W_2. Se tiene,

$$\begin{aligned}
\psi_{W_1,W_2}(s,t) &= E\left[\exp\left(s\,\frac{U-\lambda_1}{\sqrt{\lambda_1}} + t\,\frac{V-\lambda_2}{\sqrt{\lambda_2}}\right)\right] \\
&= E\left[\exp\left(\frac{sU}{\sqrt{\lambda_1}} - s\,\sqrt{\lambda_1} + \frac{tV}{\sqrt{\lambda_2}} - t\,\sqrt{\lambda_2}\right)\right] \\
&= E\left[\exp\left(\frac{sU}{\sqrt{\lambda_1}} + \frac{tV}{\sqrt{\lambda_2}}\right)\exp\left(-s\,\sqrt{\lambda_1} - t\,\sqrt{\lambda_2}\right)\right] \\
&= \exp\left(-s\,\sqrt{\lambda_1} - t\,\sqrt{\lambda_2}\right) E\left[\exp\left(\frac{sU}{\sqrt{\lambda_1}} + \frac{tV}{\sqrt{\lambda_2}}\right)\right] \\
&= \exp\left(-s\,\sqrt{\lambda_1} - t\,\sqrt{\lambda_2}\right) \psi_{U,V}\left(\frac{s}{\sqrt{\lambda_1}}, \frac{t}{\sqrt{\lambda_2}}\right),
\end{aligned}$$

donde en la última igualdad aparece la *f.g.m.* dada por (2.49). Luego,

$$\begin{aligned}
\psi_{W_1,W_2}(s,t) &= \exp\left(-s\,\sqrt{\lambda_1} - t\,\sqrt{\lambda_2}\right) \\
&\quad\times \exp\left(\lambda_{10}\left\{\exp\left(\frac{s}{\sqrt{\lambda_1}}\right) - 1\right\} + \lambda_{01}\left\{\exp\left(\frac{t}{\sqrt{\lambda_2}}\right) - 1\right\} + \lambda_{11}\left\{\exp\left(\frac{s}{\sqrt{\lambda_1}} + \frac{t}{\sqrt{\lambda_2}}\right) - 1\right\}\right).
\end{aligned}$$

Para simplificar la escritura introducimos las siguientes notaciones

$$A = \lambda_{10}\left\{\exp\left(\frac{s}{\sqrt{\lambda_1}}\right) - 1\right\}, \qquad B = \lambda_{01}\left\{\exp\left(\frac{t}{\sqrt{\lambda_2}}\right) - 1\right\}$$

$$\text{y} \qquad C = \lambda_{11}\left\{\exp\left(\frac{s}{\sqrt{\lambda_1}} + \frac{t}{\sqrt{\lambda_2}}\right) - 1\right\}.$$

Con esta notación se tiene

$$\psi_{W_1,W_2}(s,t) = e^{\left(-s\,\sqrt{\lambda_1} - t\,\sqrt{\lambda_2}\right)} e^{A+B+C}. \tag{2.50}$$

Utilizamos ahora el desarrollo en series de la función exponencial, vea la Sección 1.3, para obtener

$$\exp\left(\frac{s}{\sqrt{\lambda_1}}\right) = \sum_{i=0}^{\infty} \frac{\left(\frac{s}{\sqrt{\lambda_1}}\right)^i}{i!} = 1 + \frac{s}{\sqrt{\lambda_1}} + \frac{s^2}{2!\,\lambda_1} + \frac{s^3}{3!\,\lambda_1^{3/2}} + \frac{s^4}{4!\,\lambda_1^2} + \cdots,$$

de modo que

$$\lambda_{10}\left\{\exp\left(\frac{s}{\sqrt{\lambda_1}}\right) - 1\right\} = \frac{\lambda_{10}\,s}{\sqrt{\lambda_1}} + \frac{\lambda_{10}\,s^2}{2\,\lambda_1} + \frac{\lambda_{10}\,s^3}{3!\,\lambda_1^{3/2}} + \frac{\lambda_{10}\,s^4}{4!\,\lambda_1^2} + \cdots.$$

Similarmente,

$$\lambda_{01}\left\{\exp\left(\frac{t}{\sqrt{\lambda_2}}\right) - 1\right\} = \frac{\lambda_{01}\,t}{\sqrt{\lambda_2}} + \frac{\lambda_{01}\,t^2}{2\,\lambda_2} + \frac{\lambda_{01}\,t^3}{3!\,\lambda_2^{3/2}} + \frac{\lambda_{01}\,t^4}{4!\,\lambda_2^2} + \cdots.$$

Así mismo,

$$\begin{aligned}
\exp\left(\frac{s}{\sqrt{\lambda_1}}+\frac{t}{\sqrt{\lambda_2}}\right) &= \exp\left(\frac{s\sqrt{\lambda_2}+t\sqrt{\lambda_1}}{\sqrt{\lambda_1\lambda_2}}\right)\\
&= \sum_{i=0}^{\infty}\frac{1}{i!}\left(\frac{s\sqrt{\lambda_2}+t\sqrt{\lambda_1}}{\sqrt{\lambda_1\lambda_2}}\right)^i\\
&= 1+\left(\frac{s}{\sqrt{\lambda_1}}+\frac{t}{\sqrt{\lambda_2}}\right)+\frac{1}{2!}\left(\frac{s\sqrt{\lambda_2}+t\sqrt{\lambda_1}}{\sqrt{\lambda_1\lambda_2}}\right)^2\\
&\qquad +\frac{1}{3!}\left(\frac{s\sqrt{\lambda_2}+t\sqrt{\lambda_1}}{\sqrt{\lambda_1\lambda_2}}\right)^3+\frac{1}{4!}\left(\frac{s\sqrt{\lambda_2}+t\sqrt{\lambda_1}}{\sqrt{\lambda_1\lambda_2}}\right)^4+\cdots\\
&= 1+\frac{s}{\sqrt{\lambda_1}}+\frac{t}{\sqrt{\lambda_2}}+\frac{s^2\,\lambda_2+2st\sqrt{\lambda_1\lambda_2}+t^2\,\lambda_1}{2\,\lambda_1\lambda_2}\\
&\quad +\frac{s^3\,\lambda_2^{3/2}+3s^2t\sqrt{\lambda_1}\lambda_2+3st^2\,\lambda_1\sqrt{\lambda_2}+t^3\,\lambda_1^{3/2}}{3!\,(\lambda_1\lambda_2)^{3/2}}\\
&\qquad +\frac{s^4\,\lambda_2^2+4s^3t\sqrt{\lambda_1}\lambda_2^{3/2}+6s^2t^2\,\lambda_1\lambda_2+4st^3\,\lambda_1^{3/2}\sqrt{\lambda_2}+t^4\,\lambda_1^2}{4!\,(\lambda_1\lambda_2)^2}+\cdots\\
&= 1+\frac{s}{\sqrt{\lambda_1}}+\frac{t}{\sqrt{\lambda_2}}+\frac{s^2}{2\,\lambda_1}+\frac{st}{\sqrt{\lambda_1\lambda_2}}+\frac{t^2}{2\,\lambda_2}\\
&\quad +\frac{s^3}{3!\,\lambda_1^{3/2}}+\frac{s^2t}{2\,\lambda_1\sqrt{\lambda_2}}+\frac{st^2}{2\sqrt{\lambda_1}\lambda_2}+\frac{t^3}{3!\,\lambda_2^{3/2}}\\
&\qquad +\frac{s^4}{4!\,\lambda_1^2}+\frac{s^3t}{3!\,\lambda_1^{3/2}\sqrt{\lambda_2}}+\frac{6s^2t^2}{4!\,\lambda_1\lambda_2}+\frac{st^3}{3!\sqrt{\lambda_1}\lambda_2^{3/2}}+\frac{t^4}{4!\,\lambda_2^2}+\cdots.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\lambda_{11}\left\{\exp\left(\frac{s}{\sqrt{\lambda_1}}+\frac{t}{\sqrt{\lambda_2}}\right)-1\right\} &= \frac{\lambda_{11}s}{\sqrt{\lambda_1}}+\frac{\lambda_{11}t}{\sqrt{\lambda_2}}+\frac{\lambda_{11}s^2}{2\,\lambda_1}+\frac{\lambda_{11}st}{\sqrt{\lambda_1\lambda_2}}+\frac{\lambda_{11}t^2}{2\,\lambda_2}\\
&\quad +\frac{\lambda_{11}s^3}{3!\,\lambda_1^{3/2}}+\frac{\lambda_{11}s^2t}{2\,\lambda_1\sqrt{\lambda_2}}+\frac{\lambda_{11}st^2}{2\sqrt{\lambda_1}\lambda_2}+\frac{\lambda_{11}t^3}{3!\,\lambda_2^{3/2}}\\
&\qquad +\frac{\lambda_{11}s^4}{4!\,\lambda_1^2}+\frac{\lambda_{11}s^3t}{3!\,\lambda_1^{3/2}\sqrt{\lambda_2}}+\frac{6\lambda_{11}s^2t^2}{4!\,\lambda_1\lambda_2}+\frac{\lambda_{11}st^3}{3!\sqrt{\lambda_1}\lambda_2^{3/2}}+\frac{\lambda_{11}t^4}{4!\,\lambda_2^2}+\cdots.
\end{aligned}$$

Sumamos los desarrollos encontrados, agrupamos términos y factorizamos, con lo que obtenemos

$$\begin{aligned}
A+B+C &= \frac{(\lambda_{10}+\lambda_{11})}{\sqrt{\lambda_1}}\,s+\frac{(\lambda_{01}+\lambda_{11})}{\sqrt{\lambda_2}}\,t+\frac{(\lambda_{10}+\lambda_{11})}{2\,\lambda_1}\,s^2+\frac{(\lambda_{01}+\lambda_{11})}{2\,\lambda_2}\,t^2\\
&\quad +\frac{\lambda_{11}st}{\sqrt{\lambda_1\lambda_2}}+\frac{(\lambda_{10}+\lambda_{11})}{3!\,\lambda_1^{3/2}}\,s^3+\frac{(\lambda_{01}+\lambda_{11})}{3!\,\lambda_2^{3/2}}\,t^3+\frac{\lambda_{11}s^2t}{2\,\lambda_1\sqrt{\lambda_2}}+\frac{\lambda_{11}st^2}{2\sqrt{\lambda_1}\lambda_2}\\
&\qquad +\frac{(\lambda_{10}+\lambda_{11})}{4!\,\lambda_1^2}\,s^4+\frac{(\lambda_{01}+\lambda_{11})}{4!\,\lambda_2^2}\,t^4+\frac{\lambda_{11}s^3t}{3!\,\lambda_1^{3/2}\sqrt{\lambda_2}}+\frac{6\lambda_{11}s^2t^2}{4!\,\lambda_1\lambda_2}+\frac{\lambda_{11}st^3}{3!\sqrt{\lambda_1}\lambda_2^{3/2}}+\cdots.
\end{aligned}$$

Recordando las definiciones de λ_1 y λ_2 llegamos a que

$$\begin{aligned}
A+B+C &= s\sqrt{\lambda_1}+t\sqrt{\lambda_2}+\frac{s^2}{2}+\frac{t^2}{2}+\frac{s^3}{3!\sqrt{\lambda_1}}+\frac{t^3}{3!\sqrt{\lambda_2}}+\frac{s^4}{4!\,\lambda_1}+\frac{t^4}{4!\,\lambda_2}+\frac{\lambda_{11}st}{\sqrt{\lambda_1\lambda_2}}\\
&\qquad +\frac{\lambda_{11}s^2t}{2\,\lambda_1\sqrt{\lambda_2}}+\frac{\lambda_{11}st^2}{2\sqrt{\lambda_1}\lambda_2}+\frac{\lambda_{11}s^3t}{3!\,\lambda_1^{3/2}\sqrt{\lambda_2}}+\frac{6\lambda_{11}s^2t^2}{4!\,\lambda_1\lambda_2}+\frac{\lambda_{11}st^3}{3!\sqrt{\lambda_1}\lambda_2^{3/2}}+\cdots.
\end{aligned}$$

Sustituyendo en (2.50) obtenemos un desarrollo para la $f.g.m.$, a saber,

$$\begin{aligned}(2.51)\quad \psi_{W_1,W_2}(s,t) = \exp\Bigg(&\frac{s^2}{2}+\frac{t^2}{2}+\frac{s^3}{3!\sqrt{\lambda_1}}+\frac{t^3}{3!\sqrt{\lambda_2}}+\frac{s^4}{4!\,\lambda_1}+\frac{t^4}{4!\,\lambda_2}+\frac{\lambda_{11}st}{\sqrt{\lambda_1\lambda_2}}+\frac{\lambda_{11}s^2t}{2\,\lambda_1\sqrt{\lambda_2}}\\ &+\frac{\lambda_{11}st^2}{2\sqrt{\lambda_1}\lambda_2}+\frac{\lambda_{11}s^3t}{3!\,\lambda_1^{3/2}\sqrt{\lambda_2}}+\frac{6\lambda_{11}s^2t^2}{4!\,\lambda_1\lambda_2}+\frac{\lambda_{11}st^3}{3!\sqrt{\lambda_1}\lambda_2^{3/2}}+\cdots\Bigg).\end{aligned}$$

Ahora, tomamos el límite de la $f.g.m.$ dada por (2.51) cuando $\lambda_{10}\to\infty$, $\lambda_{01}\to\infty$ y $\lambda_{11}\to\infty$ de tal suerte que se cumpla (2.47), luego también $\lambda_1\to\infty$ y $\lambda_2\to\infty$. Por (2.48) sabemos (vea la demostración de la Proposición 2.25) que

$$\frac{\lambda_{11}}{\sqrt{\lambda_1\lambda_2}}\to\frac{1}{2}.$$

Por otro lado, utilizando (2.47), tenemos que

$$\frac{\lambda_{11}}{\lambda_1\sqrt{\lambda_2}}=\frac{\frac{\lambda_{11}}{\lambda}}{\frac{\lambda_1}{\lambda}\sqrt{\lambda_2}}\to\frac{1}{2}\cdot 0=0$$

y, similarmente,

$$\frac{\lambda_{11}}{\sqrt{\lambda_1}\lambda_2}\to\frac{1}{2}\cdot 0=0,\qquad \frac{\lambda_{11}}{\lambda_1\lambda_2}\to\frac{1}{2}\cdot 0=0,\qquad \frac{\lambda_{11}}{\lambda_1^{3/2}\lambda_2}\to\frac{1}{2}\cdot 0=0\quad\text{y}\quad \frac{\lambda_{11}}{\lambda_1\lambda_2^{3/2}}=\frac{1}{2}\cdot 0=0.$$

Se concluye pues que casi todos los términos en (2.51) convergen a cero. Por lo tanto,

$$\lim_{(\lambda_{10},\lambda_{01},\lambda_{11})\to\infty}\psi_{W_1,W_2}(s,t)=\exp\left[\frac{1}{2}(s^2+st+t^2)\right],$$

donde esta última función es la $f.g.m.$ para la distribución normal bivariada con media $\vec{\mu}=\vec{0}$ y matriz de covarianzas $\begin{bmatrix}1 & 1/2\\ 1/2 & 1\end{bmatrix}$, es decir, cuando $\lambda_{10}\to\infty$, $\lambda_{01}\to\infty$ y $\lambda_{11}\to\infty$ de tal suerte que se cumpla (2.47), se obtiene

$$(W_1\ ,\ W_2)\ \xrightarrow{\quad d\quad}\ N_2\left(\vec{0}\ ,\ \begin{bmatrix}1 & 1/2\\ 1/2 & 1\end{bmatrix}\right).\ \blacksquare$$

2.3.3. Ejemplos

Los siguientes ejemplos ilustran la forma en que se pueden construir situaciones que se ajusten con el modelo de la distribución PB dada por (2.36). Se trata de una aplicación del Teorema 2.20 donde habrá que considerar tres variables aleatorias independientes cuya distribución sea de Poisson, digamos X, Y y Z, entre las cuales debe haber algún atributo en común de modo que la suma, por ejemplo $X+Z$, describa aquel atributo.

Ejemplo 2.28 En un banco se distinguen a los clientes en dos "tipos": los clientes con cuenta en el banco y los clientes sin cuenta. El banco ofrece diversas operaciones que serán clasificadas en dos: depósito a cuentas y "otras transacciones". Los clientes sin cuenta sólo pueden realizar depósitos. Se dispone de la siguiente información:

1. Los clientes con cuenta y que realizan "otras transacciones" llegan de acuerdo a una distribución de Poisson con media $\lambda_{10} = 20$ clientes por hora.

2. Los clientes sin cuenta llegan de acuerdo a una distribución de Poisson con media $\lambda_{01} = 15$ clientes por hora.

3. Los clientes con cuenta y que realizan un depósito llegan de acuerdo a una distribución de Poisson con media $\lambda_{11} = 25$ clientes por hora.

Sean X, Y y Z v.as. tales que: X es el número de llegadas de clientes con cuenta y que realizan "otras transacciones", Y el número de llegadas de clientes sin cuenta y Z el número de llegadas de clientes con cuenta y que realizan un depósito, todas medidas durante un periodo de una hora. Entonces

$$X \sim Poisson(\lambda_{10}), \quad Y \sim Poisson(\lambda_{01}), \quad \text{y} \quad Z \sim Poisson(\lambda_{11}).$$

Suponga que X, Y y Z son independientes. Si U y V son v.as. que miden el número de llegadas, en un periodo de una hora, de clientes con cuenta y de clientes que realizan un depósito, respectivamente, entonces $U = X + Z$, $V = Y + Z$ y (U, V) tiene la distribución PB.

Con la información disponible es posible obtener el coeficiente de correlación entre U y V por medio de (2.41), se tiene que $Corr(U, V) = 0.5893$, es decir, existe una dependencia considerable entre U y V, además se cumple $Corr(U, V) > 0$ tal como se mencionó luego de la Proposición 2.23. ◀

Ejemplo 2.29 Suponga que las personas que salen de la ciudad durante el periodo vacacional de fin de año se clasifican de acuerdo a su sexo y que sus motivos se clasifican en dos tipos: por placer o compromiso. Se está particularmente interesado en el número de mujeres que salen de la ciudad y el número de personas que salen por placer durante un periodo vacacional de fin de año cualquiera, de manera que todas las variables aleatorias que se indiquen serán medidas durante un periodo vacacional de fin de año.

Sean X el número de mujeres que salen por compromiso, Y el número de hombres que salen por placer y Z el número de mujeres que salen por placer. Suponga que X, Y y Z son independientes. Si $U = X + Z$ y $V = Y + Z$, es decir, U cuenta el número de mujeres que salen de la ciudad y V el número de personas que salen por placer, entonces (U, V) es un vector aleatorio con la distribución PB dada por (2.36). ◀

Capítulo 3

Distribuciones hipergeométrica, geométrica y binomial negativa bivariadas

3.1. Distribución hipergeométrica bivariada

3.1.1. La distribución hipergeométrica bivariada

Definición 3.1 *Se dice que un vector aleatorio* (U, V) *tiene distribución* ***hipergeométrica bivariada*** *(HGB) si: para* $N, N_{10}, N_{01}, N_{11}, N_{00}, n, a \in \mathbb{N}$, $1 \leq n \leq N$, $N_{10} + N_{01} + N_{11} + N_{00} = N$, $a = 0, 1, \ldots, \text{mín}(u, v)$ *y* $u - a = 0, 1, \ldots, \text{mín}(n, N_{10})$, $v - a = 0, 1, \ldots, \text{mín}(n, N_{01})$, *se tiene que*

$$P[U = u, V = v] = \frac{\displaystyle\sum_{a=0}^{\text{mín}(u,v)} \binom{N_{10}}{u-a}\binom{N_{01}}{v-a}\binom{N_{11}}{a}\binom{N_{00}}{n-u-v+a}}{\displaystyle\binom{N}{n}}. \tag{3.1}$$

En el caso univariado la distribución hipergeométrica se describe mediante un experimento similar al de la binomial, la diferencia es el tipo de muestreo aleatorio que se realiza, el cual es, en este caso (aquí el contraste) **sin reemplazo**. El caso bivariado es similar, para obtener la distribución HGB dada por (3.1) vamos a considerar el siguiente esquema de muestreo.

Consideremos una población de tamaño N en la cual N_{ij} elementos de la población poseen una característica representada por la pareja (i, j), $i, j \in \{0, 1\}$, donde la primera componente de la pareja indica la presencia o no de un atributo, digamos A, mientras que la segunda componente indica la presencia o no de otro atributo B.

Podemos pensar en lo anterior como si se tratara de una urna con N bolas, numeradas $1, \ldots, N$, cada una de las cuales tiene asociada una pareja (i, j), donde, por ejemplo, $i = 1$ representa la presencia del atributo "blanca" y $j = 1$ representa la presencia del atributo "par", así una bola que tenga asociada la pareja $(1, 1)$ sería blanca y par.

Si se realiza un muestreo aleatorio **sin reemplazo** de tamaño n y, U y V son el número de elementos de la muestra que poseen el atributo A en la primera componente y el atributo B en la segunda componente, respectivamente, se probará que entonces (U, V) tiene la distribución (3.1). Note que si el muestreo se realiza con reemplazo, entonces (U, V) debe tener distribución binomial bivariada (vea la Sección 2.2).

Para verificar lo anterior utilizamos el método del punto muestral. Un punto muestral en Ω corresponderá a una selección única de n elementos de la población, donde los elementos de la población están numerados por $1, \ldots, N$ de tal manera que se pueden distinguir unos de otros. Recuerde que cada punto muestral tiene asociada una pareja (i, j), $i, j \in \{0, 1\}$. Así pues al seleccionar la muestra se obtendría un arreglo de tamaño n de parejas (i, j).

Se muestra a continuación un punto muestral correspondiente a las observaciones registradas en la segunda componente de un experimento en particular. Este punto típicamente se puede ver de la siguiente forma:

$$\underbrace{4, 8, 95, 1, 17, 7, 12, 35, 46, \ldots, 25, 3}_{n \text{ posiciones}}.$$

En donde se indica que la primera observación corresponde a una bola (blanca o negra) par (la bola marcada con el número 4), enseguida viene una bola par (la número 8), luego una bola impar (número 95), y así sucesivamente hasta indicar las últimas dos observaciones que corresponden a bolas impares (números 25 y 3). El arreglo de tamaño n de parejas (i, j) correspondiente a esta muestra sería

$$\underbrace{(i, 1), (i, 1), (i, 0), (i, 0), (i, 0), (i, 0), (i, 1), \ldots, (i, 0), (i, 0)}_{n \text{ posiciones}}.$$

en donde $i = 1$ si la bola seleccionada es blanca e $i = 0$ si es negra.

Así que el número total de elementos en Ω es exactamente el número de formas en las que se puede seleccionar un subconjunto de n elementos de la población, cuyo tamaño es N, es decir $\binom{N}{n}$.

Vamos a describir la población mediante la representación de la Tabla 3.1 en la que se especifica el número de elementos y sus atributos. Como ya se mencionó, una muestra aleatoria es extraída sin reemplazo de la población y las observaciones son registradas en un arreglo de tamaño n (el de la muestra) las cuales pueden ser representadas como en la Tabla 3.2, donde a es el número de ocasiones en que se observó la característica $(1, 1)$, así mismo $u - a$ corresponde a $(1, 0)$, $v - a$ para $(0, 1)$ y el resto $n - u - v + a$ para $(0, 0)$.

Nos preguntamos por la probabilidad de observar el atributo A en $U = u$ ocasiones y B en $V = v$ ocasiones tal como se indica en la Tabla 3.2.

		Primera Componente		
		A	A^c	
Segunda	B	N_{11}	N_{01}	N_{+1}
Componente	B^c	N_{10}	N_{00}	N_{+0}
		N_{1+}	N_{0+}	N

Tabla 3.1: N es el número de elementos en la población, N_{+1} es el número de elementos que poseen el atributo B, N_{+0} el número de elementos que *no* poseen el atributo B, N_{1+} el número de elementos que poseen el atributo A y N_{0+} los elementos que *no* poseen el atributo A.

		Primera Componente		
		A	A^c	
Segunda	B	a	$v-a$	v
Componente	B^c	$u-a$	$n-u-v+a$	$n-v$
		u	$n-u$	n

Tabla 3.2: Observaciones registradas en un muestreo de tamaño n. Se observa el atributo A en u ocasiones y B en v ocasiones.

Vamos a introducir las siguientes variables aleatorias que nos permitirán describir más precisamente el esquema:

X_{11} : número de elementos que poseen los atributos A y B,
X_{10} : número de elementos que poseen el atributo A, pero *no* el atributo B,
X_{01} : número de elementos que poseen el atributo B, pero *no* el atributo A,
X_{00} : número de elementos que *no* poseen los atributos A y B.

Con estas variables aleatorias el esquema de muestreo queda descrito como se muestra en la Tabla 3.3. Se tiene entonces que $X_{11}+X_{10}+X_{01}+X_{00}=n$ y la probabilidad de que $X_{11}=x_{11}$, $X_{10}=x_{10}$, $X_{01}=x_{01}$ y $X_{00}=x_{00}$, dado el esquema de muestreo descrito, debe ser

$$P[X_{11}=x_{11},X_{10}=x_{10},X_{01}=x_{01},X_{00}=x_{00}]=\frac{\binom{N_{11}}{x_{11}}\binom{N_{10}}{x_{10}}\binom{N_{01}}{x_{01}}\binom{N_{00}}{x_{00}}}{\binom{N}{n}}, \tag{3.2}$$

donde

$$\begin{aligned} x_{11} &= 0,\dots,\min(n,N_{11}) \\ x_{10} &= 0,\dots,\min(n,N_{10}) \\ x_{01} &= 0,\dots,\min(n,N_{01}) \\ x_{00} &= 0,\dots,\min(n,N_{00}) \\ x_{11}&+x_{10}+x_{01}+x_{00}=n. \end{aligned}$$

Lo anterior se puede apreciar ya que en el correspondiente arreglo de tamaño n, el número total de puntos muestrales para la ocurrencia del evento $(X_{11}X_{10},X_{01},X_{00})=(x_{11},x_{10},x_{01},x_{00})$ es obtenido de la regla del producto al formar todas las combinaciones posibles. En este caso, el número de formas de

		Primera Componente		
		A	A^c	
Segunda	B	X_{11}	X_{01}	V
Componente	B^c	X_{01}	X_{00}	$n-V$
		U	$n-U$	n

Tabla 3.3: Definición de las variables aleatorias correspondientes al esquema de muestreo.

seleccionar x_{11} elementos de un total de N_{11}, para ocupar x_{11} posiciones en el arreglo es $\binom{N_{11}}{x_{11}}$, el mismo significado tienen los factores restantes en el numerador de (3.2). Para formar todos los arreglos de tamaño n posibles, requerimos calcular el número de formas de combinar un conjunto de x_{11} elementos, con otro de x_{10} elementos, con uno más de x_{01} elementos y finalmente con un conjunto de x_{00} elementos, que de acuerdo a la regla del producto es exactamente el numerador de (3.2). Así queda justificada dicha probabilidad.

El esquema de muestreo que nos condujo a la probabilidad (3.2) nos permite deducir la función de distribución conjunta de las variables aleatorias U y V (vea la Tabla 3.2 y la definición de las variables aleatorias en la Tabla 3.3). Se debe tener

$$P[U=u, V=v] = \sum_{a=0}^{\text{mín}(u,v)} P[X_{11}=a, X_{10}=u-a, X_{01}=v-a, X_{00}=n-u-v+a]. \tag{3.3}$$

Así, sustituyendo (3.2) en (3.3) obtenemos (3.1).

3.1.2. La distribución multi-hipergeométrica

Otra distribución a veces llamada **distribución hipergeométrica multivariada**, que podría ser llamada más precisamente como **distribución multi-hipergeométrica**, es obtenida de la siguiente forma (una situación similar como cuando hablamos de la distribución multinomial).

Consideremos una población de tamaño N que contiene cuatro tipos de características, digamos I, II, III y IV de tamaños N_1, N_2, N_3 y N_4, respectivamente. Realizamos un muestreo aleatorio **sin reemplazo** de tamaño n y, para $i = I, II, III, IV$, definimos

$$X_i = \text{número de individuos con la característica } i \text{ en la muestra.}$$

Entonces, como $X_1 + X_2 + X_3 + X_4 = n$, podemos obtener la probabilidad de que x_1 individuos tengan la característica I, x_2 la II y x_3 la III, el resto de los individuos en la muestran presentan la característica IV, como

$$P[X_1=x_1, X_2=x_2, X_3=x_3] = \frac{\binom{N_1}{x_1}\binom{N_2}{x_2}\binom{N_3}{x_3}\binom{N_4}{n-x_1-x_2-x_3}}{\binom{N}{n}}, \tag{3.4}$$

donde $0 \leq x_1 + x_2 + x_3 \leq n$ y

$$\begin{aligned} x_1 =& 0, \ldots, \text{mín}(n, N_1), \\ x_2 =& 0, \ldots, \text{mín}(n, N_2), \\ x_3 =& 0, \ldots, \text{mín}(n, N_3). \end{aligned}$$

A este modelo se le puede considerar como una generalización trivariada de la distribución hipergeométrica. En el Capítulo 6 de [8] se trata un modelo correspondiente a una generalización bivariada de la distribución hipergeométrica. Aquí trabajamos con (3.4) porque la población sobre la que estamos trabajando es tal que puede ser dividida en cuatro grupos con características particulares.

Vamos a ver que la distribución HGB dada por (3.1) se puede derivar de la distribución multi-hipergeométrica dada por (3.4). En efecto, pensemos que la población está formada por individuos cuyas características se distinguen por la pareja (i,j) con $i,j \in \{0,1\}$, como antes, suponiendo que la primera componente representa la presencia o no de un atributo A y que la segunda representa la presencia o no de un atributo B. Para la derivación que nos ocupa en este momento vamos a establecer las siguientes definiciones. Sean X_1, X_2, X_3 y X_4 el número de individuos con los atributos $A \cap B, A^c \cap B, A \cap B^c$ y $A^c \cap B^c$, respectivamente. Entonces, si U y V son el número de elementos que poseen el atributo A en la primera componente y el atributo B en la segunda componente, respectivamente, se obtiene que

$$U = X_1 + X_3 \qquad \text{y} \qquad V = X_1 + X_2. \tag{3.5}$$

Nos preguntamos por la probabilidad de hallar el atributo A en u ocasiones y B en v ocasiones. Obtenemos, bajo la condición $x_1 + x_3 = u$ y $x_1 + x_2 = v$, que

$$P[U = u, V = v] = \sum_{x_1}\sum_{x_2}\sum_{x_3} P[X_1 = x_1, X_2 = x_2, X_3 = x_3] \tag{3.6}$$

Lo que indica exactamente todas las formas posibles de que en la muestra se observen los atributos $A \cap B, A^c \cap B$ y $A \cap B^c$, en x_1, x_2 y x_3 ocasiones, respectivamente, de forma tal que $x_1 + x_3 = u$ y $x_1 + x_2 = v$.

La identidad en (3.6) podemos mostrarla más precisamente si hacemos las siguientes elecciones, sean $N_{11} = N_1, N_{10} = N_3, N_{01} = N_2$ y $N_{00} = N_4$ que coincide con el esquema de muestreo descrito en la sección anterior. Por otro lado, puesto que

$$\begin{aligned} n - x_1 - x_2 - x_3 &= n - x_1 - x_2 - x_3 - x_1 + x_1 \\ &= n - (x_1 + x_2) - (x_3 + x_1) + x_1 \\ &= n - (x_1 + x_3) - (x_1 + x_2) + x_1 \\ &= n - u - v + x_1, \end{aligned}$$

vemos entonces que

$$P[X_1 = x_1, X_2 = v - x_1, X_3 = u - x_1] = \frac{\binom{N_{11}}{x_1}\binom{N_{01}}{v - x_1}\binom{N_{10}}{u - x_1}\binom{N_{00}}{n - u - v + x_1}}{\binom{N}{n}}.$$

Así podemos reescribir (3.6) como

$$P[U = u, V = v] = \sum_{x_1} P[X_1 = x_1, X_2 = v - x_1, X_3 = u - x_1]$$

que es, ahora explícitamente, tomando $a = x_1$, la distribución HGB (3.1).

Ahora podemos hacer distinciones entre (3.1) y (3.4). Como se mencionó en la sección anterior, para obtener la distribución HGB pensamos en el ejemplo de una urna cuyos elementos son bolas blancas y negras marcadas con un número par o impar, el modelo surge del estudio de una población que podemos separar en dos grupos distintivos de acuerdo a sus atributos (el de las bolas blancas y el de las negras) los cuales a su vez pueden ser divididos en otros dos grupos (bola par o impar), por lo que se obtienen cuatro subgrupos, sin embargo nos ocupamos de investigar la ocurrencia de los atributos como si dividiéramos la población en cuatro grupos (bolas blancas, bolas negras, bolas pares, bolas impares) investigando exactamente la ocurrencia de dos de esos atributos en un cierto número de ocasiones; en particular y para ilustrar los hechos, investigamos la ocurrencia de observar bola blanca en u ocasiones y par en v ocasiones. En cambio, para obtener (3.4), dividimos la población en cuatro grupos distintivos (bola blanca y par, bola blanca e impar, bola negra y par, bola negra e impar), que coinciden con los cuatro subgrupos mencionados arriba, para investigar la ocurrencia de cada uno de los elementos de estos grupos en un cierto número de ocasiones bajo un muestreo sin reemplazo.

Entonces, (3.1) es un modelo que interesa para poblaciones en las que se pueden observar exactamente **cuatro atributos** y en donde estemos investigando la ocurrencia de *dos de esos atributos* en un cierto número de ocasiones. Por otro lado, (3.4) es un modelo de interés para poblaciones que se pueden dividir en *cuatro grupos distintivos* y nos ocupamos en investigar la ocurrencia de *cada una de esas características distintivas* en un muestreo. Cuando esos cuatro grupos distintivos coinciden con los atributos en los que estemos particularmente interesados, y nos ocupemos de investigar la ocurrencia de dos de aquellos atributos, podemos utilizar cualquiera de estos modelos. Esto último es exactamente la relación entre (3.1) y (3.6).

3.1.3. Valor esperado condicional

Sea (U, V) un vector aleatorio con distribución HGB dada por (3.1). Vamos a derivar los valores esperados condicionales $E[U \mid V = v]$ y $E[V \mid U = u]$, para ello requeriremos de las distribuciones marginales para U y V que son obtenidas directamente de las Tablas 3.1 y 3.2.

Pensemos pues en el mismo experimento descrito anteriormente y tratemos únicamente con la variable V que representa el número de veces que se observa el atributo B (la segunda componente). Recordando que el experimento fue sin reemplazo, tenemos que la distribución marginal de V está dada por

$$P[V = v] = \frac{\binom{N_{+1}}{v}\binom{N_{+0}}{n-v}}{\binom{N}{n}}, \tag{3.7}$$

donde $v = 0, \ldots, \text{mín}(n, N+1)$ y $N_{+1} + N_{+0} = N$. Las Tablas 3.1 y 3.2 muestran el número de elementos en la población que poseen el atributo B y el número de observaciones, respectivamente. De manera completamente similar obtenemos la siguiente distribución marginal de U

$$P[U = u] = \frac{\binom{N_{1+}}{u}\binom{N_{0+}}{n-u}}{\binom{N}{n}}, \tag{3.8}$$

donde $u = 0, \ldots, \text{mín}(n, N_{1+})$ y $N_{1+} + N_{0+} = N$. Observe en particular que las distribuciones marginales de U y V son ambas hipergeométricas (Definición 1.72).

Proposición 3.2 *Si* (U, V) *es un vector aleatorio con distribución* HGB *dada por* (3.1)*, entonces las distribuciones condicionales de* $V = v$ *dado que* $U = u$ *y de* $U = u$ *dado que* $V = v$ *están dadas,* *respectivamente, por*

$$P[V=v\,|\,U=u]=\sum_a \frac{\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}}, \tag{3.9}$$

$$P[U=u\,|\,V=v]=\sum_a \frac{\binom{v}{a}\binom{N_{+1}-v}{N_{11}-a}\binom{N_{+0}-(n-v)}{N_{10}-(u-a)}\binom{n-v}{u-a}}{\binom{N_{+1}}{N_{11}}\binom{N_{+0}}{N_{10}}}. \tag{3.10}$$

Demostración. Vamos a obtener la distribución condicional de $V=v$ dado que $U=u$ directamente de (3.1) y (3.8) como sigue. Se tiene,

$$\begin{aligned}
P[V=v\,|\,U=u] &= \frac{P[V=v,U=u]}{P[U=u]}\\
&= \frac{\displaystyle\sum_a \frac{\binom{N_{10}}{u-a}\binom{N_{01}}{v-a}\binom{N_{11}}{a}\binom{N_{00}}{n-u-v+a}}{\binom{N}{n}}}{\dfrac{\binom{N_{1+}}{u}\binom{N_{0+}}{n-u}}{\binom{N}{n}}}\\
&= \sum_a \frac{\binom{N_{10}}{u-a}\binom{N_{01}}{v-a}\binom{N_{11}}{a}\binom{N_{00}}{n-u-v+a}}{\binom{N_{1+}}{u}\binom{N_{0+}}{n-u}},
\end{aligned}$$

donde

$$\begin{aligned}
&\frac{\binom{N_{10}}{u-a}\binom{N_{01}}{v-a}\binom{N_{11}}{a}\binom{N_{00}}{n-u-v+a}}{\binom{N_{1+}}{u}\binom{N_{0+}}{n-u}}\\
&= \frac{\frac{N_{10}!}{(u-a)!\,(N_{10}-(u-a))!}\;\frac{N_{01}!}{(v-a)!\,(N_{01}-(v-a))!}\;\frac{N_{11}!}{a!\,(N_{11}-a)!}\;\frac{N_{00}!}{(n-u-v+a)!\,(N_{00}-(n-u-v+a))!}}{\frac{N_{1+}!}{u!\,(N_{1+}-u)!}\;\frac{N_{0+}!}{(n-u)!\,(N_{0+}-(n-u))!}}\\
&= \frac{\frac{u!}{a!\,(u-a)!}\;\frac{(N_{1+}-u)!}{(N_{11}-a)!\,(N_{10}-(u-a))!}\;\frac{(N_{0+}-(n-u))!}{(N_{01}-(v-a))!\,(N_{00}-(n-u-v+a))!}\;\frac{(n-u)!}{(v-a)!\,(n-u-v+a)!}}{\frac{N_{1+}!}{N_{11}!\,N_{10}!}\;\frac{N_{0+}!}{N_{01}!\,N_{00}!}}\\
&= \frac{\frac{u!}{a!\,(u-a)!}\;\frac{(N_{1+}-u)!}{(N_{11}-a)!\,(N_{1+}-u-(N_{11}-a))!}\;\frac{(N_{0+}-(n-u))!}{(N_{01}-(v-a))!\,(N_{0+}-(n-u)-(N_{01}-(v-a)))!}\;\frac{(n-u)!}{(v-a)!\,(n-u-(v-a))!}}{\frac{N_{1+}!}{N_{11}!\,(N_{1+}-N_{11})!}\;\frac{N_{0+}!}{N_{01}!\,(N_{0+}-N_{01})!}}\\
&= \frac{\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}},
\end{aligned}$$

de manera que

$$P[V=v\,|\,U=u]=\sum_a \frac{\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}}.$$

Intercambiando los papeles de U y V se demuestra similarmente (3.10). ∎

Observación 3.3 Puesto que (3.9) y (3.10) son distribuciones de probabilidad discretas, al variar v y u, respectivamente, se cumplen las siguientes identidades

$$\sum_v P[V=v\,|\,U=u]=1, \qquad \text{y} \qquad \sum_u P[U=u\,|\,V=v]=1,$$

es decir,

$$\sum_v \frac{P[U=u,V=v]}{P[U=u]}=1 \quad \text{y} \quad \sum_u \frac{P[U=u,V=v]}{P[V=v]}=1,$$

de manera que

$$\sum_v\sum_a \frac{\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}}=1,$$

$$\sum_u\sum_a \frac{\binom{v}{a}\binom{N_{+1}-v}{N_{11}-a}\binom{N_{+0}-(n-v)}{N_{10}-(u-a)}\binom{n-v}{u-a}}{\binom{N_{+1}}{N_{11}}\binom{N_{+0}}{N_{10}}}=1.$$

En esta forma se tienen las identidades

$$\sum_v\sum_a \binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}=\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}},$$

$$\sum_u\sum_a \binom{v}{a}\binom{N_{+1}-v}{N_{11}-a}\binom{N_{+0}-(n-v)}{N_{10}-(u-a)}\binom{n-v}{u-a}=\binom{N_{+1}}{N_{11}}\binom{N_{+0}}{N_{10}},$$

que serán usadas en la derivación de los correspondientes valores esperados condicionales. ◁

Teorema 3.4 *Sea* (U,V) *un vector aleatorio con distribución hipergeométrica bivariada dada por* (3.1). *Entonces las esperanzas condicionales de* V *dado que* $U=u$ *y de* U *dado que* $V=v$ *están dadas, respectivamente, por*

$$E[V\mid U=u]=n\frac{N_{01}}{N_{0+}}+u\frac{N_{11}\,N_{00}-N_{10}\,N_{01}}{N_{1+}\,N_{0+}}, \tag{3.11}$$

$$E[U\mid V=v]=n\frac{N_{10}}{N_{+0}}+v\frac{N_{11}\,N_{00}-N_{10}\,N_{01}}{N_{+1}\,N_{+0}}. \tag{3.12}$$

Demostración. La derivación de los valores esperados que nos ocupan en esta parte se realizará utilizando las distribuciones condicionales (3.9) y (3.10). Por definición,

$$
\begin{aligned}
E[V \mid U = u] &= \sum_v vP[V = v \mid U = u] \\
&= \sum_v \sum_a \frac{v\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}},
\end{aligned}
$$

que podemos reescribir como

$$
\sum_{v,a} \frac{(v-a)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}} \tag{3.13}
$$

$$
+\sum_{v,a} \frac{a\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}}, \tag{3.14}
$$

donde

$$
\begin{aligned}
&(v-a)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a} \\
&\qquad = (v-a)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\left(\frac{(n-u)!}{(v-a)!(n-u-v+a)!}\right) \\
&\qquad = (n-u)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\left(\frac{(n-u-1)!}{(v-a-1)!(n-u-v+a)!}\right) \\
&\qquad = (n-u)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-1-(n-u-1)}{N_{01}-1-(v-a-1)}\binom{n-u-1}{v-a-1}.
\end{aligned}
$$

Pero por la Observación 3.3 tenemos que

$$
\sum_{v,a}(n-u)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-1-(n-u-1)}{N_{01}-1-(v-a-1)}\binom{n-u-1}{v-a-1} = (n-u)\binom{N_{1+}}{N_{11}}\binom{N_{0+}-1}{N_{01}-1}
$$

y, por lo tanto, (3.13) es

$$
\begin{aligned}
\sum_{v,a} \frac{(v-a)\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}} &= \frac{(n-u)\binom{N_{1+}}{N_{11}}\binom{N_{0+}-1}{N_{01}-1}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}} \\
&= (n-u)\frac{\dfrac{(N_{0+}-1)!}{(N_{01}-1)!(N_{0+}-N_{01})}}{\dfrac{N_{0+}!}{N_{01}!\,(N_{0+}-N_{01})!}} = (n-u)\frac{N_{0+}-1)!\,N_{01}!}{(N_{01}-1)!\,N_{0+}!} = (n-u)\frac{N_{01}}{N_{0+}}.
\end{aligned}
$$

Procedemos similarmente con (3.14). Se tiene,

$$\begin{aligned} a\binom{u}{a}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a} \\ &= a\binom{u!}{a!(u-a)!}\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a} \\ &= u\left(\frac{(u-1)!}{(a-1)!(u-a)!}\right)\binom{N_{1+}-u}{N_{11}-a}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a} \\ &= u\binom{u-1}{a-1}\binom{N_{1+}-1-(u-1)}{N_{11}-1-(a-1)}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}, \end{aligned}$$

luego, por la Observación 3.3, resulta

$$\sum_{v,a} u\binom{u-1}{a-1}\binom{N_{1+}-1-(u-1)}{N_{11}-1-(a-1)}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a} = u\binom{N_{1+}-1}{N_{11}-1}\binom{N_{0+}}{N_{01}}$$

y, por lo tanto,

$$\begin{aligned} \sum_{v,a}\frac{u\binom{u-1}{a-1}\binom{N_{1+}-1-(u-1)}{N_{11}-1-(a-1)}\binom{N_{0+}-(n-u)}{N_{01}-(v-a)}\binom{n-u}{v-a}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}} &= \frac{u\binom{N_{1+}-1}{N_{11}-1}\binom{N_{0+}}{N_{01}}}{\binom{N_{1+}}{N_{11}}\binom{N_{0+}}{N_{01}}} \\ &= u\frac{\frac{(N_{1+}-1)!}{(N_{11}-1)!(N_{1+}-N_{11})!}}{\frac{N_{1+}!}{N_{11}!\,(N_{1+}-N_{11})!}} = u\frac{(N_{1+}-1)!\,N_{11}!}{(N_{11}-1)!\,N_{1+}!} = u\frac{N_{11}}{N_{1+}}. \end{aligned}$$

Así, la suma de (3.13) y (3.14), es

$$(n-u)\frac{N_{01}}{N_{0+}} + u\frac{N_{11}}{N_{1+}} = n\frac{N_{01}}{N_{0+}} + u\left(\frac{N_{11}}{N_{1+}} - \frac{N_{01}}{N_{0+}}\right) = n\frac{N_{01}}{N_{0+}} + u\frac{N_{11}\,N_{00} - N_{10}\,N_{01}}{N_{1+}\,N_{0+}},$$

con lo que hemos obtenido (3.11).

Intercambiando los papeles de U y V se demuestra de manera similar (3.12). ∎

La correlación de U con V es obtenida como la media geométrica de los coeficientes de u y v en (3.11) y (3.12) como lo indica la siguiente proposición.

Proposición 3.5 *Si (U, V) es un vector aleatorio con distribución HGB dada por (3.1) entonces*

$$Corr(U,V) = \frac{N_{11}\,N_{00} \;-\; N_{10}\,N_{01}}{\sqrt{N_{+1}\,N_{+0}\,N_{1+}\,N_{0+}}}. \tag{3.15}$$

Demostración. Ya sabemos que U y V tienen distribución marginal hipergeométrica, luego sus medias y varianzas son finitas. Denotamos tales cantidades para U y V como μ_1, μ_2 y σ_1^2, σ_2^2, respectivamente, y sea $Corr(U,V) = \rho_{u,v}$. De (3.12) y (3.11) podemos decir que $E[U \mid V]$ es lineal como función de V y $E[V \mid U]$ es lineal como función de U. Entonces, por el Teorema 1.46, se tiene que

$$\begin{aligned} E[U \mid V=v] &= \mu_1 + \rho_{u,v}\,\frac{\sigma_1}{\sigma_2}(v-\mu_2) = \mu_1 - \rho_{u,v}\,\frac{\sigma_1}{\sigma_2}\mu_2 + \rho_{u,v}\,\frac{\sigma_1}{\sigma_2}\,v \\ E[V \mid U=u] &= \mu_2 + \rho_{u,v}\,\frac{\sigma_2}{\sigma_1}(u-\mu_1) = \mu_2 - \rho_{u,v}\,\frac{\sigma_2}{\sigma_1}\mu_1 + \rho_{u,v}\frac{\sigma_2}{\sigma_1}\,u, \end{aligned}$$

de donde los coeficientes de u y v son $\rho_{u,v}\,\frac{\sigma_2}{\sigma_1}$ y $\rho_{u,v}\,\frac{\sigma_1}{\sigma_2}$, respectivamente. Note que la media geométrica de estos coeficientes es la correlación entre U y V.

Comparando las expresiones anteriores con (3.12) y (3.11) obtenemos

$$\rho_{u,v}\,\frac{\sigma_1}{\sigma_2} = \frac{N_{11}\,N_{00} - N_{10}\,N_{01}}{N_{+1}\,N_{+0}} \qquad \text{y} \qquad \rho_{u,v}\,\frac{\sigma_2}{\sigma_1} = \frac{N_{11}\,N_{00} - N_{10}\,N_{01}}{N_{1+}\,N_{0+}}$$

y, por lo tanto,

$$\rho_{u,v} = \frac{N_{11}\,N_{00} \;-\; N_{10}\,N_{01}}{\sqrt{N_{+1}\,N_{+0}\,N_{1+}\,N_{0+}}},$$

que es el resultado indicado. ∎

3.1.4. Convergencia a la distribución binomial bivariada

En el caso univariado es bien sabido que la distribución binomial aproxima a la distribución hipergeométrica para poblaciones grandes. En esta sección se demostrará que esta aproximación es también válida para el caso bivariado, más precisamente, que la distribución HGB dada por (3.1) converge a la distribución binomial bivariada dada por (2.5).

Teorema 3.6 *Sea (U, V) un vector aleatorio con distribución HGB dada por (3.1). Entonces, cuando $N \to \infty$ y $N_{ij} \to \infty$ de tal suerte que*

$$\frac{N_{ij}}{N} \quad\longrightarrow\quad p_{ij}, \quad i, j \in \{0, 1\},$$

se tiene que

$$P[U = u, V = v] \quad\longrightarrow\quad P[U^* = u^*, V^* = v^*],$$

donde el vector aleatorio (U^*, V^*) tiene distribución binomial bivariada dada por (2.5).

Demostración. Si $N \to \infty$ y $N_{ij} \to \infty$ de tal forma que

$$\frac{N_{ij}}{N} \quad\longrightarrow\quad p_{ij}, \quad i, j \in \{0, 1\},$$

claramente $p_{00} + p_{01} + p_{10} + p_{11} = 1$. La demostración estará pues terminada después de demostrar que, para (3.2), si $N \to \infty$ y $N_{ij} \to \infty$ como antes, entonces

$$\frac{\binom{N_{11}}{x_{11}}\binom{N_{10}}{x_{10}}\binom{N_{01}}{x_{01}}\binom{N_{00}}{x_{00}}}{\binom{N}{n}} \quad\longrightarrow\quad \binom{n}{x_{11}, x_{10}, x_{01}, x_{00}} p_{11}^{x_{11}}\,p_{10}^{x_{10}}\,p_{01}^{x_{01}}\,p_{00}^{x_{00}} \tag{3.16}$$

para cada $(x_{11}, x_{10}, x_{01}, x_{00})$ fijo y $x_{11} + x_{10} + x_{01} + x_{00} = n$. En efecto, digamos que el primer miembro de (3.16) es H, entonces

$$\begin{aligned} H &= \frac{N_{11}!}{(N_{11} - x_{11})!\,x_{11}!}\;\frac{N_{10}!}{(N_{10} - x_{10})!\,x_{10}!}\;\frac{N_{01}!}{(N_{01} - x_{01})!\,x_{01}!}\;\frac{N_{00}!}{(N_{00} - x_{00})!\,x_{00}!}\;\frac{n!\,(N-n)!}{N!} \\ &= \frac{n!}{x_{11}!\,x_{10}!\,x_{01}!\,x_{00}!} N_{11}(N_{11}-1)\cdots(N_{11}-x_{11}+1)N_{10}(N_{10}-1)\cdots(N_{10}-x_{10}+1) \\ &\qquad\qquad \times N_{01}(N_{01}-1)\cdots(N_{01}-x_{01}+1)N_{00}(N_{00}-1)\cdots(N_{00}-x_{00}+1)\,\frac{(N-n)!}{N!}, \end{aligned}$$

donde, para cualquier $i, j \in \{0,1\}$, se tiene que

$$N_{ij}(N_{ij}-1)\cdots(N_{ij}-x_{ij}+1) = N_{ij}^{x_{ij}}\left(1-\frac{1}{N_{ij}}\right)\cdots\left(1-\frac{x_{ij}-1}{N_{ij}}\right)$$

y

$$\frac{(N-n)!}{N!} = \frac{1}{N(N-1)\cdots(N-n+1)} = \frac{1}{N^n\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)}.$$

Al escribir $N^n = N^{x_{11}}N^{x_{10}}N^{x_{01}}N^{x_{00}}$ y haciendo tender $N \to \infty$ y $N_{ij} \to \infty$ como antes, resulta que

$$\begin{aligned}
H = \frac{n!}{x_{11}!\,x_{10}!\,x_{01}!\,x_{00}!} & N_{11}^{x_{11}}\left(1-\frac{1}{N_{11}}\right)\cdots\left(1-\frac{x_{11}-1}{N_{11}}\right)N_{10}^{x_{10}}\left(1-\frac{1}{N_{10}}\right)\cdots\left(1-\frac{x_{10}-1}{N_{10}}\right) \\
& \times N_{01}^{x_{01}}\left(1-\frac{1}{N_{01}}\right)\cdots\left(1-\frac{x_{01}-1}{N_{01}}\right)N_{00}^{x_{00}}\left(1-\frac{1}{N_{00}}\right)\cdots\left(1-\frac{x_{00}-1}{N_{00}}\right) \\
& \times \frac{1}{N^{x_{11}}N^{x_{10}}N^{x_{01}}N^{x_{00}}\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)} \\
= \binom{n}{x_{11}, x_{10}, x_{01}, x_{00}} & \left(\frac{N_{11}}{N}\right)^{x_{11}}\left(\frac{N_{10}}{N}\right)^{x_{10}}\left(\frac{N_{01}}{N}\right)^{x_{01}}\left(\frac{N_{00}}{N}\right)^{x_{00}}\left(1-\frac{1}{N_{11}}\right)\left(1-\frac{1}{N_{10}}\right)\left(1-\frac{1}{N_{01}}\right) \\
& \times \left(1-\frac{1}{N_{00}}\right)\cdots\left(1-\frac{x_{11}-1}{N_{11}}\right)\left(1-\frac{x_{10}-1}{N_{10}}\right)\left(1-\frac{x_{01}-1}{N_{01}}\right)\left(1-\frac{x_{00}-1}{N_{00}}\right) \\
& \times \frac{1}{\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)}
\end{aligned}$$

debe tender a

$$\begin{aligned}
& \binom{n}{x_{11}, x_{10}, x_{01}, x_{00}} p_{11}^{x_{11}}\,p_{10}^{x_{10}}\,p_{01}^{x_{01}}\,p_{00}^{x_{00}}\left(1-\frac{1}{N_{11}}\right)\left(1-\frac{1}{N_{10}}\right)\left(1-\frac{1}{N_{01}}\right)\left(1-\frac{1}{N_{00}}\right) \\
& \quad \times \cdots\left(1-\frac{x_{11}-1}{N_{11}}\right)\left(1-\frac{x_{10}-1}{N_{10}}\right)\left(1-\frac{x_{01}-1}{N_{01}}\right)\left(1-\frac{x_{00}-1}{N_{00}}\right)\frac{1}{\left(1-\frac{1}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)},
\end{aligned}$$

pues basta observar que, para todo $i, j \in \{0,1\}$,

$$\lim_{N_{ij}\to\infty}\left(1-\frac{\alpha}{N_{ij}}\right) = 1 \qquad \text{y} \qquad \lim_{N\to\infty}\left(1-\frac{\beta}{N}\right) = 1$$

con $1 \leq \alpha \leq x_{ij}-1$ y $1 \leq \beta \leq n-1$. Por lo tanto, se cumple (3.16).

Finalmente, tomando $x_{11} = a,\ x_{10} = u-a,\ x_{01} = v-a$ y $x_{00} = n-u-v+a$ en (3.16) y haciendo tender $N \to \infty$ y $N_{ij} \to \infty$ como antes, se obtiene

$$\frac{\binom{N_{11}}{a}\binom{N_{10}}{u-a}\binom{N_{01}}{v-a}\binom{N_{00}}{n-u-v+a}}{\binom{N}{n}} \longrightarrow \binom{n}{a, u-a, v-a, n-u-v+a} p_{11}^{a}\,p_{10}^{u-a}\,p_{01}^{v-a}\,p_{00}^{n-u-v+a}.$$

Como esta convergencia es independiente de $a = 0, \ldots, \text{mín}(u, v)$, al sumar con respecto a a y tomar límite, obtenemos

$$\sum_{a=0}^{\text{mín}(u,v)} \frac{\binom{N_{11}}{a}\binom{N_{10}}{u-a}\binom{N_{01}}{v-a}\binom{N_{00}}{n-u-v+a}}{\binom{N}{n}}$$

$$\longrightarrow \quad \sum_{a=0}^{\text{mín}(u,v)} \binom{n}{a, u-a, v-a, n-u-v+a} p_{11}^{a}\, p_{10}^{u-a}\, p_{01}^{v-a}\, p_{00}^{n-u-v+a},$$

es decir,

$$P[U = u, V = v] \quad \longrightarrow \quad P[U^* = u^*, V^* = v^*],$$

donde (U^*, V^*) tiene la distribución binomial bivariada dada por (2.5). ∎

3.1.5. Ejemplos

Ejemplo 3.7 Considere el estudio de los accidentes automovilísticos del Ejemplo 2.5. Suponga que la población fue de $N = 70$ casos, donde en $N_{11} = 31$ casos el niño sobrevivió y usaba cinturón; en $N_{10} = 22$ casos el niño sobrevivió y no usaba cinturón; en $N_{01} = 5$ casos el niño no sobrevivió y usaba cinturón; y en $N_{00} = 12$ casos el niño no sobrevivió y no usaba cinturón.

Se realizó un muestreo sin reemplazo de $n = 20$ casos. Si las variables aleatorias U y V cuentan el número de veces en que se observó que el niño sobrevivió y las veces en que se observó que utilizaba el cinturón, respectivamente, entonces el vector aleatorio (U, V) tiene distribución HGB dada por (3.1).

El coeficiente de correlación entre U y V es determinado mediante (3.15). Con la información del primer párrafo se obtuvo que $Corr(U, V) = 0.2495$, de modo que se tiene poca dependencia positiva entre U y V. ◄

Ejemplo 3.8 Aproximación binomial bivariada a la distribución hipergeométrica bivariada. En el Ejemplo 2.16 se obtuvo $P[U^* = 15, V^* = 10] = 3.65718965 \times 10^{-2}$ para una muestra de tamaño $n = 20$, donde el vector aleatorio (U^*, V^*) tiene la distribución binomial bivariada dada por (2.5). De acuerdo al Teorema 3.6, la cantidad anterior puede ser considerada como una aproximación de $P[U = 15, V = 10]$ cuando (U, V) tiene la distribución HGB para una población muy grande y una muestra del mismo tamaño.

En la Tabla 3.4 se muestran los resultados obtenidos para poblaciones muy grandes donde $p(15, 10) = P[U = 15, V = 10]$. Se puede apreciar que a mayor tamaño N de la población, menor es $p(15, 10)$. El lector puede comprobar además que, en todos los casos, $\frac{N_{ij}}{N} \approxeq p_{ij}$ (aproximadamente iguales) para cada $i, j \in \{0, 1\}$, vea la Tabla 2.1. Note que desde que el tamaño de la población es de 100,000 casos, se tiene que $p(15, 10)$ coincide con $P[U^* = 15, V^* = 10]$ en cuatro dígitos. En dicha tabla se puede apreciar que los valores de $p(15, 10)$ se van aproximando cada vez más al valor $P[U^* = 15, V^* = 10]$ cuando $N \to \infty$ y $N_{ij} \to \infty$ para todo $i, j \in \{0, 1\}$, confirmando numéricamente el resultado del Teorema 3.6. ◄

N	N_{11}	N_{10}	N_{01}	N_{00}	$p(15, 10)$
70	31	22	5	12	5.0559×10^{-2}
80	35	26	6	13	4.7827×10^{-2}
90	40	29	6	15	4.6333×10^{-2}
100	44	32	7	17	4.5433×10^{-2}
200	88	64	14	34	4.0524×10^{-2}
300	132	96	21	51	3.9115×10^{-2}
1000	440	320	70	170	3.7299×10^{-2}
10,000	4,400	3,200	700	1,700	3.6643×10^{-2}
100,000	44,000	32,000	7,000	17,000	3.6579×10^{-2}
1,000,000	440,000	320,000	70,000	170,000	$3.65726154 \times 10^{-2}$
10,000,000	4,400,000	3,200,000	700,000	1,700,000	$3.65719732 \times 10^{-2}$
100,000,000	44,000,000	32,000,000	7,000,000	17,000,000	$3.65719089 \times 10^{-2}$
1,000,000,000	440,000,000	320,000,000	70,000,000	170,000,000	$3.65719026 \times 10^{-2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	∞	∞	∞	∞	$3.65718965 \times 10^{-2}$

Tabla 3.4: $P[U = 15, V = 10]$ para poblaciones muy grandes, (U, V) tiene distribución HGB. Cuando $N \to \infty$ y $N_{ij} \to \infty$ para todo $i, j \in \{0, 1\}$ se tiene que $p(15, 10) \to P[U^* = 15, V^* = 10]$ donde (U^*, V^*) tiene distribución binomial bivariada.

3.2. Distribución geométrica bivariada

Consideremos una sucesión $\{(X_i, Y_i)\}_{i=1}^{\infty}$ de vectores aleatorios con distribución de Bernoulli bivariada. Recordemos, por definición, que las variables aleatorias de Bernoulli pueden tomar alguno de los dos valores cero o uno, donde podemos pensar que cuando se observa un uno en la variable aleatoria se trata de un "éxito" en el experimento. La generalización más natural de la distribución geométrica es que se estudie el número de pruebas hasta observar el primer éxito en la primer componente del vector (X_i, Y_i) y el primer éxito en la segunda componente. Definimos pues U y V como el número de ceros (fracasos) antes del primer uno (éxito) en la sucesión $\{X_i\}_{i=1}^{\infty}$ y en $\{Y_i\}_{i=1}^{\infty}$, respectivamente. Claramente, por la descripción indicada, cada una de las variables U y V tienen la distribución geométrica pero, en general, estas variables no serán independientes.

3.2.1. La distribución geométrica bivariada

Teorema y Definición 3.9 *Sea $\{(X_i, Y_i)\}_{i=1}^{\infty}$ una sucesión de vectores aleatorios independientes tal que (X_i, Y_i) tiene distribución de Bernoulli bivariada, sean U y V el número de pruebas antes de observar el* *primer éxito en la sucesión* $\{X_i\}_{i=1}^{\infty}$ *y en* $\{Y_i\}_{i=1}^{\infty}$*, respectivamente. Entonces*

$$P[U=u,V=v]=\begin{cases} p_{00}^{u}\,p_{10}\,p_{+0}^{v-u-1}\,p_{+1} & si \quad 0\le u<v,\\ p_{00}^{v}\,p_{01}\,p_{0+}^{u-v-1}\,p_{1+} & si \quad 0\le v<u,\\ p_{00}^{u}\,p_{11} & si \quad 0\le u=v, \end{cases} \tag{3.17}$$

equivalentemente,

$$P[U\ge u,V\ge v]=\begin{cases} p_{00}^{u}\,p_{+0}^{v-u} & si \quad 0\le u\le v,\\ p_{00}^{v}\,p_{0+}^{u-v} & si \quad 0\le v\le u. \end{cases} \tag{3.18}$$

Cualquier vector aleatorio (U,V) *con esta distribución se dice que tiene* **distribución geométrica bivariada** *(GB).*

Demostración. Se demostrará primeramente (3.17) considerando distintos posibles casos.

Caso $0\le u<v$. Para desglosar ideas supongamos primero que $u=0$. Esto significa que en la primer prueba se observó un éxito, es decir, $X_1=1$, otra forma de decirlo es que no se observó algún cero antes del primer éxito, el experimento concluye hasta que se observa $Y_{v+1}=1$, como $v>0$ entonces no se puede observar $Y_1=1$, de manera que la primer prueba corresponde exactamente al vector $(1,0)$ cuya probabilidad es p_{10}. Las siguientes pruebas pueden corresponder a cualquiera de los vectores $(1,0)$ y $(0,0)$ hasta en $v-1$ ocasiones (como ya se realizó una prueba antes, en este momento se han realizado v pruebas en las que se observó cero en la segunda componente), es decir, justo antes de observar un éxito en la segunda componente, lo que puede suceder observando $(0,1)$ o $(1,1)$. Lo que acaba de ser descrito puede quedar expresado en la siguiente forma

$$\begin{aligned} P[U=0,V=v] &= P[(X_1,Y_1)=(1,0),Y_2=0,\cdots,Y_v=0,Y_{v+1}=1]\\ &= P[(X_1,Y_1)=(1,0)]P[Y_2=0]\cdots P[Y_v=0]P[Y_{v+1}=1] \end{aligned}$$

que por independencia está justificada la segunda igualdad, luego entonces, recordando la distribución marginal de las variables aleatorias $Y_2,\ldots,Y_{v+1}$ dada en (2.1) se obtiene

$$P[U=0,V=v]=p_{10}\,p_{+0}^{v-1}\,p_{+1}.$$

Supongamos ahora que $0<u<v$, en la primera componente no se observa un 1 sino hasta después de u pruebas y como $u<v$ tampoco se observa 1 en la segunda componente hasta después de v pruebas, esto significa que las primeras u pruebas corresponden exactamente al vector $(0,0)$ y la $(u+1)$-ésima al vector $(1,0)$, la probabilidad hasta este momento es $p_{00}^{u}\,p_{10}$; se llevan $u+1$ observaciones. El experimento continúa observando ceros en la segunda componente hasta en $v-(u+1)$ ocasiones que corresponden a cualquiera de los vectores $(0,0)$ o $(1,0)$, es decir, con probabilidad $(p_{00}+p_{10})^{v-(u+1)}$ y finalmente se observa un éxito en la segunda componente, que corresponde a $(0,1)$ o $(1,1)$, cuya probabilidad es $(p_{01}+p_{11})$, así que, mediante la siguiente expresión

$$\begin{aligned} &P[U=u,V=v]\\ &\quad= P[(X_1,Y_1)=(0,0),\ldots,(X_u,Y_u)=(0,0),(X_{u+1},Y_{u+1})=(1,0),Y_{u+2}=0,\ldots,Y_v=0,Y_{v+1}=1]\\ &\quad= P[(X_1,Y_1)=(0,0)]\cdots P[(X_u,Y_u)=(0,0)]P[(X_{u+1},Y_{u+1})=(1,0)] \end{aligned}$$

$$\times P[Y_{u+2} = 0] \cdots P[Y_v = 0] P[Y_{v+1} = 1]$$

$$= \underbrace{p_{00} \cdots p_{00}}_{u \text{ veces}} \; p_{10} \; \underbrace{p_{+0} \cdots p_{+0}}_{v-(u+1) \text{ veces}} \; p_{+1},$$

hemos obtenido el primer caso de (3.17), es decir,

$$P[U = u, V = v] = p_{00}^{u}\, p_{10}\, p_{+0}^{v-u-1}\, p_{+1} \quad \text{si} \quad 0 \leq u < v.$$

Con argumentos similares a los anteriores se demuestra que en el caso $0 \leq v < u$ se cumple

$$P[U = u, V = v] = p_{00}^{v}\, p_{01}\, p_{0+}^{u-v-1}\, p_{1+}.$$

Caso $u = v$. Aquí se observa uno en ambas componentes a la vez. Si $u = v = 0$ entonces la primer prueba corresponde al vector $(1, 1)$. Por otro lado si $u > 0$ entonces $v > 0$, esto significa que en u ocasiones se observa el vector $(0, 0)$ con probabilidad p_{00}^{u} y justo después se observa $(1, 1)$. Haciendo uso de la independencia entre los vectores aleatorios con la distribución de Bernoulli bivariada subyacente, la probabilidad queda expresada de la siguiente forma

$$\begin{aligned} P[U = u, V = v] &= P[(X_1, Y_1) = (0, 0), \ldots, (X_u, Y_u) = (0, 0), (X_{u+1}, Y_{u+1}) = (1, 1)] \\ &= P[(X_1, Y_1) = (0, 0)] \cdots P[(X_u, Y_u) = (0, 0)] P[(X_{u+1}, Y_{u+1}) = (1, 1)] \\ &= \underbrace{p_{00} \cdots p_{00}}_{u \text{ veces}} \; p_{11} \end{aligned}$$

por lo que

$$P[U = u, V = v] = p_{00}^{u}\, p_{11} \quad \text{si} \quad 0 \leq u = v.$$

Esto completa la prueba de (3.17).

Comprobaremos (3.18) considerando también distintos casos.

Caso $0 \leq u \leq v$. Note que $P[U \geq u, V \geq v]$ es la probabilidad de observar u ceros en ambas componentes hasta la u-ésima prueba y observar $v - u$ ceros en la segunda componente entre la $(u + 1)$-ésima y la v-ésima pruebas, no siendo relevante lo que ocurra en la primer componente para estas últimas. De modo que, recordando (2.1), se obtiene

$$\begin{aligned} P[U \geq u, V \geq v] &= P[(X_1, Y_1) = (0, 0), \ldots, (X_u, Y_u) = (0, 0), Y_{u+1} = 0, \ldots, Y_v = 0] \\ &= P[(X_1, Y_1) = (0, 0)] \cdots P[(X_u, Y_u) = (0, 0)] P[Y_{u+1} = 0] \cdots P[Y_v = 0] \\ &= \underbrace{p_{00} \cdots p_{00}}_{u \text{ veces}} \; \underbrace{p_{+0} \cdots p_{+0}}_{v-u \text{ veces}} \\ &= p_{00}^{u}\, p_{+0}^{v-u}, \end{aligned}$$

donde la segunda igualdad está justificada por independencia entre los vectores aleatorios (X_i, Y_i).

Con argumentos similares a los anteriores se demuestra que en el caso $0 \leq v \leq u$ se cumple

$$\begin{aligned} P[U \geq u, V \geq v] &= P[(X_1, Y_1) = (0, 0), \ldots, (X_v, Y_v) = (0, 0), X_{v+1} = 0, \ldots, X_u = 0] \\ &= P[(X_1, Y_1) = (0, 0)] \cdots P[(X_v, Y_v) = (0, 0)] P[X_{v+1} = 0] \cdots P[X_u = 0] \\ &= p_{00}^{v}\, p_{0+}^{u-v}. \end{aligned}$$

Esto completa la prueba de (3.18). ∎

Proposición 3.10 *Si* (U, V) *es un vector aleatorio con la distribución geométrica bivariada* (3.17), *entonces su f.g.m.f. es*

$$m_{U,V}(s,t) = \frac{p_{11} - (s+t)\tau - st(p_{01}\,p_{10} \,-\, p_{00}\tau)}{(1-p_{0+}s)(1-p_{+0}t)(1-p_{00}st)}, \tag{3.19}$$

donde τ *es la covarianza de la distribución de Bernoulli bivariada subyacente dada por* (2.2).

Demostración. De acuerdo con la definición de $f.g.m.f.$ tenemos que

$$\begin{aligned}
m_{U,V}(s,t) &= E\left[s^U t^V\right] \\
&= \sum_{u=0}^{\infty}\sum_{v=0}^{\infty} s^u t^v P[U=u, V=v] \\
&= \sum_{u=0}^{\infty}\sum_{v=u+1}^{\infty} s^u t^v P[U=u, V=v] + \sum_{u=0, v=u}^{\infty} s^u t^v P[U=u, V=v] \\
&\qquad + \sum_{v=0}^{\infty}\sum_{u=v+1}^{\infty} s^u t^v P[U=u, V=v] \\
&= \sum_{u=0}^{\infty}\sum_{v=u+1}^{\infty} s^u t^v p_{00}^u\, p_{10}\, p_{+0}^{v-u-1}\, p_{+1} + \sum_{u=0}^{\infty} s^u t^u p_{00}^u\, p_{11} \\
&\qquad + \sum_{v=0}^{\infty}\sum_{u=v+1}^{\infty} s^u t^v p_{00}^v\, p_{01}\, p_{0+}^{u-v-1}\, p_{1+},
\end{aligned}$$

donde los tres miembros de la suma en la última igualdad están justificados porque $u < v$, $v = u$ y $v < u$, respectivamente. Así es que

$$\begin{aligned}
m_{U,V}(s,t) &= p_{10}\, p_{+1} \sum_{u=0}^{\infty} (sp_{00})^u \sum_{v=u+1}^{\infty} t^v p_{+0}^{v-u-1} + p_{11} \sum_{u=0}^{\infty} (stp_{00})^u \\
&\qquad + p_{01}\, p_{1+} \sum_{v=0}^{\infty} (tp_{00})^v \sum_{u=v+1}^{\infty} s^u p_{0+}^{u-v-1} \\
&= p_{10}\, p_{+1} t \sum_{u=0}^{\infty} (stp_{00})^u \sum_{v=u+1}^{\infty} (tp_{+0})^{v-(u+1)} + p_{11} \sum_{u=0}^{\infty} (stp_{00})^u \\
&\qquad + p_{01}\, p_{1+} s \sum_{v=0}^{\infty} (stp_{00})^v \sum_{u=v+1}^{\infty} (sp_{0+})^{u-(v+1)}.
\end{aligned}$$

Las sumas mostradas convergen siempre que se tome $|s| < 1$ y $|t| < 1$, con lo que se obtiene

$$\begin{aligned}
m_{U,V}(s,t) &= p_{10}\, p_{+1} t \left(\frac{1}{1-p_{00}st}\right)\left(\frac{1}{1-p_{+0}t}\right) + p_{11}\left(\frac{1}{1-p_{00}st}\right) \\
&\qquad + p_{01}\, p_{1+} s \left(\frac{1}{1-p_{00}st}\right)\left(\frac{1}{1-p_{0+}s}\right) \\
&= \frac{p_{10}\, p_{+1} t}{(1-p_{00}st)(1-p_{+0}t)} + \frac{p_{11}}{1-p_{00}st} + \frac{p_{01}\, p_{1+} s}{(1-p_{00}st)(1-p_{0+}s)} \\
&= \frac{p_{10}\, p_{+1}(1-p_{0+}s)t + p_{11}(1-p_{0+}s)(1-p_{+0}t) + p_{01}\, p_{1+}(1-p_{+0}t)s}{(1-p_{0+}s)(1-p_{+0}t)(1-p_{00}st)}.
\end{aligned}$$

Desarrollando los tres miembros del numerador de la última expresión se tiene

$$\begin{aligned}
&p_{10}\,p_{+1}(1-p_{0+}s)t = p_{10}\,p_{+1}t - p_{10}\,p_{+1}\,p_{0+}st\\
&p_{11}(1-p_{0+}s)(1-p_{+0}t) = p_{11} - p_{11}\,p_{+0}t - p_{11}\,p_{0+}s + p_{11}\,p_{+0}\,p_{0+}st\\
&p_{01}\,p_{1+}(1-p_{+0}t)s = p_{01}\,p_{1+} - p_{01}\,p_{1+}\,p_{+0}st.
\end{aligned}$$

Sumando ahora tales miembros y agrupando términos, resulta

$$\begin{aligned}
(3.20)\quad & p_{10}\,p_{+1}(1-p_{0+}s)t + p_{11}(1-p_{0+}s)(1-p_{+0}t) + p_{01}\,p_{1+}(1-p_{+0}t)s\\
&= p_{11} + (p_{10}\,p_{+1} - p_{11}\,p_{+0})t + (p_{01}\,p_{1+} - p_{11}\,p_{0+})s\\
&\qquad + (p_{11}\,p_{+0}\,p_{0+} - p_{10}\,p_{+1}\,p_{0+} - p_{01}\,p_{1+}\,p_{+0})st,
\end{aligned}$$

donde, recordando la definición de τ dada en (2.2), es posible simplificar de la siguiente forma

$$\begin{aligned}
p_{10}\,p_{+1} - p_{11}\,p_{+0} &= p_{10}(p_{01}+p_{11}) - p_{11}(p_{00}+p_{10})\\
&= p_{10}p_{01} + p_{10}p_{11} - p_{11}p_{00} - p_{11}p_{10} = p_{10}p_{01} - p_{11}p_{00} = -\tau
\end{aligned}$$

y

$$\begin{aligned}
p_{01}\,p_{1+} - p_{11}\,p_{0+} &= p_{01}(p_{10}+p_{11}) - p_{11}(p_{00}+p_{01})\\
&= p_{01}p_{10} + p_{01}p_{11} - p_{11}p_{00} - p_{11}p_{01} = p_{01}p_{10} - p_{11}p_{00} = -\tau,
\end{aligned}$$

por lo que se puede reescribir (3.20) como

$$\begin{aligned}
(3.21)\quad & p_{10}\,p_{+1}(1-p_{0+}s)t + p_{11}(1-p_{0+}s)(1-p_{+0}t) + p_{01}\,p_{1+}(1-p_{+0}t)s\\
&= p_{11} - t\tau - s\tau + (p_{11}\,p_{+0}\,p_{0+} - p_{10}\,p_{+1}\,p_{0+} - p_{01}\,p_{1+}\,p_{+0})st\\
&= p_{11} - (s+t)\tau + (p_{11}\,p_{+0}\,p_{0+} - p_{10}\,p_{+1}\,p_{0+} - p_{01}\,p_{1+}\,p_{+0})st.
\end{aligned}$$

Por otro lado, haciendo uso nuevamente de (2.2), se tiene

$$\begin{aligned}
&p_{11}\,p_{+0}\,p_{0+} - p_{10}\,p_{+1}\,p_{0+} - p_{01}\,p_{1+}\,p_{+0}\\
&\quad= p_{0+}(p_{11}\,p_{+0} - p_{10}\,p_{+1}) - p_{01}\,p_{1+}\,p_{+0}\\
&\quad= p_{0+}(p_{11}p_{00} + p_{11}p_{10} - p_{10}p_{01} - p_{10}p_{11}) - p_{01}\,p_{1+}\,p_{+0}\\
&\quad= p_{0+}(p_{11}p_{00} - p_{10}p_{01}) - p_{01}(p_{10}p_{00} + p_{10}p_{10} + p_{11}p_{00} + p_{11}p_{10})\\
&\quad= p_{0+}\tau - p_{01}p_{10}p_{00} - p_{01}p_{10}p_{10} - p_{01}p_{11}p_{00} - p_{01}p_{11}p_{10}\\
&\quad= p_{0+}\tau - p_{01}p_{10}(p_{00} + p_{10} + p_{11}) - p_{01}p_{11}p_{00}\\
&\quad= p_{0+}\tau - p_{01}p_{10}(1 - p_{01}) - p_{01}p_{11}p_{00}\\
&\quad= p_{0+}\tau - p_{01}p_{10} + p_{01}p_{10}p_{01} - p_{01}p_{11}p_{00}\\
&\quad= p_{0+}\tau - p_{01}p_{10} - p_{01}(p_{11}p_{00} - p_{10}p_{01})\\
&\quad= p_{00}\tau + p_{01}\tau - p_{01}p_{10} - p_{01}\tau\\
&\quad= p_{00}\tau - p_{01}p_{10}.
\end{aligned}$$

Finalmente, sustituyendo en (3.21), se obtiene

$$
\begin{aligned}
(3.22)\qquad p_{10}\,p_{+1}(1-p_{0+}s)t + p_{11}(1-p_{0+}s)(1-p_{+0}t) + p_{01}\,p_{1+}(1-p_{+0}t)s \\
&= p_{11} - (s+t)\tau + (p_{00}\tau - p_{01}p_{10})st \\
&= p_{11} - (s+t)\tau - (p_{01}p_{10} - p_{00}\tau)st.
\end{aligned}
$$

Como (3.22) es el numerador de $m_{U,V}(s,t)$, queda demostrada la afirmación. ∎

Observe que las distribuciones marginales para U y V son geométricas de parámetros p_{1+} y p_{+1}, respectivamente, como se mencionó al principio de esta sección. En efecto, dado que U es el número de ceros antes del primer éxito en la primer componente cuya probabilidad es p_{1+}, y dado que V es el número de ceros antes del primer éxito en la segunda componente con probabilidad de éxito p_{+1}, se tienen las siguientes expresiones para las respectivas distribuciones marginales

$$
P[U = u] = p_{0+}^u\, p_{1+}, \tag{3.23}
$$

$$
P[V = v] = p_{+0}^v\, p_{+1}. \tag{3.24}
$$

Proposición 3.11 *Se tiene*

$$
Corr(U,V) = \frac{\rho\,\sqrt{p_{1+}\,p_{+1}}}{1-p_{00}}\,, \tag{3.25}
$$

donde ρ es el coeficiente de correlación de la distribución de Bernoulli bivariada subyacente dado por (2.3). *Además se cumplen:*

i. *Si* $p_{10} = p_{01} = 0$, *entonces* $\quad Corr(U,V) = 1,$

ii. *Si* $p_{00} = p_{11} = 0$, *entonces* $\quad Corr(U,V) = -\sqrt{p_{10}\,p_{01}},$

iii. *En general,*

$$
-\frac{1}{2} \le Corr(U,V) \le 1.
$$

Demostración. Se procederá como en las Secciones 2.2 y 2.3, se calculará la correlación por definición, para ello se requiere de la covarianza entre U y V que como sabemos esta dada por $E[UV] - E[U]E[V]$, donde $E[UV]$ será calculado recurriendo a (3.19), que denotaremos simplemente como $m(s,t)$, sabiendo que

$$
E[UV] = \left.\frac{\partial^2 m(s,t)}{\partial s\,\partial t}\right|_{s=1,t=1}.
$$

Habrá que tener mucha paciencia para hacer a mano el cálculo correspondiente de las derivadas parciales, si el lector desea comprobar los cálculos se le recomienda que haga uso de algún software especializado en ese sentido, aquí se ha hecho uso de *"Wolfram Mathematica"*. Se muestra el resultado correspondiente a la primer derivada parcial que sí fue calculado a mano

$$
\begin{aligned}
\frac{\partial m(s,t)}{\partial t} = \frac{-\tau - (p_{01}\,p_{10} - p_{00}\,\tau)s}{(1-p_{0+}s)(1-p_{+0}t)(1-p_{00}st)} + \frac{p_{00}\,s\,(p_{11} - \tau(s+t)) - p_{00}(p_{01}\,p_{10} - p_{00}\,\tau)s^2\,t}{(1-p_{0+}s)(1-p_{+0}t)(1-p_{00}st)^2} \\
+ \frac{p_{+0}(p_{11} - \tau(s+t)) - (p_{01}\,p_{10} - p_{00}\,\tau)s\,t}{(1-p_{0+}s)\,(1-p_{+0}t)^2\,(1-p_{00}st)},
\end{aligned}
$$

con lo que es posible apreciar la magnitud del desarrollo que habría que realizar a mano para obtener la segunda derivada parcial. Aquí solo se muestra el resultado obtenido para la segunda derivada parcial, luego de hacer la evaluación correspondiente,

$$\left.\frac{\partial^2 m(s,t)}{\partial s\,\partial t}\right|_{s=1,t=1}$$
$$= \frac{-p_{01}\,p_{10}(1-p_{+0}+p_{0+}(p_{00}\,p_{+0}-1))+p_{11}[p_{0+}\,p_{+0}+p_{00}^2\,p_{0+}\,p_{+0}-p_{00}(p_{0+}+p_{+0}+p_{0+}\,p_{+0}-1)]}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}.$$

Vamos a simplificar el numerador de la expresión anterior, desarrollando y recordando (2.2) obtenemos

$$\begin{aligned}
&-p_{01}\,p_{10}+p_{01}\,p_{10}\,p_{+0}-p_{01}\,p_{10}\,p_{0+}\,p_{00}\,p_{+0}+p_{01}\,p_{10}\,p_{0+}+p_{11}\,p_{0+}\,p_{+0}\\
&\quad+p_{11}\,p_{00}^2\,p_{0+}\,p_{+0}-p_{11}\,p_{00}\,p_{0+}-p_{11}\,p_{00}\,p_{+0}-p_{11}\,p_{00}\,p_{0+}\,p_{+0}+p_{11}\,p_{00}\\
&\qquad=(p_{11}\,p_{00}-p_{01}\,p_{10})-p_{0+}(p_{11}\,p_{00}-p_{01}\,p_{10})-p_{+0}(p_{11}\,p_{00}-p_{01}\,p_{10})\\
&\qquad\qquad+p_{00}\,p_{0+}\,p_{+0}(p_{11}\,p_{00}-p_{01}\,p_{10})+(1-p_{00})p_{11}\,p_{0+}\,p_{+0}\\
&\qquad=\tau-p_{0+}\,\tau-p_{+0}\,\tau+p_{00}\,p_{0+}\,p_{+0}\,\tau+(1-p_{00})p_{11}\,p_{0+}\,p_{+0}\\
&\qquad=\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{11}\,p_{0+}\,p_{+0},
\end{aligned}$$

por lo que

$$E[UV]=\frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{11}\,p_{0+}\,p_{+0}}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\ ,$$

luego, de (3.23) y (3.24), se tiene

$$E[U]=\frac{p_{0+}}{p_{1+}}=\frac{p_{0+}}{(1-p_{0+})}\qquad \text{y}\qquad E[V]=\frac{p_{+0}}{p_{+1}}=\frac{p_{+0}}{(1-p_{+0})},$$

también

$$Var[U]=\frac{p_{0+}}{p_{1+}^2}=\frac{p_{0+}}{(1-p_{0+})^2}\qquad \text{y}\qquad Var[V]=\frac{p_{+0}}{p_{+1}^2}=\frac{p_{+0}}{(1-p_{+0})^2}.$$

Así es que como

$$Corr(U,V)=\frac{E[UV]-E[U]E[V]}{\sqrt{Var[U]Var[V]}},$$

donde

$$\sqrt{Var[U]Var[V]}=\frac{\sqrt{p_{+0}\,p_{0+}}}{(1-p_{+0})(1-p_{0+})}$$

implica

$$\frac{1}{\sqrt{Var[U]Var[V]}}=\frac{(1-p_{+0})(1-p_{0+})}{\sqrt{p_{+0}\,p_{0+}}},$$

y dado que (véase (2.2)) $\tau=p_{11}-p_{1+}\,p_{+1}$, resulta

$$\begin{aligned}
&E[UV]-E[U]E[V]\\
&\quad=\frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{11}\,p_{0+}\,p_{+0}}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}-\frac{p_{0+}\,p_{+0}}{(1-p_{0+})(1-p_{+0})}\\
&\quad=\frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{11}\,p_{0+}\,p_{+0}-p_{0+}\,p_{+0}(1-p_{00})(1-p_{0+})(1-p_{+0})}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\\
&\quad=\frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{0+}\,p_{+0}(p_{11}-(1-p_{0+})(1-p_{+0}))}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{0+}\,p_{+0}(p_{11}-p_{1+}\,p_{+1})}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\\
&= \frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0})+(1-p_{00})p_{0+}\,p_{+0}\,\tau}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\\
&= \frac{\tau(1-p_{0+}-p_{+0}+p_{00}\,p_{0+}\,p_{+0}+(1-p_{00})p_{0+}\,p_{+0})}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\\
&= \frac{\tau(1-p_{0+}-p_{+0}+p_{0+}\,p_{+0})}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\\
&= \frac{\tau(1-p_{0+})(1-p_{+0})}{(1-p_{00})(1-p_{0+})^2\,(1-p_{+0})^2}\\
&= \frac{\tau}{(1-p_{00})(1-p_{0+})(1-p_{+0})},
\end{aligned}
$$

entonces

$$
Corr(U,V) = \left(\frac{\tau}{(1-p_{00})(1-p_{0+})(1-p_{+0})}\right)\left(\frac{(1-p_{+0})(1-p_{0+})}{\sqrt{p_{+0}\,p_{0+}}}\right) = \frac{\tau}{(1-p_{00})\sqrt{p_{+0}\,p_{0+}}}. \tag{3.26}
$$

Multiplicando por $\sqrt{p_{1+}\,p_{+1}}$ tanto en el numerador como en el denominador y recordando (2.3), se obtiene

$$
\frac{\tau\sqrt{p_{1+}\,p_{+1}}}{(1-p_{00})\sqrt{p_{+0}\,p_{0+}\,p_{1+}\,p_{+1}}} = \frac{\rho\sqrt{p_{1+}\,p_{+1}}}{1-p_{00}}
$$

que es exactamente (3.25).

Nos ocupamos ahora de demostrar (i), (ii) y (iii) utilizando la expresión (3.26) equivalente a la correlación que obtuvimos recientemente, es decir,

$$
Corr(U,V) = \frac{p_{11}\,p_{00}-p_{10}\,p_{01}}{\sqrt{p_{0+}\,p_{+0}}\,(1-p_{00})}.
$$

i. Si $p_{10}=p_{01}=0$, entonces $p_{0+}=p_{+0}=p_{00}$ y $p_{00}+p_{11}=1$. Así pues,

$$
Corr(U,V) = \frac{p_{11}\,p_{00}}{p_{00}\,(1-p_{00})} = \frac{p_{11}}{1-p_{00}} = \frac{p_{11}}{p_{11}} = 1.
$$

ii. Si $p_{00}=p_{11}=0$, entonces $p_{0+}=p_{01}$, $p_{+0}=p_{10}$, luego $p_{01}+p_{10}=1$ y

$$
Corr(U,V) = \frac{-p_{10}\,p_{01}}{\sqrt{p_{01}\,p_{10}}} = -\sqrt{p_{01}\,p_{10}}.
$$

iii. Es bien sabido de la teoría que la correlación entre cualquier par de variables aleatorias toma como valor mínimo -1 y máximo 1, en el presente caso se afirma que el valor mínimo corresponde a una cantidad mayor a -1. En (ii) se obtuvo una expresión para la correlación entre U y V estrictamente menor que cero, el resultado es consecuencia de dicha expresión; podremos obtener el valor mínimo de la correlación resolviendo el problema de *minimizar*

$$
f(x) = -\sqrt{x(1-x)}
$$

sobre $[0,1]$ interpretando $x=p_{01}$. Este problema es equivalente a *maximizar*

$$
g(x) = x(1-x)
$$

sobre $[0,1]$. Como $g'(x) = -2x+1$, se tiene que $x = 1/2$ es un punto crítico y como, además, $g''(1/2) = -2$, dicho punto crítico es un punto máximo local el cual se ve de inmediato que, de hecho, en ese punto la función alcanza su valor máximo absoluto sobre $[0,1]$. Por lo tanto $x = 1/2$ minimiza a $f(x)$, es decir,

$$p_{01} = \frac{1}{2}$$

minimiza la correlación obtenida en (ii) y por lo tanto, en tal caso,

$$Corr(U,V) = -\sqrt{\left(\frac{1}{2}\right)^2} = -\frac{1}{2}$$

es el mínimo valor que puede adquirir la correlación, esto demuestra la afirmación. ∎

3.2.2. Valor esperado condicional

En esta sección, para un vector aleatorio (U,V) con la distribución GB dada por (3.17), calcularemos los valores esperados condicionales de U dado V y de V dado U, respectivamente, donde recuerde que las distribuciones marginales de U y V están dadas por (3.23) y (3.24), respectivamente.

Teorema 3.12 *Si* (U,V) *es un vector aleatorio con distribución geométrica bivariada dada por* (3.17), *entonces*

$$E[U \mid V = v] = \begin{cases} \dfrac{p_{00}}{p_{10}} - \left(\dfrac{p_{00}}{p_{+0}}\right)^v \dfrac{(1-p_{00})\,\tau}{p_{10}\,p_{1+}p_{+1}} & si \quad p_{10} \neq 0, \\ v + \dfrac{p_{01}}{p_{11}\,(1-p_{00})} & si \quad p_{10} = 0, \end{cases} \tag{3.27}$$

$$E[V \mid U = u] = \begin{cases} \dfrac{p_{00}}{p_{01}} - \left(\dfrac{p_{00}}{p_{0+}}\right)^u \dfrac{(1-p_{00})\,\tau}{p_{01}\,p_{1+}p_{+1}} & si \quad p_{01} \neq 0, \\ u + \dfrac{p_{10}}{p_{11}\,(1-p_{00})} & si \quad p_{01} = 0. \end{cases} \tag{3.28}$$

Demostración. Se probará (3.27). Supongamos primero que $p_{10} \neq 0$. Se tiene

$$E[U \mid V = v] = \sum_{u=0}^{\infty} uP[U = u \mid V = v].$$

Como el valor de v está dado, podemos partir la sumatoria en tres partes, a saber, cuando $u < v$, $u = v$ y $u > v$, de manera que

$$E[U \mid V = v] = \sum_{u=0}^{v-1} uP[U = u \mid V = v] + vP[U = v \mid V = v] + \sum_{u>v} uP[U = u \mid V = v]. \tag{3.29}$$

En los tres casos siguientes haremos uso de (3.24). Así que, cuando $u < v$, obtenemos

$$\begin{aligned} \sum_{u=0}^{v-1} uP[U = u \mid V = v] &= \sum_u u\frac{P[U = u, V = v]}{P[V = v]} = \sum_u u\frac{p_{00}^u\, p_{10}\, p_{+0}^{v-u-1}\; p_{+1}}{p_{+0}^v\, p_{+1}} \\ &= \sum_u up_{00}^u\, p_{10}\, p_{+0}^{-u-1} = \frac{p_{10}}{p_{+0}} \sum_u u\left(\frac{p_{00}}{p_{+0}}\right)^u = \frac{p_{10}\, p_{00}}{p_{+0}^2} \sum_u u\left(\frac{p_{00}}{p_{+0}}\right)^{u-1}. \end{aligned} \tag{3.30}$$

Tomando $r = \frac{p_{00}}{p_{+0}}$ y dado que $\frac{d}{dr}(r^u) = ur^{u-1}$ tenemos

$$\sum_u ur^{u-1} = \sum_u \frac{d}{dr}(r^u) = \frac{d}{dr}\left(\sum_u r^u\right)$$

donde esta última suma es convergente puesto que, como $p_{+0} > p_{00}$, necesariamente $0 \leq r < 1$. Ya que

$$\sum_{u=0}^{v-1} r^u = \frac{1-r^v}{1-r},$$

$$\frac{d}{dr}\left(\frac{1-r^v}{1-r}\right) = \frac{1-v(1-r)r^{v-1}-r^v}{(1-r)^2}$$

y

$$1-r = \frac{p_{10}}{p_{+0}},$$

llegamos a que

$$\begin{aligned}\sum_u u\left(\frac{p_{00}}{p_{+0}}\right)^{u-1} &= \left(\frac{p_{+0}}{p_{10}}\right)^2\left\{1 - v\frac{p_{10}}{p_{+0}}\left(\frac{p_{00}}{p_{+0}}\right)^{v-1} - \left(\frac{p_{00}}{p_{+0}}\right)^v\right\} \\ &= \left(\frac{p_{+0}}{p_{10}}\right)^2\left\{1 - \left(\frac{p_{00}}{p_{+0}}\right)^v\left(v\frac{p_{10}}{p_{+0}}\left(\frac{p_{00}}{p_{+0}}\right)^{-1}+1\right)\right\} \\ &= \left(\frac{p_{+0}}{p_{10}}\right)^2\left\{1 - \left(\frac{p_{00}}{p_{+0}}\right)^v\left(v\frac{p_{10}}{p_{00}}+1\right)\right\}\end{aligned} \tag{3.31}$$

Sustituyendo (3.31) en (3.30) se tiene

$$\begin{aligned}\sum_{u=0}^{v-1} uP[U=u \mid V=v] &= \frac{p_{10}\,p_{00}}{p_{+0}^2}\left(\frac{p_{+0}}{p_{10}}\right)^2\left\{1-\left(\frac{p_{00}}{p_{+0}}\right)^v\left(v\frac{p_{10}}{p_{00}}+1\right)\right\} \\ &= \frac{p_{00}}{p_{10}}\left\{1-\left(\frac{p_{00}}{p_{+0}}\right)^v\left(v\frac{p_{10}}{p_{00}}+1\right)\right\} \\ &= \frac{p_{00}}{p_{10}} - \left(\frac{p_{00}}{p_{+0}}\right)^v\left(v+\frac{p_{00}}{p_{10}}\right).\end{aligned} \tag{3.32}$$

Ahora, cuando $u = v$, se tiene que

$$vP[U=v \mid V=v] = v\frac{P[U=v, V=v]}{P[V=v]} = v\frac{p_{00}^v\,p_{11}}{p_{+0}^v\,p_{+1}} = v\left(\frac{p_{00}}{p_{+0}}\right)^v\left(\frac{p_{11}}{p_{+1}}\right). \tag{3.33}$$

Finalmente, cuando $u > v$,

$$\begin{aligned}\sum_{u>v} uP[U=u \mid V=v] &= \sum_{u>v} u\frac{P[U=u, V=v]}{P[V=v]} \\ &= \sum_{u>v} u\frac{p_{00}^v\;p_{01}\;p_{0+}^{u-v-1}\;p_{1+}}{p_{+0}^v\,p_{+1}} \\ &= \sum_{u>v} u\left(\frac{p_{00}}{p_{+0}\,p_{0+}}\right)^v\left(\frac{p_{01}\,p_{1+}}{p_{+1}}\right)p_{0+}^{u-1} \\ &= \left(\frac{p_{00}}{p_{+0}\,p_{0+}}\right)^v\left(\frac{p_{01}\,p_{1+}}{p_{+1}}\right)\sum_{u>v} up_{0+}^{u-1}.\end{aligned} \tag{3.34}$$

Sea $s = p_{0+}$. Como $\frac{d}{ds}(s^u) = us^{u-1}$, se tiene que

$$\sum_{u>v} us^{u-1} = \sum_{u>v} \frac{d}{ds}(s^u) = \frac{d}{ds}\left(\sum_{u>v} s^u\right),$$

donde

$$\sum_{u>v} s^u = \sum_{u=0}^{\infty} s^u - \sum_{u=0}^{v} s^u = \frac{1}{1-s} - \frac{1-s^{v+1}}{1-s} = \frac{s^{v+1}}{1-s}$$

cuya convergencia está justificada porque $s < 1$ dado que $p_{0+} < 1$. Luego,

$$\frac{d}{ds}\left(\frac{s^{v+1}}{1-s}\right) = \frac{(1-s)(v+1)s^v + s^{v+1})}{(1+s)^2} = \frac{s^v[(v+1)(1-s)+s]}{(1+s)^2}$$

y, dado que $1 - s = p_{1+}$, obtenemos

$$\text{(3.35)} \qquad \sum_{u>v} up_{0+}^{u-1} = \sum_{u>v} us^{u-1} = \frac{s^v[(v+1)(1-s)+s]}{(1+s)^2} = \frac{p_{0+}^v[(v+1)p_{1+} + p_{0+}]}{(p_{1+})^2} = \frac{p_{0+}^v(vp_{1+}+1)}{(p_{1+})^2}.$$

Sustituyendo (3.35) en (3.34) tenemos

$$\text{(3.36)} \qquad \begin{aligned} \sum_{u>v} uP[U = u \mid V = v] &= \left(\frac{p_{00}}{p_{+0}\, p_{0+}}\right)^v \left(\frac{p_{01}\, p_{1+}}{p_{+1}}\right) \frac{p_{0+}^v(vp_{1+}+1)}{(p_{1+})^2} \\ &= \left(\frac{p_{00}}{p_{+0}}\right)^v \frac{p_{01}\,(vp_{1+}+1)}{p_{+1}\, p_{1+}}. \end{aligned}$$

Sumando (3.32), (3.33) y (3.36) obtenemos exactamente el valor esperado condicional (3.29). En efecto, se tiene

$$\text{(3.37)} \qquad E[U \mid V = v] = \frac{p_{00}}{p_{10}} - \left(\frac{p_{00}}{p_{+0}}\right)^v \left(v + \frac{p_{00}}{p_{10}} - v\frac{p_{11}}{p_{+1}} - \frac{p_{01}\,(vp_{1+}+1)}{p_{+1}\, p_{1+}}\right).$$

Desarrollando la última expresión es posible simplificar

$$\begin{aligned} v + \frac{p_{00}}{p_{10}} - v\frac{p_{11}}{p_{+1}} - \frac{p_{01}\,(vp_{1+}+1)}{p_{+1}\, p_{1+}} &= v + \frac{p_{00}}{p_{10}} - v\frac{p_{11}}{p_{+1}} - v\frac{p_{01}}{p_{+1}} - \frac{p_{01}}{p_{+1}\, p_{1+}} \\ &= v\left(1 - \frac{p_{11}}{p_{+1}} - \frac{p_{01}}{p_{+1}}\right) + \frac{p_{00}}{p_{10}} - \frac{p_{01}}{p_{+1}\, p_{1+}} \\ &= v\left(1 - \frac{p_{11}+p_{01}}{p_{+1}}\right) + \frac{p_{00}\, p_{+1}\, p_{1+} - p_{01}\, p_{10}}{p_{10}\, p_{+1}\, p_{1+}} \\ &= v\left(1 - \frac{p_{+1}}{p_{+1}}\right) + \frac{p_{00}\, p_{+1}\, p_{1+} - p_{01}\, p_{10}}{p_{10}\, p_{+1}\, p_{1+}} \\ &= \frac{p_{00}\,(p_{01}+p_{11})\,(p_{11}+p_{10}) - p_{01}\, p_{10}}{p_{10}\, p_{+1}\, p_{1+}}, \end{aligned}$$

donde

$$
\begin{aligned}
(3.38)\quad p_{00}\,(p_{01}+p_{11})\,(p_{11}+p_{10})-p_{01}\,p_{10} &= p_{00}\,p_{01}\,p_{10}+p_{00}\,p_{01}\,p_{11}+p_{00}\,p_{11}\,p_{10}+p_{00}\,p_{11}\,p_{11}-p_{01}\,p_{10}\\
&= p_{00}\,p_{01}\,p_{10}+p_{00}\,p_{11}\,(p_{01}+p_{10}+p_{11})-p_{01}\,p_{10}\\
&= p_{00}\,p_{01}\,p_{10}+p_{00}\,p_{11}\,(p_{00}+p_{01}+p_{10}+p_{11}-p_{00})-p_{01}\,p_{10}\\
&= p_{00}\,p_{01}\,p_{10}+p_{00}\,p_{11}\,(1-p_{00})-p_{01}\,p_{10}\\
&= p_{00}\,p_{11}\,(1-p_{00})-p_{01}\,p_{10}+p_{00}\,p_{01}\,p_{10}\\
&= p_{00}\,p_{11}\,(1-p_{00})-p_{01}\,p_{10}\,(1-p_{00})\\
&= (1-p_{00})(p_{00}\,p_{11}-p_{01}\,p_{10})\\
&= (1-p_{00})\,\tau
\end{aligned}
$$

con τ dada por (2.2). Entonces

$$v+\frac{p_{00}}{p_{10}}-v\frac{p_{11}}{p_{+1}}-\frac{p_{01}\,(vp_{1+}+1)}{p_{+1}\,p_{1+}}=\frac{(1-p_{00})\,\tau}{p_{10}\,p_{+1}\,p_{1+}},$$

por lo que, sustituyendo en (3.37), obtenemos

$$E[U\mid V=v]=\frac{p_{00}}{p_{10}}-\left(\frac{p_{00}}{p_{+0}}\right)^{v}\frac{(1-p_{00})\,\tau}{p_{10}\,p_{+1}\,p_{1+}}.$$

Para completar la prueba de (3.27) falta tratar el caso $p_{10}=0$. Para ello comenzamos con (3.29) donde, por (3.30),

$$\sum_{u=0}^{v-1}uP[U=u\mid V=v]=0,$$

luego, por (3.33),

$$vP[U=v\mid V=v]=v\frac{p_{11}}{p_{+1}}$$

y de (3.36) obtenemos

$$\sum_{u>v}uP[U=u\mid V=v]=\frac{p_{01}\,(vp_{11}+1)}{p_{+1}\,p_{11}}.$$

Por lo tanto, concluimos que si $p_{10}=0$, entonces

$$
\begin{aligned}
E[U\mid V=v] &= v\frac{p_{11}}{p_{+1}}+\frac{p_{01}\,(vp_{11}+1)}{p_{+1}\,p_{11}}\\
&= v\frac{p_{11}}{p_{+1}}+v\frac{p_{01}}{p_{+1}}+\frac{p_{01}}{p_{+1}\,p_{11}}\\
&= v\left(\frac{p_{11}}{p_{+1}}+\frac{p_{01}}{p_{+1}}\right)+\frac{p_{01}}{(1-p_{+0})\,p_{11}}\\
&= v\left(\frac{p_{11}+p_{01}}{p_{+1}}\right)+\frac{p_{01}}{(1-p_{00})\,p_{11}}\\
&= v+\frac{p_{01}}{(1-p_{00})\,p_{11}},
\end{aligned}
$$

quedando demostrado en esta forma (3.27).

Con argumentos completamente similares se demuestra (3.28), es decir, que si $p_{01}\neq 0$, entonces

$$E[V\mid U=u]=\frac{p_{00}}{p_{01}}-\left(\frac{p_{00}}{p_{0+}}\right)^{u}\frac{(1-p_{00})\,\tau}{p_{01}\,p_{+1}\,p_{1+}},$$

y si $p_{01} = 0$, entonces

$$E[V \mid U = u] = u + \frac{p_{10}}{(1 - p_{00})\, p_{11}}. \blacksquare$$

Terminamos esta sección mencionando que en el caso univariado es posible obtener la distribución exponencial como límite de la distribución geométrica. El procedimiento que se utiliza para dicha derivación puede ser extendido a dos dimensiones con la distribución geométrica bivariada dada por (3.18), de hecho, se puede demostrar el resultado siguiente.

Proposición 3.13 *Sea $\{(X_i, Y_i)\}_{i=1}^{\infty}$ una sucesión de vectores aleatorios independientes, tal que (X_i, Y_i) tiene distribución de Bernoulli bivariada, que son observados en los instantes $\left\{\frac{i}{n}\right\}_{i=1}^{\infty}$, con $n \in \mathbb{N}$, y sean U y V el número de ceros observados antes del primer uno en las sucesiones $\{X_i\}_{i=1}^{\infty}$ y $\{Y_i\}_{i=1}^{\infty}$, respectivamente. Si nt_1 y nt_2 no son enteros, se sabe que*

$$P[U > t_1, V > t_2] = \begin{cases} p_{00}^{[nt_1]}\; p_{+0}^{[nt_2]-[nt_1]} & si \quad t_1 < t_2, \\ p_{00}^{[nt_2]}\; p_{0+}^{[nt_1]-[nt_2]} & si \quad t_1 > t_2. \end{cases}$$

Si se hacen tender $n \to \infty$ y $p_{ij} \to 0$ de tal suerte que $np_{ij} \to \lambda_{ij}$, entonces

$$P[U > t_1, V > t_2] \quad \longrightarrow \quad \exp[-\lambda_{10} t_1 - \lambda_{01} t_2 - \lambda_{11} \max(t_1, t_2)], \quad t_1, t_2 \geq 0,$$

donde esta última distribución es la distribución exponencial bivariada propuesta en [10].

3.2.3. Ejemplos

Se ilustran algunas situaciones comunes donde aparece de manera natural la distribución geométrica bivariada.

Ejemplo 3.14 *Lanzamiento de moneda bivariado.* Un juego en parejas, digamos "Jugador 1" y "Jugador 2" que abreviaremos, respectivamente, como J_1 y J_2, consiste en lanzar una moneda al mismo tiempo indefinidamente. J_1 debe decir "águila" y J_2 "sol" en cuanto obtengan el resultado por vez primera, los jugadores no dejan de lanzar su moneda hasta el momento en que ambos han observado su resultado. Puede suceder que J_1 observe "águila" antes de que J_2 observe "sol", quizá suceda al contrario, o incluso que ambos obtengan el resultado en el mismo instante. Este juego puede ser analizado con la distribución GB dada en el Teorema y Definición 3.9.

Considere una sucesión de vectores aleatorios $\{(X_i, Y_i)\}_{i=1}^{\infty}$ con distribución de Bernoulli bivariada. Donde, para todo i, se definen

$$X_i = \begin{cases} 1 & \text{si la moneda de } J_1 \text{ cae "águila"}, \\ 0 & \text{si no,} \end{cases} \qquad \text{y} \quad Y_i = \begin{cases} 1 & \text{si la moneda de } J_2 \text{ cae "sol"}, \\ 0 & \text{si no.} \end{cases}$$

Cada prueba, (X_i, Y_i), puede considerarse independiente. Sean U el número de lanzamientos que realiza J_1 antes de observar por primera vez "águila" y V el número de lanzamientos que realiza J_2 antes de obtener "sol", entonces (U, V) es un vector aleatorio con distribución GB.

El coeficiente de correlación entre U y V está dado por (3.25). Las probabilidades de los cuatro valores que toma cada ve.a. (X_i, Y_i) son iguales: $p_{ij} = \frac{1}{4}$ y $p_{i+} = p_{+j} = \frac{1}{2}$ para todo $i, j \in \{0, 1\}$, así es que $Corr(U, V) = 0$ (pues $\rho = 0$, vea (2.3)). ◄

Ejemplo 3.15 *El juego de la ruleta Americana de doble cero.* Se trata de un juego muy popular en los casinos, está constituido por 38 casillas, 36 de las cuales están numeradas y son de color rojo o negro, el par de casillas restantes son de color verde o azul y corresponden al 0 y 00. Las casillas de color rojo están constituidas por 8 números pares y 10 impares, las de color negro por 10 números pares y 8 impares. La ruleta se hace girar y el "arbitro" lanza una bola blanca en dirección contraria a la que se hizo girar la ruleta, gana el jugador que haya apostado por la casilla en la que caiga la bola.

Suponga que un jugador está particularmente interesado en apostar por casillas rojas y por números pares, considere que U es el número de jugadas antes de que la bola caiga en una casilla roja y que V es el número de jugadas antes de que la bola caiga en una casilla de número par, entonces el vector aleatorio (U, V) tendrá distribución GB. Considere la sucesión de vectores aleatorios $\{(X_i, Y_i)\}_{i=1}^{\infty}$, que se supondrán independientes, con distribución de Bernoulli bivariada definidos, para todo i, como

$$X_i = \begin{cases} 1 & \text{si la bola cae en una casilla roja,} \\ 0 & \text{en otro caso,} \end{cases} \quad \text{y} \quad Y_i = \begin{cases} 1 & \text{si la bola cae en una casilla par,} \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces cada uno de los ve.as. (X_i, Y_i) tiene parámetros $p_{11} = \frac{4}{19}$, $p_{10} = \frac{5}{19}$, $p_{01} = \frac{5}{19}$, y $p_{00} = \frac{5}{19}$, donde p_{11} es la probabilidad de que la bola caiga en una casilla roja y de numero par (lo cual sucede en 8 de 38 resultados posibles); p_{10} es la probabilidad de que la bola caiga en una casilla roja y de número impar (10 de 38 casos); p_{01} es la probabilidad de que la bola caiga en una casilla no roja (negra o verde) y de número par (las casillas de número par no rojas corresponden a las negras, 10 de 38 casos); p_{00} es la probabilidad de que la bola caiga en una casilla no roja y de cualquier número no par (en este caso es posible que caiga en cualquiera de las 8 casillas negras de número impar o en cualquiera de las 2 casillas verdes que corresponden al 0 y 00, lo cual sucede en 10 de 38 casos). Con esta información es posible obtener el coeficiente de correlación entre U y V dado por (3.25), como en el ejemplo anterior, $Corr(U, V) = -0.0357$, significa que hay muy poca dependencia, en sentido negativo, entre las v.as. U y V. ◄

Ejemplo 3.16 Considere el Ejemplo 2.6. Suponga que de un lote de producción se seleccionan en sucesión y de manera aleatoria los productos, con el objeto de someterlos a prueba. Es posible, como en el ejemplo anterior, describir las variables aleatorias U y V de manera que (U, V) tenga la distribución GB; la primera de ellas es el número de pruebas antes de encontrar el primer producto que presente fallas del tipo I, mientras que la otra es el número de pruebas antes de encontrar el primer producto con fallas del tipo II. ◄

3.3. La distribución binomial negativa bivariada

La distribución binomial negativa bivariada generaliza a la distribución geométrica bivariada de manera análoga al caso univariado. Para una sucesión $\{(X_i, Y_i)\}_{i=1}^{\infty}$ de vectores aleatorios independientes con distribución de Bernoulli bivariada, y para $r, s \in \mathbb{Z}$, $r \geq 0$ y $s \geq 0$, definimos U y V como el número de ceros antes de observar el r-ésimo éxito en la sucesión $\{X_i\}_{i=1}^{\infty}$ y antes de observar el s-ésimo éxito en la sucesión $\{Y_i\}_{i=1}^{\infty}$, respectivamente. Es posible ver que las variables U y V tienen, por separado, distribución binomial negativa; sin embargo, en general, como para la distribución GB, estas variables no serán independientes.

Teorema y Definición 3.17 *Sea $\{(X_i, Y_i)\}_{i=1}^{\infty}$ una sucesión de vectores aleatorios independientes tal que (X_i, Y_i) tiene distribución de Bernoulli bivariada, sean U y V el número de pruebas antes de observar el r-ésimo éxito en la sucesión $\{X_i\}_{i=1}^{\infty}$ y el número de pruebas antes de observar el s-ésimo éxito en la sucesión $\{Y_i\}_{i=1}^{\infty}$, respectivamente. La distribución conjunta del vector aleatorio (U, V) está dada por*

$$(3.39)\quad P[U \geq u, V \geq v] = \sum_{a=0}^{s-1}\sum_{l=0}^{r-1}\sum_{i=0}^{\text{mín}(a,l)}\sum_{m=0}^{s-1-a}\binom{r+u-1}{i, l-i, a-i, r+u-1-a-l+i} p_{11}^{i}\, p_{10}^{l-i}\, p_{01}^{a-i}\, p_{00}^{r+u-1-a-l+i} \times \binom{s+v-r-u}{m} p_{+1}^{m}\, p_{+0}^{s+v-r-u-m} \qquad si \quad r+u \leq s+v,$$

$$(3.40)\quad P[U \geq u, V \geq v] = \sum_{a=0}^{r-1}\sum_{l=0}^{s-1}\sum_{i=0}^{\text{mín}(a,l)}\sum_{m=0}^{r-1-a}\binom{s+v-1}{i, a-i, l-i, s+v-1-a-l+i} p_{11}^{i}\, p_{10}^{a-i}\, p_{01}^{l-i}\, p_{00}^{s+v-1-a-l+i} \times \binom{r+u-s-v}{m} p_{1+}^{m}\, p_{0+}^{r+u-s-v-m} \qquad si \quad s+v \leq r+u,$$

equivalentemente,

$$(3.41)\quad P[U = u, V = v] = \sum_{a=\text{máx}(r+u-1-v,0)}^{\text{mín}(s-2,r+u-1)}\sum_{i=0}^{\text{mín}(r-1,a)}\binom{r+u-1}{i, r-1-i, a-i, u-a+i} p_{11}^{i+1} p_{10}^{r-1-i} p_{01}^{a-i} p_{00}^{u-a+i} \times \binom{s+v-r-u-1}{s-a-2} p_{+0}^{v-u-r+a+1}\, p_{+1}^{s-a-1} + \sum_{a=\text{máx}(r+u-1-v,0)}^{\text{mín}(s-1,r+u-1)}\sum_{i=0}^{\text{mín}(r-1,a)}\binom{r+u-1}{i, r-1-i, a-i, u-a+i} p_{11}^{i}\, p_{10}^{r-i}\, p_{01}^{a-i}\, p_{00}^{u-a+i} \times \binom{s+v-r-u-1}{s-a-1} p_{+0}^{v-u-r+a}\, p_{+1}^{s-a} \qquad si \quad r+u < s+v,$$

$$(3.42)\quad P[U = u, V = v] = \sum_{a=\text{máx}(s+v-1-u,0)}^{\text{mín}(r-2,s+v-1)}\sum_{i=0}^{\text{mín}(s-1,a)}\binom{s+v-1}{i, s-1-i, a-i, v-a+i} p_{11}^{i+1}\, p_{10}^{s-1-i}\, p_{01}^{a-i}\, p_{00}^{v-a+i} \times \binom{r+u-s-v-1}{r-a-2} p_{0+}^{u-v-s+a+1}\, p_{1+}^{r-a-1} + \sum_{a=\text{máx}(s+v-1-u,0)}^{\text{mín}(r-1,s+v-1)}\sum_{i=0}^{\text{mín}(s-1,a)}\binom{s+v-1}{i, s-1-i, a-i, v-a+i} p_{11}^{i}\, p_{10}^{s-i}\, p_{01}^{a-i}\, p_{00}^{v-a+i} \times \binom{r+u-s-v-1}{r-a-1} p_{0+}^{u-v-s+a}\, p_{1+}^{r-a} \qquad si \quad s+v < r+u,$$

$$(3.43)\quad P[U = u, V = v] = \sum_{i=0}^{\text{mín}(r-1,s-1)}\binom{r+u-1}{i, r-1-i, s-1-i, u-s+1+i} p_{11}^{i+1}\, p_{10}^{r-1-i}\, p_{01}^{s-1-i}\, p_{00}^{u-s+1+i} \qquad si \quad s+v = r+u,$$

donde los coeficientes binomiales anteriores se anulan si alguna de sus entradas es un número negativo. Esta distribución conjunta es conocida como la **distribución binomial negativa bivariada** *(BNB).*

Demostración. La derivación de las identidades del enunciado se hará pensando en un experimento similar al que se desarrolló en la sección anterior. En este caso, el experimento concluye una vez que se observan r éxitos en la primer componente y s en la segunda componente. Como siempre, pensaremos que un éxito corresponde al evento en el que las variables aleatorias toman el valor uno, es decir, $X = 1$ y/o $Y = 1$, respectivamente.

Se trabajará con los siguientes cinco casos que corresponden a las cinco expresiones numeradas de la (3.39) a la (3.43). Los dos primeros corresponden a la probabilidad de observar al menos u ceros en la primer componente antes del r-ésimo éxito y al menos v ceros en la segunda componente antes del s-ésimo éxito.

Caso 1: $r + u \leq s + v$. En este caso, $P[U \geq u, V \geq v]$ es la probabilidad de que se observan a lo más $r-1$ éxitos en la primer componente entre las primeras $r+u-1$ pruebas y a lo más $s-1$ éxitos en la segunda componente entre todas las $s+v-1$ pruebas de las que consiste el experimento. Un evento favorable corresponde exactamente a un total de $s+v-1$ pruebas, por lo que un punto muestral completo será representado por un arreglo de tamaño $s+v-1$. Se hará uso de un índice auxiliar cuyo papel es contar el número de unos que aparecen en la segunda componente entre las primeras $r+u-1$ pruebas, digamos $a \in \{0, 1, \ldots, s-1\}$. Podemos describir abreviadamente lo que acabamos de argumentar como

$$\begin{aligned} P[U \geq u, V \geq v] = \sum_{a=0}^{s-1} P[&\text{a lo más } r-1 \text{ unos entre } X_1, \ldots, X_{r+u-1}, \\ &\text{exactamente } a \text{ unos entre } Y_1, \ldots, Y_{r+u-1}, \\ &\text{a lo más } s-1-a \text{ unos entre } Y_{r+u}, \ldots, Y_{s+v-1}]. \end{aligned}$$

Sean los eventos

$$E = \left\{ (x_1, \ldots, x_{r+u-1}, y_1, \ldots, y_{r+u-1}) \;\middle|\; \sum_{j=1}^{r+u-1} x_j \leq r-1 \quad \text{y} \quad \sum_{j=1}^{r+u-1} y_j = a \right\}$$

y

$$F = \left\{ (y_{r+u}, \ldots, y_{s+v-1}) \;\middle|\; \sum_{j=r+u}^{s+v-1} y_j \leq s-1-a \right\}$$

de modo tal que podemos reescribir la expresión anterior, por independencia entre vectores aleatorios y variables aleatorias, de la siguiente forma

$$\begin{aligned} P[U \geq u, V \geq v] = \sum_{a=0}^{s-1} &P[(X_1, \ldots, X_{r+u-1}, Y_1, \ldots, Y_{r+u-1}) \in E] \\ &\times P[(Y_{r+u}, \ldots, Y_{s+v-1}) \in F]. \end{aligned}$$

Para el primer factor del lado derecho (el que corresponde al evento E) de la última igualdad consideramos $l \in \{0, 1, \ldots, r-1\}$ e $i \in \{0, 1, \ldots, \text{mín}(a, l)\}$ índices auxiliares cuyo trabajo es el de contar el número de unos observados en la primer componente y en ambas componentes, respectivamente. Un arreglo de tamaño $r+u-1$ es de la siguiente forma

$$\underbrace{p_{10}\,p_{10}\cdots p_{10}}_{\substack{l-i \text{ unos en} \\ \text{la primer} \\ \text{componente}}}\ \underbrace{p_{11}\,p_{11}\cdots p_{11}}_{\substack{i \text{ unos en} \\ \text{ambas} \\ \text{componentes}}}\ \underbrace{p_{01}\,p_{01}\cdots p_{01}}_{\substack{a-i \text{ unos en} \\ \text{la segunda} \\ \text{componente}}}\ \underbrace{p_{00}\,p_{00}\cdots p_{00}}_{\substack{r+u-1-(l+a-i) \\ \text{términos}}}$$

$$= p_{11}^{i}\,p_{10}^{l-i}\,p_{01}^{a-i}\,p_{00}^{r+u-1-(l+a-i)}.$$

Notemos que $p_{11}^{i}\,p_{10}^{l-i}$ significa que se han observado exactamente l unos en la primer componente, donde l puede contar a lo más $r-1$ unos, y por otro lado, $p_{11}^{i}\,p_{01}^{a-i}$ significa que se han observado exactamente a unos en la segunda componente, siendo que a puede contar a lo más $s-1$ unos, pero entonces se deben haber incluido por los menos $r+u-1-a$ ceros, cantidad que no debe exceder a v, es decir, a debe ser por lo menos $r+u-1-v$. Así pues, el rango de valores verdaderos de a debe ser de $\text{máx}(r+u-1-v,0)$ a $\text{mín}(r+u-1,s-1)$. Esta probabilidad describe la primer parte correspondiente al evento E. Como el número de formas en que puede ocurrir tal evento es

$$\binom{r+u-1}{i,l-i,a-i,r+u-1-a-l+i},$$

entonces

$$P[(X_1,\ldots,X_{r+u-1},Y_1,\ldots,Y_{r+u-1})\in E]=\sum_{l=0}^{r-1}\sum_{i=0}^{\text{mín}(a,l)}\binom{r+u-1}{i,l-i,a-i,r+u-1-a-l+i}$$
$$\times\, p_{11}^{i}\,p_{10}^{l-i}\,p_{01}^{a-i}\,p_{00}^{r+u-1-a-l+i}.$$

La segunda parte, correspondiente al evento F, consta de $s+v-(r+u)$ observaciones. Consideramos $m\in\{0,1,\ldots,s-1-a\}$ como el número de unos que aparecen en la segunda componente entre las $s+v-r-u$ observaciones restantes que comprenden al experimento. Se tiene que un arreglo, cuyo tamaño es el número de observaciones indicadas, puede observarse en la siguiente forma

$$p_{+1}^{m}\,p_{+0}^{s+v-r-u-m}.$$

Este evento puede ocurrir en

$$\binom{s+v-r-u}{m}$$

ocasiones y por lo tanto

$$P[(Y_{r+u},\ldots,Y_{s+v-1})\in F]=\sum_{m=0}^{s-1-a}\binom{s+v-r-u}{m}p_{+1}^{m}\,p_{+0}^{s+v-r-u-m}.$$

Se sigue de todo lo anterior que

$$P[U\geq u,V\geq v]$$
$$=\sum_{a=0}^{s-1}\sum_{l=0}^{r-1}\sum_{i=0}^{\text{mín}(a,l)}\sum_{m=0}^{s-1-a}\binom{r+u-1}{i,l-i,a-i,r+u-1-a-l+i}p_{11}^{i}\,p_{10}^{l-i}\,p_{01}^{a-i}\,p_{00}^{r+u-1-a-l+i}$$
$$\times\binom{s+v-r-u}{m}p_{+1}^{m}\,p_{+0}^{s+v-r-u-m}\,,$$

que es exactamente (3.39).

Caso 2: $s+v \leq r+u$. En este caso, $P[U \geq u, V \geq v]$ es la probabilidad de que se observan a lo más $s-1$ éxitos en la segunda componente entre las primeras $s+v-1$ pruebas y a lo más $r-1$ éxitos en la primer componente entre todas las $r+u-1$ pruebas de las que consiste el experimento. Un evento favorable corresponde exactamente a un total de $r+u-1$ pruebas, por lo que un punto muestral completo se puede representar mediante a un arreglo de tamaño $r+u-1$. Este caso es completamente análogo al anterior, intercambiando los papeles de u con v y de r con s obtendremos la distribución indicada, además de las elecciones para los índices auxiliares.

El índice auxiliar con el que comenzamos ahora contará el número de unos observados en la primer componente entre las primeras $s+v-1$ observaciones, ésta es tal que $a \in \{0, 1, \ldots, r-1\}$. Luego

$$
\begin{aligned}
P[U \geq u, V \geq v] = \sum_{a=0}^{r-1} P[&\text{a lo más } s-1 \text{ unos entre } Y_1, \ldots, Y_{s+v-1},\\
&\text{exactamente } a \text{ unos entre } X_1, \ldots, X_{s+v-1},\\
&\text{a lo más } r-1-a \text{ unos entre } X_{s+v}, \ldots, X_{r+u-1}].
\end{aligned}
$$

Un punto muestral correspondiente a las primeras $s+v-1$ pruebas puede tener la siguiente forma

$$
\underbrace{p_{01}\, p_{01} \cdots p_{01}}_{\substack{l-i \text{ unos en}\\ \text{la segunda}\\ \text{componente}}}\;
\underbrace{p_{11}\, p_{11} \cdots p_{11}}_{\substack{i \text{ unos en}\\ \text{ambas}\\ \text{componentes}}}\;
\underbrace{p_{10}\, p_{10} \cdots p_{10}}_{\substack{a-i \text{ unos en}\\ \text{la primer}\\ \text{componente}}}\;
\underbrace{p_{00}\, p_{00} \cdots p_{00}}_{\substack{s+v-1-(l+a-i)\\ \text{términos}}}
$$

$$
= p_{11}^{i}\, p_{10}^{a-i}\, p_{01}^{l-i}\, p_{00}^{s+v-1-(l+a-i)},
$$

donde $l \in \{0, 1, \ldots, s-1\}$ es el número de unos observados en la segunda componente, $i \in \{0, 1, \ldots, \text{mín}(a, l)\}$ el número de unos observados en ambas componentes. Este punto muestral puede ocurrir en el siguiente número de ocasiones

$$
\binom{s+v-1}{i, a-i, l-i, s+v-1-a-l+i}.
$$

Las observaciones restantes corresponden a los $r+u-(s+v)$ lugares del arreglo completo (un punto muestral) de tamaño $r+u-1$. En este momento se han observado a unos en la primer componente, de modo que, entre las variables $X_{s+v}, \ldots, X_{r+u-1}$ se observan a lo más $r-1-a$ unos, teniendo así, en total, observados a lo más $r-1$ unos en la primer componente. No siendo relevante lo que ocurra, desde este punto, en la segunda componente, un arreglo de tamaño $r+u-s-v$ tendrá la siguiente forma

$$
p_{1+}^{m}\, p_{0+}^{r+u-s-v-m}
$$

cuyo número de posibilidades es

$$
\binom{r+u-s-v}{m},
$$

donde $m \in \{0, 1, \ldots, r-1-a\}$ y por lo tanto se obtiene que

$$
\begin{aligned}
&P[U \geq u, V \geq v] \\
&\quad = \sum_{a=0}^{r-1} \sum_{l=0}^{s-1} \sum_{i=0}^{\min(a,l)} \sum_{m=0}^{r-1-a} \binom{s+v-1}{i, a-i, l-i, s+v-1-a-l+i} p_{11}^{i}\, p_{10}^{a-i}\, p_{01}^{l-i}\, p_{00}^{s+v-1-a-l+i} \\
&\qquad\qquad \times \binom{r+u-s-v}{m} p_{1+}^{m}\, p_{0+}^{r+u-s-v-m},
\end{aligned}
$$

que es exactamente (3.40).

En los siguientes tres casos se trabaja con un número exacto de observaciones. Se trata de $P[U = u, V = v]$ que es la probabilidad de observar exactamente u ceros antes del r-ésimo uno en la primer componente y exactamente v ceros antes del s-ésimo éxito en la segunda componente. Las expresiones dependen de u, v, r y s.

Caso 3: $r + u < s + v$. En este caso el experimento consta de $s + v$ observaciones, si pensamos en las distribuciones marginales entones el experimento que trata las observaciones en la primer componente termina antes que el experimento que trata con la segunda componente. La cantidad $r+u$ es exactamente el número de ceros y unos en la primer componente, r unos y u ceros, por supuesto que el r-ésimo uno es observado en la prueba número $r + u$; así mismo $s + v$ es el número de ceros y unos observados en la segunda componente, en la prueba $r + u$-ésima es importante considerar que la segunda componente puede ser éxito o fracaso y finalmente habrá que considerar que la última observación, la $s + v$-ésima, corresponde al s-ésimo éxito. Se deberá tener cuidado, dado que $r + u < s + v$, de no contar s o más unos y de no contar v o más ceros en las primeras $r + u$ pruebas, esto será controlado mediante un índice auxiliar que llamaremos a.

Describimos la probabilidad buscada mediante la siguiente expresión

$$
\begin{aligned}
&P[U = u, V = v] \\
&= \sum_{a} P[\text{exactamente } r-1 \text{ unos entre } X_1, \ldots, X_{r+u-1}, \qquad (3.44)\\
&\qquad \text{exactamente } a \text{ unos entre } Y_1, \ldots, Y_{r+u-1},\ (X_{r+u}, Y_{r+u}) = (1,1), \\
&\qquad\qquad \text{exactamente } s-(a+1)-1 \text{ unos entre } Y_{r+u+1}, \ldots, Y_{s+v-1},\ Y_{s+v} = 1] \\
&+ \sum_{a} P[\text{exactamente } r-1 \text{ unos entre } X_1, \ldots, X_{r+u-1}, \qquad (3.45)\\
&\qquad \text{exactamente } a \text{ unos entre } Y_1, \ldots, Y_{r+u-1},\ (X_{r+u}, Y_{r+u}) = (1,0), \\
&\qquad\qquad \text{exactamente } s-a-1 \text{ unos entre } Y_{r+u+1}, \ldots, Y_{s+v-1},\ Y_{s+v} = 1].
\end{aligned}
$$

En (3.44) se considera el caso donde $Y_{r+u} = 1$ por lo que, en las primeras $r + u$ pruebas se han observado $a + 1$ unos en la segunda componente, de modo que en las restantes observaciones se tendrán exactamente $s - (a + 1) - 1$ unos para tal componente. Pero entonces se habrán observado $r + u - 1 - a$ ceros en las primeras $r + u - 1$ pruebas, luego $r + u - 1 - a \leq v$, es decir, $r + u - 1 - v \leq a$. Por ello es que el índice auxiliar correspondiente satisface

$$a \in \{\max(r + u - 1 - v, 0), \ldots, \min(s - 2, r + u - 1)\}.$$

La cantidad de éxitos hasta la $r+u-1$-ésima prueba no puede ser mayor a $s-2$. Piense, por ejemplo, en el caso extremo, cuando $a=s-2$. Recuerde que a cuenta el número de unos en las primeras $r+u-1$ pruebas para la segunda componente, si $Y_{r+u}=1$ entonces se han observado $a+1$ unos, es decir, $s-1$ éxitos en la segunda componente, por lo tanto es de esperarse que entre las variables $Y_{r+u+1},\ldots,Y_{s+v-1}$ no sea observado éxito alguno como está bien indicado por la expresión anterior, $s-(a+1)-1=0$ para $a=s-2$, el siguiente éxito (el r-ésimo) debe ser el observado en la última prueba, es decir, $Y_{r+v}=1$, con lo que se tienen r éxitos observados.

Por otro lado (3.45) corresponde al caso en que $Y_{r+u}=0$, esto significa que en las primeras $r+u$ pruebas se han observado a unos en la segunda componente, quedando por ser observados justo $s-a-1$ unos en la misma. En este caso el índice auxiliar cambia a

$$a \in \{\max(r+u-1-v,0),\ldots,\min(s-1,r+u-1)\}.$$

Hasta la $r+u-1$-ésima prueba se tiene el mismo punto muestral tanto para el caso de (3.44) como de (3.45), lo cual puede ser observado en la siguiente forma

$$\underbrace{p_{10}\,p_{10}\cdots p_{10}}_{\substack{(r-1)-i \text{ unos en}\\ \text{la primer}\\ \text{componente}}}\quad \underbrace{p_{11}\,p_{11}\cdots p_{11}}_{\substack{i \text{ unos en}\\ \text{ambas}\\ \text{componentes}}}\quad \underbrace{p_{01}\,p_{01}\cdots p_{01}}_{\substack{a-i \text{ unos en}\\ \text{la segunda}\\ \text{componente}}}\quad \underbrace{p_{00}\,p_{00}\cdots p_{00}}_{\substack{u-a+i\\ \text{términos}}}$$

$$= p_{11}^{i}\,p_{10}^{r-1-i}\,p_{01}^{a-i}\,p_{00}^{u-a+i},$$

donde $i\in\{0,1,\ldots,\min(r-1,a)\}$ lleva el control en el número de posibilidades para observar uno en ambas componentes. Note que $p_{11}^{i}\,p_{10}^{r-1-i}$ significa que se han observado exactamente $r-1$ unos en la primer componente y que $p_{11}^{i}\,p_{01}^{a-i}$ significa que se han observado exactamente a unos en la segunda componente. Este punto muestral puede suceder en el siguiente número de ocasiones

$$\binom{r+u-1}{i,r-1-i,a-i,u-a+i}.$$

Por lo tanto se tiene hasta el momento una probabilidad igual a

$$\binom{r+u-1}{i,r-1-i,a-i,u-a+i}p_{11}^{i}\,p_{10}^{r-1-i}\,p_{01}^{a-i}\,p_{00}^{u-a+i}\,. \tag{3.46}$$

La observación que le sigue, la $r+u$-ésima, es donde ocurre exactamente el r-ésimo éxito en la primer componente, es decir, $X_{r+u}=1$. En este punto existen, como ya se indicó anteriormente, las dos posibilidades siguientes para Y_{r+u}.

i) Caso: $Y_{r+u}=1$. El resultado observado corresponde a la variable $(X_{r+u},Y_{r+u})=(1,1)$ con probabilidad p_{11}. De manera que la probabilidad contada hasta el momento, multiplicando por (3.46), es

$$\binom{r+u-1}{i,r-1-i,a-i,u-a+i}p_{11}^{i+1}\,p_{10}^{r-1-i}\,p_{01}^{a-i}\,p_{00}^{u-a+i}\,. \tag{3.47}$$

El número restante de pruebas para el experimento es $s+v-1-(r+u)$ en el que se observan exactamente $s-(a+1)-1=s-a-2$ unos entre las variables aleatorias $Y_{r+u+1},\ldots,Y_{s+v-1}$. La probabilidad de cada uno de estos puntos muestrales es pues

$$\underbrace{p_{+1}\,p_{+1}\cdots p_{+1}}_{\substack{s-a-2\text{ unos en}\\ \text{la segunda}\\ \text{componente}}}\quad \underbrace{p_{+0}\,p_{+0}\cdots p_{+0}}_{\substack{v-u-r+a+1\\ \text{términos}}} \;=\; p_{+1}^{s-a-2}\,p_{+0}^{v-u-r+a+1}\,,$$

que ocurre el siguiente número de ocasiones

$$\binom{s+v-1-r-u}{s-a-2},$$

por lo que la probabilidad correspondiente es

$$\binom{s+v-1-r-u}{s-a-2}p_{+1}^{s-a-2}\,p_{+0}^{v-u-r+a+1}$$

a la cual multiplicamos por la probabilidad de la última observación que es p_{+1}, correspondiente a $Y_{s+v}=1$. Así es que obtenemos

$$\binom{s+v-1-r-u}{s-a-2}p_{+1}^{s-a-1}\,p_{+0}^{v-u-r+a+1}\,, \tag{3.48}$$

que, junto con (3.47) y considerando todos las posibilidades para a e i, llegamos a la expresión correspondiente dados los argumentos en (3.44)

$$\sum_{a=\text{máx}(r+u-1-v,0)}^{\text{mín}(s-2,r+u-1)}\ \sum_{i=0}^{\text{mín}(r-1,a)}\binom{r+u-1}{i,r-1-i,a-i,u-a+i}p_{11}^{i+1}\,p_{10}^{r-1-i}\,p_{01}^{a-i}\,p_{00}^{u-a+i} \times\binom{s+v-1-r-u}{s-a-2}p_{+1}^{s-a-1}\,p_{+0}^{v-u-r+a+1}\,. \tag{3.49}$$

ii) Caso: $Y_{r+u}=0$. El resultado observado corresponde a la variable $(X_{r+u},Y_{r+u})=(1,0)$ con probabilidad p_{10}, de manera que, multiplicando por (3.46), obtenemos

$$\binom{r+u-1}{i,r-1-i,a-i,u-a+i}p_{11}^{i}\,p_{10}^{r-i}\,p_{01}^{a-i}\,p_{00}^{u-a+i}\,. \tag{3.50}$$

El número restante de pruebas para el experimento es $s+v-1-(r+u)$ en el que ahora se observan exactamente $s-a-1$ unos entre las variables aleatorias $Y_{r+u+1},\ldots,Y_{s+v-1}$, que de modo similar al que se ha procedido en el punto anterior llegamos al siguiente resultado

$$p_{+1}^{s-a-1}\,p_{+0}^{v-u-r+a}$$

que es posible en el siguiente número de ocasiones

$$\binom{s+v-1-r-u}{s-a-1}.$$

Así pues, la probabilidad correspondiente es

$$\binom{s+v-1-r-u}{s-a-1}p_{+1}^{s-a-1}\,p_{+0}^{v-u-r+a}$$

a la cual multiplicamos por la probabilidad de la última observación, $Y_{s+v} = 1$; p_{+1}. Se obtiene en esta forma

$$\binom{s+v-1-r-u}{s-a-1} p_{+1}^{s-a}\, p_{+0}^{v-u-r+a} \tag{3.51}$$

que, junto con (3.50) y considerando todos las posibilidades para a e i, llegamos a la expresión correspondiente dados los argumentos en (3.45), es decir,

$$\begin{aligned} \sum_{a=\max(r+u-1-v,0)}^{\min(s-1,r+u-1)} \sum_{i=0}^{\min(r-1,a)} & \binom{r+u-1}{i, r-1-i, a-i, u-a+i} p_{11}^{i}\, p_{10}^{r-i}\, p_{01}^{a-i}\, p_{00}^{u-a+i} \\ & \times \binom{s+v-1-r-u}{s-a-1} p_{+1}^{s-a}\, p_{+0}^{v-u-r+a} \,. \end{aligned} \tag{3.52}$$

Así es como, finalmente, por (3.49) y (3.52), obtenemos (3.41).

Caso 4: $s+v < r+u$. En este caso (3.42) se demuestra de manera análoga al caso anterior intercambiando los papeles de las variables v por u y s por r.

Caso 5: $r+u = s+v$. Pensemos en que el experimento corresponde a $r+u$ pruebas, en las que son contabilizados r éxitos en la primer componente y s en la segunda, el experimento termina al mismo tiempo para ambas componentes (pensando en las distribuciones marginales), esto significa que la última observación corresponde al vector $(1,1)$. Es posible describir la probabilidad mediante los siguientes argumentos

$$\begin{aligned} P[U=u, V=v] = P[&\text{exactamente } r-1 \text{ unos entre } X_1, \ldots, X_{r+u-1}, \\ & \text{exactamente } s-1 \text{ unos entre } Y_1, \ldots, Y_{r+u-1},\ (X_{r+u}, Y_{r+u}) = (1,1)]. \end{aligned}$$

Son observados $r-1$ unos en la primer componente y $s-1$ unos en la segunda componente hasta la $r+u-1$-ésima prueba, seguido de la última observación $(X_{r+u}, Y_{r+u}) = (1,1)$. Un punto muestral ordenado para tal experimento, excluyendo la última observación, corresponde a un arreglo de tamaño $r+u-1$ que puede ser observado en la siguiente forma

$$\begin{aligned} & \underbrace{p_{10}\, p_{10} \cdots p_{10}}_{(r-1)-i \text{ unos}} \quad \underbrace{p_{11}\, p_{11} \cdots p_{11}}_{i \text{ unos}} \quad \underbrace{p_{01}\, p_{01} \cdots p_{01}}_{(s-1)-i \text{ unos}} \quad \underbrace{p_{00}\, p_{00} \cdots p_{00}}_{\substack{u-s+1+i \\ \text{términos}}} \\ & = p_{11}^{i}\, p_{10}^{r-1-i}\, p_{01}^{s-1-i}\, p_{00}^{u-s+1+i} \,, \end{aligned}$$

donde $i \in \{0, 1, \ldots, \min(r-1, s-1)\}$ cuenta el número de éxitos observados en ambas componentes. Note que $p_{11}^{i}\, p_{10}^{r-1-i}$ significa que se han observado $r-1$ unos en la primer componente y $p_{11}^{i}\, p_{01}^{s-1-i}$ significa que se tienen $s-1$ unos en la segunda componente. Este punto muestral es posible en

$$\binom{r+u-1}{i, r-1-i, s-1-i, u-s+1+i}$$

ocasiones, luego entonces, agregando la probabilidad de la última observación en el arreglo anterior, p_{11}, obtenemos

$$\binom{r+u-1}{i, r-1-i, s-1-i, u-s+1+i} p_{11}^{i+1}\, p_{10}^{r-1-i}\, p_{01}^{s-1-i}\, p_{00}^{u-s+1+i} \,,$$

por lo que considerando todas las posibilidades para i llegamos a

$$P[U=u,V=v]=\sum_{i=0}^{\min(r-1,s-1)}\binom{r+u-1}{i,r-1-i,s-1-i,u-s+1+i}p_{11}^{i+1}\,p_{10}^{r-1-i}\,p_{01}^{s-1-i}\,p_{00}^{u-s+1+i}$$

con lo que se ha obtenido (3.43). ∎

3.3.1. Convergencia a la distribución de Poisson bivariada.

En esta sección, así como se ha venido haciendo en la mayoría de las secciones anteriores, discutiremos un resultado de convergencia. En el caso univariado es posible obtener la distribución de Poisson como límite de la binomial negativa, vamos a ver aquí que el resultado se preserva en el caso bivariado.

Teorema 3.18 *Sean $\lambda_{11},\lambda_{10},\lambda_{01}>0$ y $0<c<1$ constantes y sean $u,v\in\mathbb{N}$ arbitrarios. Se considera $s\in\mathbb{N}$ arbitrariamente grande pero de tal suerte que $r=cs\in\mathbb{N}$. Definimos*

$$p_{1+}=\frac{cs}{\lambda_1+cs}\,,\qquad p_{+1}=\frac{s}{\lambda_2+s}\qquad y\qquad p_{11}=\frac{cs}{\lambda+cs},$$

donde $\lambda_1=\lambda_{10}+\lambda_{11}$, $\lambda_2=\lambda_{01}+\lambda_{11}$, $\lambda=\lambda_{01}+\lambda_{10}+\lambda_{11}$. Si (U,V) es un vector aleatorio con distribución binomial negativa bivariada BNB (con parámetros p_{10},p_{01},p_{11}) dada por (3.41), entonces

$$\lim_{s\to\infty}P[U=u,V=v]=P[U^*=u,V^*=v]$$

donde (U^*,V^*) es un vector aleatorio con distribución de Poisson bivariada PB (con parámetros $\lambda_{11},\lambda_{10},\lambda_{01}$) dada por (2.36).

Demostración. Se calcularán primero las correspondientes p_{10}, p_{01}, $p_{01}+p_{10}+p_{11}$, p_{0+} y p_{+0} de acuerdo a la elección tomada para p_{1+},p_{+1} y p_{11} (note, por cierto, que todas estas cantidades son en realidad funciones de s). Se tiene que

$$p_{10}=p_{1+}-p_{11}\qquad \text{y}\qquad p_{01}=p_{+1}-p_{11},$$

es decir,

$$\begin{aligned}p_{10}=\frac{cs}{\lambda_1+cs}-\frac{cs}{\lambda+cs}=\frac{cs(\lambda+cs-\lambda_1-cs)}{(\lambda_1+cs)(\lambda+cs)}&=\frac{cs(\lambda-\lambda_1)}{(\lambda_1+cs)(\lambda+cs)}\\&=\frac{cs(\lambda-\lambda_1)}{(cs)^2\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}=\frac{\lambda-\lambda_1}{cs\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\end{aligned}$$

y

$$\begin{aligned}p_{01}=\frac{s}{\lambda_2+s}-\frac{cs}{\lambda+cs}=\frac{s(\lambda+cs)-cs(\lambda_2+s)}{(\lambda_2+s)(\lambda+cs)}&=\frac{s(\lambda-c\lambda_2)}{(\lambda_2+s)(\lambda+cs)}\\&=\frac{s(\lambda-c\lambda_2)}{s\left(1+\frac{\lambda_2}{s}\right)(\lambda+cs)}=\frac{\lambda-c\lambda_2}{cs\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\,.\end{aligned}$$

Vemos ahora que

$$
\begin{aligned}
p_{01}+p_{10}+p_{11} &= \left(\frac{s}{\lambda_2+s}-\frac{cs}{\lambda+cs}\right)+\left(\frac{cs}{\lambda_1+cs}-\frac{cs}{\lambda+cs}\right)+\frac{cs}{\lambda+cs}\\
&= \frac{s}{\lambda_2+s}+\frac{cs}{\lambda_1+cs}-\frac{cs}{\lambda+cs}\\
&= \frac{s(\lambda_1+cs)+cs(\lambda_2+s)}{(\lambda_2+s)(\lambda_1+cs)}-\frac{cs}{\lambda+cs}\\
&= \frac{(s\lambda_1+2scs+cs\lambda_2)(\lambda+cs)-cs(\lambda_2+s)(\lambda_1+cs)}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{s\lambda_1\lambda+s\lambda_1cs+2scs\lambda+2s(cs)^2+cs\lambda_2\lambda+(cs)^2\lambda_2-cs\lambda_2\lambda_1-(cs)^2\lambda_2-s\lambda_1cs-s(cs)^2}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{s\lambda_1\lambda+2scs\lambda+s(cs)^2+cs\lambda_2\lambda-cs\lambda_2\lambda_1}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}
\end{aligned}
$$

y como $p_{00}=1-(p_{01}+p_{10}+p_{11})$, se tiene

$$
\begin{aligned}
p_{00} &= 1-\frac{s\lambda_1\lambda+2scs\lambda+s(cs)^2+cs\lambda_2\lambda-cs\lambda_2\lambda_1}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}-\frac{s\lambda_1\lambda+2scs\lambda+s(cs)^2+cs\lambda_2\lambda-cs\lambda_2\lambda_1}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{\lambda_1\lambda\lambda_2+s\lambda_1\lambda+cs\lambda_1\lambda_2+scs\lambda_1+cs\lambda\lambda_2+scs\lambda+(cs)^2\lambda_2+s(cs)^2}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&\quad -\frac{s\lambda_1\lambda+2scs\lambda+s(cs)^2+cs\lambda_2\lambda-cs\lambda_2\lambda_1}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{\lambda_1\lambda\lambda_2+2cs\lambda_1\lambda_2+scs\lambda_1+(cs)^2\lambda_2-scs\lambda}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{\lambda_1\lambda_2(2cs+\lambda)+cs(s\lambda_1+cs\lambda_2-s\lambda)}{(\lambda_2+s)(\lambda_1+cs)(\lambda+cs)}\\
&= \frac{cs\left(\lambda_1\lambda_2(2+\frac{\lambda}{cs})+(s\lambda_1+cs\lambda_2-s\lambda)\right)}{cs(\lambda_2+s)(\lambda_1+cs)(1+\frac{\lambda}{cs})}\\
&= \frac{\lambda_1\lambda_2(2+\frac{\lambda}{cs})+s(\lambda_1+c\lambda_2-\lambda)}{(\lambda_2+s)(\lambda_1+cs)(1+\frac{\lambda}{cs})}\ .
\end{aligned}
$$

Por otro lado, también se tienen

$$
p_{0+}=1-p_{1+}=\frac{\lambda_1}{\lambda_1+cs} \qquad \text{y} \qquad p_{+0}=1-p_{+1}=\frac{\lambda_2}{\lambda_2+s}.
$$

Recordemos que, para $r+u<s+v$, se tiene que $P[U=u,V=v]$, dada por (3.41), queda expresada como la suma de dos sumatorias, digamos A y B, definidas como

$$
\begin{aligned}
A = \sum_{a=\text{máx}(r+u-v-1,0)}^{\text{mín}(r+u-1,s-2)}\ \sum_{i=0}^{\text{mín}(r-1,a)} &\binom{r+u-1}{i,r-1-i,a-i,u-a+i} p_{11}^{i+1}\,p_{10}^{r-1-i}\,p_{01}^{a-i}\,p_{00}^{u-a+i}\\
&\times\binom{s+v-r-u-1}{s-a-2} p_{+0}^{v-u-r+a+1}\,p_{+1}^{s-a-1}
\end{aligned}
$$

y

$$B = \sum_{a=\text{máx}(r+u-v-1,0)}^{\text{mín}(r+u-1,s-1)} \sum_{i=0}^{\text{mín}(r-1,a)} \binom{r+u-1}{i, r-1-i, a-i, u-a+i} p_{11}^{i}\, p_{10}^{r-i}\, p_{01}^{a-i}\, p_{00}^{u-a+i} \times \binom{s+v-r-u-1}{s-a-1} p_{+0}^{v-u-r+a}\, p_{+1}^{s-a}\,.$$

Siendo s arbitrariamente grande y $r = cs$, con $0 < c < 1$ fijo, y u y v fijos, eventualmente se tendrá que $\text{máx}(r+u-v-1,0) = cs+u-v-1$, $\text{mín}(r+u-1,s-2) = cs+u-1$ y $\text{mín}(r+u-1,s-1) = cs+u-1$, luego A y B pueden ser reescritas en este caso en la forma

$$A = \sum_{a=cs+u-v-1}^{cs+u-1} \sum_{i=0}^{\text{mín}(cs-1,a)} \binom{r+u-1}{i, r-1-i, a-i, u-a+i} p_{11}^{i+1}\, p_{10}^{r-1-i}\, p_{01}^{a-i}\, p_{00}^{u-a+i} \times \binom{s+v-r-u-1}{s-a-2} p_{+0}^{v-u-r+a+1}\, p_{+1}^{s-a-1}$$

y

$$B = \sum_{a=cs+u-v-1}^{cs+u-1} \sum_{i=0}^{\text{mín}(cs-1,a)} \binom{r+u-1}{i, r-1-i, a-i, u-a+i} p_{11}^{i}\, p_{10}^{r-i}\, p_{01}^{a-i}\, p_{00}^{u-a+i} \times \binom{s+v-r-u-1}{s-a-1} p_{+0}^{v-u-r+a}\, p_{+1}^{s-a}\,.$$

Trabajaremos primero con A. Definimos

$$\beta = a - cs - u + v + 1 \qquad \text{y} \qquad j = i - cs + v + 1,$$

donde

$$cs+u-v-1 \le a \le cs+u-1 \qquad \text{y} \qquad 0 \le i \le \text{mín}(cs-1,a).$$

Al variar a de $cs+u-v-1$ a $cs+u-1$ e i de cero a $\text{mín}(cs-1,a)$, las variaciones de β e j deben ser

$$\begin{aligned} &0 = cs+u-v-1-cs-u+v+1 \le \beta \le cs+u-1-cs-u+v+1 = v, \\ &-cs+v+1 \le j \le \text{mín}(cs-1,a) - cs + v + 1. \end{aligned}$$

Otra vez, por ser s arbitrariamente grande y u y v fijos, eventualmente se tendrá $-cs+v+1 < 0$ y puesto que hemos convenido en que un coeficiente binomial se anula si alguna de sus entradas es negativa, se puede suponer, de hecho, que las variaciones de β y j se reducen a

$$0 \le \beta \le v \qquad \text{y} \qquad 0 \le j \le \text{mín}(cs-1,a) - cs + v + 1.$$

Ahora bien, cuando a varía de $cs+u-v-1$ a $cs+u-1$, la variación de j es de cero a

$$\text{mín}(cs-1, cs+u-1) - cs + v + 1 = v.$$

En conclusión, se tiene que

$$0 \le \beta \le v \qquad \text{y} \qquad 0 \le j \le v.$$

En términos de β y j, podemos reexpresar los índices que aparecen en A como

$$\begin{aligned}
i &= j + cs - v - 1 \\
cs - 1 - i &= v - j \\
a - i &= \beta + u - j \\
u - a + i &= j - \beta
\end{aligned}$$

y, por todo lo anterior, se tiene que

$$A = \sum_{\beta=0}^{v} \sum_{j=0}^{v} \binom{cs+u-1}{j+cs-v-1, v-j, \beta+u-j, j-\beta} p_{11}^{j+cs-v}\, p_{10}^{v-j}\, p_{01}^{\beta+u-j}\, p_{00}^{j-\beta} \times \binom{s+v-cs-u-1}{\beta} p_{+0}^{\beta}\, p_{+1}^{s+v-cs-u-\beta}\ .$$

Haciendo el cambio de variables $\alpha = j - \beta$ y $j = j$ en la expresión anterior (donde $0 \leq \beta \leq v$ y $0 \leq j \leq v$), las nuevas variables deben satisfacer $0 \leq \alpha \leq v$ y $\alpha \leq j \leq v$. Con este cambio de variables A queda pues reescrita en la forma

$$A = \sum_{\alpha=0}^{v} \sum_{j=\alpha}^{v} \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} p_{11}^{j+cs-v}\, p_{10}^{v-j}\, p_{01}^{u-\alpha}\, p_{00}^{\alpha} \times \binom{s+v-cs-u-1}{j-\alpha} p_{+0}^{j-\alpha}\, p_{+1}^{s+v-cs-u-(j-\alpha)}\ .$$

Ahora bien, aplicando la convención de que un coeficiente binomial se anula si alguna de sus entradas es negativa, esta suma se reduce a

$$A = \sum_{\alpha=0}^{\min(u,v)} \sum_{j=\alpha}^{v} \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} p_{11}^{j+cs-v}\, p_{10}^{v-j}\, p_{01}^{u-\alpha}\, p_{00}^{\alpha} \times \binom{s+v-cs-u-1}{j-\alpha} p_{+0}^{j-\alpha}\, p_{+1}^{s+v-cs-u-(j-\alpha)}\ .$$

Así es que

$$\begin{aligned}
A = \sum_{\alpha,j} &\binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} \\
&\times \left(\frac{cs}{\lambda+cs}\right)^{j+cs-v} \left(\frac{\lambda-\lambda_1}{cs\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right)^{v-j} \\
&\times \left(\frac{\lambda-c\lambda_2}{cs\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right)^{u-\alpha} \left(\frac{\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+s(\lambda_1+c\lambda_2-\lambda)}{(\lambda_2+s)(\lambda_1+cs)\left(1+\frac{\lambda}{cs}\right)}\right)^{\alpha} \\
&\times \binom{s+v-cs-u-1}{j-\alpha} \left(\frac{\lambda_2}{\lambda_2+s}\right)^{j-\alpha} \left(\frac{s}{\lambda_2+s}\right)^{s+v-cs-u-(j-\alpha)}\ ,
\end{aligned}$$

luego

$$A = \sum_{\alpha,j} \frac{(cs+u-1)\cdots(cs-v+j)(cs-v+j-1)!}{(j+cs-v-1)!(v-j)!(u-\alpha)!\alpha!}$$

$$\times \left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{j+cs-v} \frac{(\lambda-\lambda_1)^{v-j}}{(cs)^{v-j}} \left[\frac{1}{\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{v-j}$$

$$\times \frac{(\lambda-c\lambda_2)^{u-\alpha}}{(cs)^{u-\alpha}} \left[\frac{1}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{u-\alpha} \frac{1}{(cs)^{\alpha}} \left[\frac{\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{\alpha}$$

$$\times \frac{(v+(1-c)s-u-1)!}{(j-\alpha)!(v+(1-c)s-u-1-(j-\alpha))!} \left(\frac{\lambda_2}{\lambda_2+s}\right)^{j-\alpha} \left(\frac{s}{\lambda_2+s}\right)^{-(j-\alpha)} \left(\frac{s}{\lambda_2+s}\right)^{v+(1-c)s-u},$$

de modo que

$$A=\sum_{\alpha,j} \frac{(cs+u-1)\cdots(cs-v+j)}{(v-j)!(u-\alpha)!\alpha!}$$

$$\times \frac{(\lambda-\lambda_1)^{v-j}(\lambda-c\lambda_2)^{u-\alpha}}{(cs)^{v-j}(cs)^{u-\alpha}(cs)^{\alpha}} \left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{j+cs-v} \left[\frac{1}{\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{v-j}$$

$$\times \left[\frac{1}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{u-\alpha} \left[\frac{\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{\alpha}$$

$$\times \frac{(v+(1-c)s-u-1)\cdots(v+(1-c)s-u-(j-\alpha))}{(j-\alpha)!}$$

$$\times \left(\frac{\lambda_2}{s}\right)^{j-\alpha} \left(\frac{1}{\lambda_2+s}\right)^{j-\alpha} \left(\frac{1}{\lambda_2+s}\right)^{-(j-\alpha)} \left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{(1-c)s} \left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{v-u}$$

$$=\sum_{\alpha,j} \frac{(cs+u-1)\cdots(cs-v+j)}{(v-j)!(u-\alpha)!\alpha!}\cdot\frac{(\lambda-\lambda_1)^{v-j}(\lambda-c\lambda_2)^{u-\alpha}}{(cs)^{u+v-j}}$$

$$\times \left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{cs} \left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{j-v} \left[\frac{1}{\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{v-j}$$

$$\times \left[\frac{1}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{u-\alpha} \left[\frac{\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{\alpha}$$

$$\times \frac{(v+(1-c)s-u-1)\cdots(v+(1-c)s-u-(j-\alpha))}{(j-\alpha)!}\cdot\left(\frac{\lambda_2}{s}\right)^{j-\alpha}$$

$$\times \left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{(1-c)s} \left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{v-u},$$

donde

$$(cs+u-1)\cdots(cs-v+j)$$

cuenta con $(cs+u-1)-(cs-v+j-1)=u+v-j$ factores, por lo que

$$(cs+u-1)\cdots(cs-v+j)=(cs)^{u+v-j}\left(1+\frac{u-1}{cs}\right)\cdots\left(1-\frac{v-j}{cs}\right).$$

También

$$(v+(1-c)s-u-1)\cdots(v+(1-c)s-u-(j-\alpha))$$

es un producto de $(v+(1-c)s-u-1)-(v+(1-c)s-u-(j-\alpha)-1)=j-\alpha$ factores, luego

$$(v+(1-c)s-u-1)\cdots(v+(1-c)s-u-(j-\alpha)) = ((1-c)s)^{j-\alpha}\left(1+\frac{v-u-1}{(1-c)s}\right)\cdots\left(1+\frac{v-u-(j-\alpha)}{(1-c)s}\right).$$

Vemos así que

$$\begin{aligned}
A = \sum_{\alpha,j} & \frac{(cs)^{u+v-j}\left(1+\frac{u-1}{cs}\right)\cdots\left(1-\frac{v-j}{cs}\right)}{(v-j)!(u-\alpha)!\alpha!}\cdot\frac{(\lambda-\lambda_1)^{v-j}(\lambda-c\lambda_2)^{u-\alpha}}{(cs)^{u+v-j}} \\
& \times\left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{cs}\left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{j-v}\left[\frac{1}{\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{v-j} \\
& \times\left[\frac{1}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{u-\alpha}\left[\frac{\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{\alpha} \\
& \times\frac{((1-c)s)^{j-\alpha}\left(1+\frac{v-u-1}{(1-c)s}\right)\cdots\left(1+\frac{v-u-(j-\alpha)}{(1-c)s}\right)}{(j-\alpha)!}\cdot\left(\frac{\lambda_2}{s}\right)^{j-\alpha} \\
& \times\left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{(1-c)s}\left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{v-u} \\
= \sum_{\alpha,j} & \frac{\left(1+\frac{u-1}{cs}\right)\cdots\left(1-\frac{v-j}{cs}\right)}{(v-j)!(u-\alpha)!\alpha!}\cdot(\lambda-\lambda_1)^{v-j}(\lambda-c\lambda_2)^{u-\alpha} \\
& \times\left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{cs}\left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{j-v}\left[\frac{1}{\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{v-j} \\
& \times\left[\frac{1}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{u-\alpha}\left[\frac{\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{\alpha} \\
& \times\frac{(1-c)^{j-\alpha}\left(1+\frac{v-u-1}{(1-c)s}\right)\cdots\left(1+\frac{v-u-(j-\alpha)}{(1-c)s}\right)}{(j-\alpha)!}\cdot\lambda_2^{j-\alpha} \\
& \times\left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{(1-c)s}\left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{v-u}.
\end{aligned}$$

Ahora tomaremos el límite cuando $s\to\infty$. La mayoría de los factores se aproximan a uno, por ejemplo,

$$\lim_{s\to\infty}\left(1+\frac{u-1}{cs}\right)=1$$

o, también,

$$\lim_{s\to\infty}\left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{j-v}=1,$$

por lo que sólo serán mostrados aquellos límites que no sea inmediato ver hacia donde convergen. Se tiene pues

$$\lim_{s\to\infty}\left[\frac{1}{\left(1+\frac{\lambda}{cs}\right)}\right]^{cs}=\lim_{s\to\infty}\frac{1}{\left(1+\frac{\lambda}{cs}\right)^{cs}}=e^{-\lambda},$$

$$\begin{aligned}
&\lim_{s\to\infty}\left[\frac{\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)}{\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{j-\beta}\\
&\qquad=\left[\frac{\lim_{s\to\infty}\left(\left(\frac{1}{s}\right)\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+(\lambda_1+c\lambda_2-\lambda)\right)}{\lim_{s\to\infty}\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right]^{j-\beta}=(\lambda_1+c\lambda_2-\lambda)^{j-\beta}
\end{aligned}$$

y

$$\lim_{s\to\infty}\left(\frac{1}{1+\frac{\lambda_2}{s}}\right)^{(1-c)s}=\left[\lim_{s\to\infty}\frac{1}{\left(1+\frac{\lambda_2}{s}\right)^{s}}\right]^{1-c}=e^{-(1-c)\lambda_2}.$$

Así que,

$$\begin{aligned}
\lim_{s\to\infty}A&=\sum_{\alpha,j}e^{-\lambda}\frac{1}{(v-j)!(u-\alpha)!\alpha!}\\
&\qquad\times(\lambda-\lambda_1)^{v-j}(\lambda-c\lambda_2)^{u-\alpha}(\lambda_1+c\lambda_2-\lambda)^{\alpha}\cdot\frac{(1-c)^{j-\alpha}}{(j-\alpha)!}\lambda_2^{j-\alpha}e^{-(1-c)\lambda_2}\\
&=\sum_{\alpha,j}e^{-\lambda-(1-c)\lambda_2}\frac{(\lambda-\lambda_1)^{v-j}}{(v-j)!}\cdot\frac{(\lambda-c\lambda_2)^{u-\alpha}}{(u-\alpha)!}\cdot\frac{(\lambda_1+c\lambda_2-\lambda)^{\alpha}}{\alpha!}\cdot\frac{(\lambda_2(1-c))^{j-\alpha}}{(j-\alpha)!}\\
&=e^{-\lambda-(1-c)\lambda_2}\sum_{\alpha=0}^{\min(u,v)}\frac{(\lambda-c\lambda_2)^{u-\alpha}}{(u-\alpha)!}\cdot\frac{(\lambda_1+c\lambda_2-\lambda)^{\alpha}}{\alpha!}\sum_{j=\alpha}^{v}\frac{(\lambda-\lambda_1)^{v-j}}{(v-j)!}\cdot\frac{(\lambda_2(1-c))^{j-\alpha}}{(j-\alpha)!}\\
&=e^{-\lambda-(1-c)\lambda_2}\sum_{\alpha=0}^{\min(u,v)}\frac{(\lambda-c\lambda_2)^{u-\alpha}}{(u-\alpha)!}\cdot\frac{(\lambda_1+c\lambda_2-\lambda)^{\alpha}}{\alpha!}\cdot\frac{[\lambda-\lambda_1+\lambda_2(1-c)]^{v-\alpha}}{(v-\alpha)!},
\end{aligned}$$

donde la última igualdad está justificada por el *teorema del binomio* aplicado con $a=\lambda-\lambda_1$, $b=\lambda_2(1-c)$ y $n=v-\alpha$, es decir,

$$[\lambda-\lambda_1+\lambda_2(1-c)]^{v-\alpha}=\sum_{k=0}^{v-\alpha}\binom{v-\alpha}{k}(\lambda-\lambda_1)^{v-\alpha-k}(\lambda_2(1-c))^{k}.$$

Reexpresando el binomio seleccionando con $j=\alpha+k$ resulta

$$\begin{aligned}
[\lambda-\lambda_1+\lambda_2(1-c)]^{v-\alpha}&=\sum_{j=\alpha}^{v}\binom{v-\alpha}{j-\alpha}(\lambda-\lambda_1)^{v-j}(\lambda_2(1-c))^{j-\alpha}\\
&=\sum_{j=\alpha}^{v}\frac{(v-\alpha)!}{(j-\alpha)!(v-j)!}(\lambda-\lambda_1)^{v-j}(\lambda_2(1-c))^{j-\alpha}\\
&=(v-\alpha)!\sum_{j=\alpha}^{v}\frac{(\lambda-\lambda_1)^{v-j}}{(v-j)!}\cdot\frac{(\lambda_2(1-c))^{j-\alpha}}{(j-\alpha)!},
\end{aligned}$$

de modo que

$$\sum_{j=\alpha}^{v} \frac{(\lambda-\lambda_1)^{v-j}}{(v-j)!} \cdot \frac{(\lambda_2(1-c))^{j-\alpha}}{(j-\alpha)!} = \frac{[\lambda-\lambda_1+\lambda_2(1-c)]^{v-\alpha}}{(v-\alpha)!}.$$

Tomando $\lambda_{11} = \lambda_1 + c\lambda_2 - \lambda$, $\lambda_{10} = \lambda - c\lambda_2$ y $\lambda_{01} = \lambda - \lambda_1 + \lambda_2(1-c)$ se tiene que

$$\lim_{s\to\infty} A = e^{-(\lambda_{11}+\lambda_{10}+\lambda_{01})} \sum_{\alpha=0}^{\min(u,v)} \frac{\lambda_{10}^{u-\alpha}}{(u-\alpha)!} \cdot \frac{\lambda_{11}^{\alpha}}{\alpha!} \cdot \frac{\lambda_{01}^{v-\alpha}}{(v-\alpha)!}\,,$$

es decir, obtenemos (2.36) que es la distribución de Poisson bivariada PB dada por

$$\sum_{\alpha=0}^{\min(u,v)} \frac{\lambda_{11}^{\alpha}\,\lambda_{10}^{u-\alpha}\,\lambda_{01}^{v-\alpha}}{\alpha!\,(u-\alpha)!\,(v-\alpha)!} e^{-(\lambda_{11}+\lambda_{10}+\lambda_{01})}.$$

Ahora trabajamos con la parte B. El desarrollo es análogo al que se ha realizado en lo anterior salvo algunos detalles que serán mostrados en lo siguiente. Empezamos con la misma restricción sobre los argumentos de las sumas, por lo que seleccionando β y j como antes, los índices en B son como los que se obtuvieron para A, así que escribimos

$$B = \sum_{\beta=0}^{v}\sum_{j=0}^{v} \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} p_{11}^{j+cs-v-1}\, p_{10}^{v-j+1}\, p_{01}^{u-\alpha}\, p_{00}^{\alpha} \\ \times \binom{s+v-cs-u-1}{\alpha} p_{+0}^{j-\alpha}\, p_{+1}^{s+v-cs-u-(j-\alpha)}\,.$$

Es posible apreciar que, dada tal expresión, A y B difieren únicamente en los exponentes de p_{11} y p_{10}. Luego, haciendo el cambio de variables $\alpha = j-\beta$ y $j = j$, como en la parte anterior, se obtiene la siguiente expresión

$$B = \sum_{\alpha=0}^{v}\sum_{j=0}^{v} \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} p_{11}^{j+cs-v-1}\, p_{10}^{v-j+1}\, p_{01}^{u-\alpha}\, p_{00}^{\alpha} \\ \times \binom{s+v-cs-u-1}{\alpha} p_{+0}^{j-\alpha}\, p_{+1}^{s+v-cs-u-(j-\alpha)}\,.$$

Por convención un coeficiente binomial se anula si alguna de sus entradas es negativa, de modo que la suma se reduce a

$$B = \sum_{\alpha=0}^{\min(u,v)}\sum_{j=0}^{v} \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} p_{11}^{j+cs-v-1}\, p_{10}^{v-j+1}\, p_{01}^{u-\alpha}\, p_{00}^{\alpha} \\ \times \binom{s+v-cs-u-1}{\alpha} p_{+0}^{j-\alpha}\, p_{+1}^{s+v-cs-u-(j-\alpha)}\,.$$

Así que

$$
\begin{aligned}
B = \sum_{\alpha,j} & \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} \\
& \times \left(\frac{cs}{\lambda+cs}\right)^{j+cs-v-1} \left(\frac{\lambda-\lambda_1}{cs\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right)^{v-j+1} \\
& \times \left(\frac{\lambda-c\lambda_2}{cs\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right)^{u-\alpha} \left(\frac{\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+s(\lambda_1+c\lambda_2-\lambda)}{(\lambda_2+s)(\lambda_1+cs)\left(1+\frac{\lambda}{cs}\right)}\right)^{\alpha} \\
& \times \binom{s+v-cs-u-1}{j-\alpha} \left(\frac{\lambda_2}{\lambda_2+s}\right)^{j-\alpha} \left(\frac{s}{\lambda_2+s}\right)^{s+v-cs-u-(j-\alpha)} .
\end{aligned}
$$

Entonces,

$$
\begin{aligned}
B = & \left(\frac{cs}{\lambda+cs}\right)^{-1} \left(\frac{\lambda-\lambda_1}{cs\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right) \\
& \times \sum_{\alpha,j} \binom{cs+u-1}{j+cs-v-1, v-j, u-\alpha, \alpha} \\
& \times \left(\frac{cs}{\lambda+cs}\right)^{j+cs-v} \left(\frac{\lambda-\lambda_1}{cs\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right)^{v-j} \\
& \times \left(\frac{\lambda-c\lambda_2}{cs\left(1+\frac{\lambda_2}{s}\right)\left(1+\frac{\lambda}{cs}\right)}\right)^{u-\alpha} \left(\frac{\lambda_1\lambda_2\left(2+\frac{\lambda}{cs}\right)+s(\lambda_1+c\lambda_2-\lambda)}{(\lambda_2+s)(\lambda_1+cs)\left(1+\frac{\lambda}{cs}\right)}\right)^{\alpha} \\
& \times \binom{s+v-cs-u-1}{j-\alpha} \left(\frac{\lambda_2}{\lambda_2+s}\right)^{j-\alpha} \left(\frac{s}{\lambda_2+s}\right)^{s+v-cs-u-(j-\alpha)} .
\end{aligned}
$$

Nótese que B puede ser escrito en términos de A recogiendo los factores que los distinguen, es decir,

$$
\begin{aligned}
B &= \left(\frac{cs}{\lambda+cs}\right)^{-1} \left(\frac{\lambda-\lambda_1}{cs\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right) A \\
&= \left[\frac{1}{(1+\frac{\lambda}{cs})}\right]^{-1} \left(\frac{\lambda-\lambda_1}{cs}\right) \left[\frac{1}{\left(1+\frac{\lambda_1}{cs}\right)\left(1+\frac{\lambda}{cs}\right)}\right] A.
\end{aligned}
$$

Tomando el límite cuando $s \to \infty$ y dado que

$$\lim_{s\to\infty} \left(\frac{1}{cs}\right) = 0\,,$$

se tiene

$$\lim_{s\to\infty} B = 0.$$

Así es que, como $P[U=u, V=v] = A + B$,

$$\lim_{s\to\infty} P[U=u, V=v] = \lim_{s\to\infty} (A+B).$$

Por todo el desarrollo anterior el límite anterior existe y es igual a la suma de los límites, es decir,

$$\lim_{s\to\infty} P[U=u, V=v] = \lim_{s\to\infty} A + \lim_{s\to\infty} B = \sum_{\alpha} \frac{\lambda_{11}^{\alpha}\, \lambda_{10}^{u-\alpha}\, \lambda_{01}^{v-\alpha}}{\alpha!\,(u-\alpha)!\,(v-\alpha)!} e^{-(\lambda_{11}+\lambda_{10}+\lambda_{01})}\,.$$

Por lo tanto,

$$\lim_{s\to\infty} P[U=u, V=v] = P[U^*=u, V^*=v]$$

donde (U^*, V^*) es un vector aleatorio con la distribución de Poisson bivariada PB dada por (2.36). ∎

3.3.2. Ejemplos

Cualquiera de los ejemplos que se mostraron para la distribución GB puede ser ampliado al modelo de la distribución BNB dada en el Teorema y Definición 3.17. El Ejemplo 3.14 puede ser analizado de modo tal que U y V sean el número posible de lanzamientos que realiza J_1 y J_2, respectivamente, para observar r veces “águila” y s veces “sol”; en el Ejemplo 3.15 del juego de la ruleta americana, se pueden considerar a U y V como el número de jugadas antes de que la bola haya caído r veces en una casilla roja y s veces en un número par, respectivamente; finalmente, en el Ejemplo 3.16 suponga que se sospecha que el lote de producción es “malo” debido a fallas significativas en las máquinas que ensamblan los productos y se pide que sean extraídas dos muestras, una de r productos que presenten fallas del tipo I y s que presenten fallas del tipo II, para determinar las posibles causas de que el lote haya salido mal, en este caso U y V son el número de pruebas que posiblemente deberán ser realizadas para obtener una muestra de r productos con fallas del tipo I y otra muestra de s con fallas del tipo II.

A continuación se muestran otro par de ejemplos que pueden ser modelados con la distribución BNB.

Ejemplo 3.19 Suponga que se deben seleccionar a dos tipos de pacientes clasificados de acuerdo a su tipo de sangre, digamos A o B, para ser incluidos en un ensayo clínico. Si U y V son el número de pacientes que deben ser examinados para obtener una muestra de r con el tipo A y otra muestra de s con el tipo B, entonces (U, V) tiene la distribución BNB con $p_{00} = p_{11} = 0$, pues no es posible que un paciente presente los dos tipos de sangre o que no tenga alguno de los dos (p_{11} es la probabilidad de observar ambos atributos en una misma componente, en este caso se trataría de la probabilidad de observar que un paciente presenta los dos tipos de sangre). ◄

Ejemplo 3.20 Suponga que una persona tiene como propósito proponer un par de peticiones al gobierno en beneficio de la sociedad. Para comprobar que las peticiones son de interés significativo para la sociedad requiere obtener r firmas para la primera y s para la segunda. Suponga que cada persona consultada toma su decisión independientemente y que podría o no firmar una o ambas peticiones, en este caso, si U y V son el número de personas que posiblemente deberán ser consultadas para obtener la cantidad de firmas mencionada, entonces (U, V) tiene la distribución BNB. ◄

Apéndice

Códigos en Wolfram Mathematica

En este apartado solamente se mostrará una forma en la que se deberá escribir el código que calcula $P[U = u, V = v]$ en *Mathematica* para las distintas distribuciones bivariadas mostradas en esta tesis. Si el lector no está familiarizado con el lenguaje se le recomienda, por ejemplo, consultar [15].

Para calcular $P[U = u, V = v]$, lo primero que habrá que hacer es asignar valores a las variables que se repiten en los cálculos, por ejemplo, el tamaño de la muestra y los parámetros de la distribución; luego hay que escribir la expresión general para $P[U = u, V = v]$ como una función de dos variables (u, v) y, finalmente, asignar a otra variable la cantidad que se desea calcular $p = P[u, v]$.

Distribución binomial bivariada.

Para el Ejemplo 2.16 se calculó $P[U = 15, V = 10]$ en la siguiente forma:

In[1]:= $n = 20$

$p11 = 0.44$

$p10 = 0.32$

$p01 = 0.07$

$p00 = 0.17$

Out[1]:= 20

Out[2]:= 0.44

Out[3]:= 0.32

Out[4]:= 0.07

Out[5]:= 0.17

In[6]:= $P[u\_, v\_] := \text{Sum}[\text{Multinomial}[a, u - a, v - a, n - u - v + a] * p11\hat{}(a) * p10\hat{}(u - a)$
$* p01\hat{}(v - a) * p00\hat{}(n - u - v + a), \{a, 0, \text{Min}[u, v]\}]$

In[7]:= $p = P[15, 10]$

Out[7]:= 0.0365719

donde la expresión dada para $P[u\_, v\_]$ corresponde a (2.5). ◄

Distribución binomial bivariada generalizada.

En el Ejemplo 2.17 fue requerida $P[U^* = 70, V^* = 60]$, el código es:

In[1]:= $k = 50$

$n = 80$

$m = 75$

$p11 = 0.75$

$p10 = 0.07$

$p01 = 0.10$

$p00 = 0.08$

$p1y = p10 + p11$

$p0y = p00 + p01$

$px1 = p01 + p11$

$px0 = p00 + p10$

Out[1]:= 50

Out[2]:= 80

Out[3]:= 75

Out[4]:= 0.75

Out[5]:= 0.07

Out[6]:= 0.10

Out[7]:= 0.08

Out[8]:= 0.82

Out[9]:= 0.18

Out[10]:= 0.85

Out[11]:= 0.15

In[12]:= $P[u\_, v\_] := \text{Sum}[\text{Multinomial}[a, u - i - a, v - j - a, k - u - v + i + j + a]$

$* p11\text{^}(a) * p10\text{^}(u - i - a) * p01\text{^}(v - j - a) * p00\text{^}(k - u - v + i + j + a)$

$* \text{Binomial}[n - k, i] * p1y\text{^}(i) * p0y\text{^}(n - k - i)$

$* \text{Binomial}[m - k, j] * px1\text{^}(j) * px0\text{^}(m - k - j),$

$\{i, 0, n - k\}, \{j, 0, m - k\}, \{a, 0, \text{Min}[u - i, v - j]\}]$

In[13]:= $p = P[70, 60]$

Out[13]:= 0.0019433

donde la expresión dada para $P[u\_, v\_]$ corresponde a (2.7) y se tienen $p1y = p_{1+}$, $p0y = p_{0+}$, $px1 = p_{+1}$ y $px0 = p_{+0}$. ◄

Distribución de Poisson bivariada.

En el siguiente código $P[u\_, v\_]$ corresponde a (2.36), requiere como cantidades de entrada a λ_{01}, λ_{10} y λ_{11}. Utilizando la información del Ejemplo 2.28 puede escribir:

$$
\begin{aligned}
\textit{In[1]:=} \quad & lambda11 = 25 \\
& lambda10 = 20 \\
& lambda01 = 15 \\
\textit{Out[1]:=} \quad & 25 \\
\textit{Out[2]:=} \quad & 20 \\
\textit{Out[3]:=} \quad & 15 \\
\textit{In[4]:=} \quad & P[u\_, v\_] := \mathrm{Sum}[(lambda11\hat{}(a) * lambda10\hat{}(u-a) * lambda01\hat{}(v-a)) \\
& \qquad /((a!) * ((u-a)!) * ((v-a)!)), \{a, 0, \mathrm{Min}[u, v]\}] \\
& \qquad * \mathrm{Exp}[-lambda01 - lambda10 - lambda11] \\
\textit{In[5]:=} \quad & p = P[30, 35] \\
\textit{Out[5]:=} \quad & 0.000269567 \blacktriangleleft
\end{aligned}
$$

Distribución hipergeométrica bivariada.

De acuerdo a la información del Ejemplo 3.7 puede escribir:

$$
\begin{aligned}
\textit{In[1]:=} \quad & M = 70 \\
& N11 = 31 \\
& N10 = 22 \\
& N01 = 5 \\
& N00 = 12 \\
& n = 20 \\
\textit{Out[1]:=} \quad & 70 \\
\textit{Out[2]:=} \quad & 31 \\
\textit{Out[3]:=} \quad & 22 \\
\textit{Out[4]:=} \quad & 5 \\
\textit{Out[5]:=} \quad & 12 \\
\textit{Out[6]:=} \quad & 20 \\
\textit{In[7]:=} \quad & P[u\_, v\_] := \mathrm{Sum}[\mathrm{Binomial}[N10, u-a] * \mathrm{Binomial}[N01, v-a] * \mathrm{Binomial}[N11, a] \\
& \qquad * \mathrm{Binomial}[N00, n-u-v+a], \{a, 0, \mathrm{Min}[u, v]\}]/\mathrm{Binomial}[M, n] \\
\textit{In[8]:=} \quad & p = P[15, 10] \\
\textit{Out[8]:=} \quad & 0.0505595
\end{aligned}
$$

donde $P[u\_, v\_]$ corresponde a (3.1), $M = N$ es el tamaño de la población, $N11 = N_{11}$, $N10 = N_{10}$, $N01 = N_{01}$ y $N00 = N_{00}$. ◄

Distribución geométrica bivariada

Así como en las distribuciones anteriores habrá que escribir primero los datos de entrada que se requieren para hacer los cálculos, en este caso, de acuerdo con (3.17) y (3.18) son necesarios p_{ij}, p_{i+} y p_{+i} para cada $i, j \in \{0, 1\}$ (puede hacerlo como en el caso de la binomial bivarada generalizada). El código entonces para $P[U = u, V = v]$ puede ser escrito en la siguiente forma:

$$
\begin{aligned}
&\textit{In[1]:=} && (****\,\text{Caso} \quad u < v ****)\\
& && P1[u\_, v\_] := p00\hat{}(u) * p10 * px0\hat{}(v - u - 1) * px1\\
&\textit{In[2]:=} && (****\,\text{Caso} \quad v < u ****)\\
& && P2[u\_, v\_] := p00\hat{}(v) * p01 * p0y\hat{}(u - v - 1) * p1y\\
&\textit{In[3]:=} && (****\,\text{Caso} \quad u = v ****)\\
& && P3[u\_, v\_] := p00\hat{}(u) * p11
\end{aligned}
$$

mientras que para $P[U \geq u, V \geq v]$ puede escribir

$$
\begin{aligned}
&\textit{In[4]:=} && (****\,\text{Caso} \quad u < v ****)\\
& && P4[u\_, v\_] := p00\hat{}(u) * px0\hat{}(v - u)\\
&\textit{In[5]:=} && (****\,\text{Caso} \quad v < u ****)\\
& && P5[u\_, v\_] := p00\hat{}(v) * p0y\hat{}(u - v) \blacktriangleleft
\end{aligned}
$$

Distribución binomial negativa bivariada.

En este caso debe ingresar las mismas cantidades que requiere la geométrica bivariada junto con r y s, según el Teorema y Definición 3.17. Serán mostrados tres casos, los dos restantes son enteramente similares.

Para $P[U = u, V = v]$:

$$
\begin{aligned}
&\textit{In[1]:=} \quad (****\,\text{Caso} \quad r + u < s + v ****)\\
&\qquad P1[u\_, v\_] := \text{Sum}[\text{Multinomial}[i, r - 1 - i, a - i, u - a + i] * p11\hat{}(i + 1)\\
&\qquad\quad * p10\hat{}(r - 1 - i) * p01\hat{}(a - i) * p00\hat{}(u - a + i)\\
&\qquad\qquad * \text{Binomial}[s + v - r - u - 1, s - a - 2] * px0\hat{}(v - u - r + a + 1) * px1\hat{}(s - a - 1),\\
&\qquad\qquad\quad \{a, \text{Max}[r + u - 1 - v, 0], \text{Min}[s - 2, r + u - 1]\}, \{i, 0, \text{Min}[r - 1, a]\}]\\
&\qquad\quad + \text{Sum}[\text{Multinomial}[i, r - 1 - i, a - i, u - a + i] * p11\hat{}(i)\\
&\qquad\qquad * p10\hat{}(r - i) * p01\hat{}(a - i) * p00\hat{}(u - a + i)\\
&\qquad\qquad\quad * \text{Binomial}[s + v - r - u - 1, s - a - 1] * px0\hat{}(v - u - r + a) * px1\hat{}(s - a),\\
&\qquad\qquad\qquad \{a, \text{Max}[r + u - 1 - v, 0], \text{Min}[s - 1, r + u - 1]\}, \{i, 0, \text{Min}[r - 1, a]\}]\\
&\textit{In[2]:=} \quad (****\,\text{Caso} \quad s + v = r + u ****)\\
&\qquad P2[u\_, v\_] := \text{Sum}[\text{Multinomial}[i, r - 1i, s - 1 - i, u - s + 1 + i] * p11\hat{}(i + 1) * p10\hat{}(r - 1 - i)\\
&\qquad\qquad * p01\hat{}(s - 1 - i) * p00\hat{}(u - s + 1 + i), \{i, 0, \text{Min}[r - 1, s - 1]\}]
\end{aligned}
$$

Para $P[U \geq u, V \geq v]$:

$$\begin{aligned}
\textit{In[3]:=} \quad & (****\text{Caso} \quad r+u \leq s+v****) \\
& P3[u_, v_] := \text{Sum}[\text{Multinomial}[i, l-i, a-i, r+u-1-a-l+i] \\
& \qquad * p11^\wedge(i) * p10^\wedge(l-i) * p01^\wedge(a-i) * p00^\wedge(r+u-1-a-l+i) \\
& \qquad\quad * Binomial[s+v-r-u, m] * px0^\wedge(s+v-r-u-m) * px1^\wedge(m), \\
& \qquad\qquad \{a, 0, s-1\}, \{l, 0, r-1\}, \{i, 0, \text{Min}[a, l]\}, \{m, 0, s-1-a\}] \blacktriangleleft
\end{aligned}$$

Nota. Los coeficientes de correlación mostrados en los ejemplos fueron calculados en *Mathematica* siguiendo de forma similar los códigos mostrados en este apartado. ◁

Conclusiones

Se ha definido una familia de distribuciones bivariadas discretas, a partir de la distribución de Bernoulli bivariada, siguiendo un camino análogo al caso univariado para obtener las distribuciones binomial, de Poisson, hipergeométrica, geométrica y binomial negativa, desde la distribución de Bernoulli. También, se ha demostrado que las siguientes aproximaciones se preservan en el caso bivariado: aproximación normal a las distribuciones binomial y de Poisson, aproximación binomial a la distribución hipergeométrica, aproximación de Poisson a las distribuciones binomial y binomial negativa. Las demostraciones, aquí desarrolladas, de dichas aproximaciones se han hecho de forma similar y esencialmente con las mismas ideas que funcionan para las del caso univariado, basta con hacer un ejercicio de comparación entre ellas para estar convencido de ello.

Cualquiera de los ejemplos clásicos o problemas que se proponen en la literatura para trabajar las distribuciones en una variable pueden ser ampliados al caso bivariado con la finalidad de incluir más atributos en el estudio de una situación en la que se esté interesado. El trabajado desarrollado es una muestra de la naturalidad con la que es posible extender las distribuciones y los resultados de aproximación.

Los modelos que aquí se han mostrado son de amplia utilidad para todo aquel interesado en estudiar experimentos en los que cada observación corresponde a un elemento clasificado por dos atributos distinguibles, de manera tal que cada elemento puede o no presentar uno o ambos atributos. Estas observaciones pueden ser modeladas con la distribución de Bernoulli bivariada y luego, entonces, trabajar con cualquiera de las distribuciones expuestas y, por ejemplo, obtener el grado de dependencia entre las variables aleatorias mediante los coeficientes de correlación también proporcionados.

Todo lo anterior puede extenderse sin dificultades, siguiendo el mismo orden de ideas, al caso general multivariado, esperamos se pueda hacer esto en detalle en un futuro trabajo.

Bibliografía

[1] Edward J. Dudewicz y Satya N. Mishra, *Modern Mathematical Statistics*, John Wiley & Sons, 1988.

[2] Paul G. Hoel, Sidney C. Port y Charles J. Stone, *Introduction to Probability Theory*, Houghton Mifflin Company, 1971.

[3] Dennis D. Wackerly, William Mendenhall III y Richard L. Scheaffer, *Estadística Matemática con Aplicaciones*, Cengage Learning Editores, 2010.

[4] Geoffrey R. Grimmett y David R. Stirzaker, *Probability and Random Processes*, Oxford University Press, 2001.

[5] Robert V. Hogg, Joseph W. McKean y Allen T. Craig, *Introduction to Mathematical Statistics*, Pearson Education, 2005.

[6] Franklin A. Graybill, *An Introduction to Linear Statistical Models Vol. 1*, McGraw-Hill, 1961.

[7] Alexander M. Mood, Franklin A. Graybill y Duane C. Boes, *Introduction to the Theory of Statistics*, McGraw-Hill, 1974.

[8] Subrahmaniam Kocherlakota y Kathleen Kocherlakota, *Bivariate Discrete Distributions*, Marcel Dekker, Inc, 1992.

[9] Albert W. Marshall e Ingram Olkin, *A Family of Bivariate Distributions Generated by the Bivariate Bernoulli Distribution*, Journal of the American Statistical Association, Vol. 80, No. 390, p. 332-338, 1985.

[10] Albert W. Marshall e Ingram Olkin, *A Multivariate Exponential Distribution*, Journal of the American Statistical Association, Vol. 62, No. 317, p. 30-44, 1967.

[11] Wicksell S. D, *Some Theorems in the Theory of Probability, With Special Reference to Their Importance in the Theory of Homograde Correlation*, Meddelanden fran Lunds Astronomiska Observatorium Series I, Vol. 74, p. 1-49, 1916.

[12] Wicksell S. D, *Contributions to the Analytical Theory of Sampling*, Meddelanden fran Lunds Astronomiska Observatorium Series I, Vol. 102, p. 1-46, 1922.

[13] Hamdan M. A, y Jensen D. R, *A Bivariate Binomial Distribution and Some Applications*, Australian Journal of Statistics, Vol. 18, No. 3, p. 163-169, 1976.

[14] Holgate P, *Estimation for the bivariate Poisson distribution*, Biometrika, Vol. 51, No. 1/2, p. 241-245, 1964.

[15] Stephen Wolfram, *The Mathematica Book*, Wolfram Media, 2003.