

HÉCTOR FRANCISCO
PONCE RENOVA

ESTADÍSTICA ELEMENTAL

PARA INVESTIGACIÓN EDUCATIVA:
PROBABILIDAD, DISTRIBUCIONES
Y CORRELACIÓN

UNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ

UNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ

Juan Ignacio Camargo Nassar

Rector

Daniel Constandse Cortez

Secretario General

Alonso Morales Muñoz

Director del Instituto de Ciencias Sociales y Administración

Jesús Meza Vega

Director General de Comunicación Universitaria

HÉCTOR FRANCISCO PONCE RENOVA

ESTADÍSTICA
ELEMENTAL PARA
INVESTIGACIÓN
EDUCATIVA

Probabilidad, distribuciones y correlación

UNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ

DR © Héctor Francisco Ponce Renova

© Universidad Autónoma de Ciudad Juárez
Avenida Plutarco Elías Calles 1210
Fovissste Chamizal, CP 32310
Ciudad Juárez, Chihuahua, México
Tels. +52 (656) 688 2100 al 09

Ponce Renova, Héctor Francisco

Estadística elemental para investigación educativa: probabilidad, distribuciones y correlación / Héctor Francisco Ponce Renova.— Primera edición -- Ciudad Juárez, Chihuahua, México: Universidad Autónoma de Ciudad Juárez, 2020.
149 páginas; 22 centímetros.

ISBN: 978-607-520-383-6

Contenido: Introducción.— Resumen.— Capítulo 1: El área y la probabilidad.— Capítulo 2: Distribución binomial y su relación con la distribución normal y estandarizada.— Capítulo 3: Correlación.— Capítulo 4: Coeficiente de Correlación r de Spearman.— Índice de tablas .— Índice de figuras.

Estadística aplicada — Educación
Estadística — Investigación educativa
Estadística — Estudio y enseñanza

LC – HA29.5S63 P65 2020

La edición, diseño y producción editorial de este documento estuvo a cargo de la Dirección General de Comunicación Universitaria, a través de la Subdirección de Editorial y Publicaciones.

Coordinación editorial:

Mayola Renova González

Cuidado editorial:

Subdirección de Editorial y Publicaciones

Diseño de portada y diagramación:

Karla María Rascón

Primera edición, 2020

<https://elibros.uacj.mx>



Índice

Introducción	11
Resumen	11
Descripción de la obra y su relación con la investigación educativa	12
Dirigido a lectoras y lectores	14
Conocimientos previos	14
Aporte principal de la obra.....	14
Razón para la creación y publicación de la obra	15
Objetivos de aprendizaje de la obra	16
Términos clave	17

Capítulo 1

El área y la probabilidad	19
a. Probabilidad	20
b. Permutaciones	24
c. Combinatorias.....	26
Opinión del autor.....	29

Capítulo 2

Distribución binomial y su relación con la distribución normal y estandarizada	31
a. Ejemplo para construir la distribución binomial.....	32
b. Estadísticas de la distribución binomial.....	37
c. Ejemplo de combinatorias y probabilidades con dados	38
d. Características de la distribución binomial.....	41
e. Curva de la distribución normal ...	42
f. Distribución normal estandarizada.....	50
Preguntas para reflexionar	54
Opinión del autor.....	55

Capítulo 3

Correlación	57
a. Correlaciones r de Pearson y policóricas.....	58
b. Características del r	62
c. Fórmulas del Coeficiente r	66
d. Coeficiente de Determinación (r^2) (<i>Determination Coefficient</i>) y el r	75
e. Requisitos para usar el Coeficiente r de Pearson.....	81
f. Intervalo de confianza para la r de Pearson	89
g. Test de la hipótesis nula	103
g. Un Test t de la hipótesis nula para la r de Pearson.....	116
i. Resumen de Test de Significancia Estadística para una y dos colas.....	129
Preguntas para reflexionar	131
Opinión del autor.....	132

Capítulo 4

Coeficiente de Correlación ρ de Spearman.....	135
Preguntas para reflexionar	140
Opinión del autor.....	141
Referencias	143

Índice de tablas

Tabla 1.1 Permutaciones con tres personas y tres puestos.....	24
Tabla 1.2 Combinatorias para el comité	26
Tabla 2.1 Probabilidades	34
Tabla 2.2 Combinatorias con una sola cara	35
Tabla 2.3 Probabilidades y combinatorias	35
Tabla 2.4 Dos lanzamientos de dados en combinatorias	40
Tabla 2.5 Valores atípicos y asimetría.....	48
Tabla 2.6 Valores z y su interpretación.....	52
Tabla 3.1 Datos en pares para usarlos en correlaciones.....	63
Tabla 3.2 Fórmula del r de Pearson con valores z	68
Tabla 3.3 Fórmula del r de Pearson con la desviación de los escores.....	69
Tabla 3.4 Fórmula del escore crudo.	71
Tabla 3.5 Usando la covarianza para encontrar el r de Pearson.....	72
Tabla 3.6 Tamaños del r y sus IC del 95% y 99%	93
Tabla 3.7 IC del 95% y 99%	95
Tabla 3.8 Transformación de Fisher de la r a la z	97
Tabla 3.9 Errores tipo I y II en la correlación de Pearson	108
Tabla 3.10 Pasos para poner a prueba una hipótesis nula.....	116
Tabla 3.11 Extracto de una Tabla de distribución t	118
Tabla 4.1 Mejores prácticas de Brophy (2001).....	138
Tabla 4.2 Resultados de las evaluaciones de mejores prácticas de Brophy (2001).....	139

Índice de figuras

Figura 1.1 Área de un cuadrado.....	22
Figura 1.2 Probabilidad de cara o cruz	23
Figura 2.1 Distribución binomial de la moneda.....	36
Figura 2.2 Distribución binomial de la moneda con 100 tiradas	38
Figura 2.3 Curva normal.....	43
Figura 2.4 Curtosis.....	45
Figura 2.5 Asimetría.....	46
Figura 2.6. Observaciones atípicas ...	49
Figura 2.7 Distribución estandarizada y sus proporciones de áreas y de probabilidad	51
Figura 2.8 Puntos en la curva estandarizada.....	53
Figura 3.1 Datos bivariados.....	59
Figura 3.2 Algunos posibles escenarios para usar el r de Pearson	65
Figura 3.3 Histograma del Tiempo en completar el test y Puntajes en un test de matemáticas (r)	74
Figura 3.4 Varianza explicada.....	76
Figura 3.5 r con evaluación del idioma español con el nivel educativo de las madres y padres, y el ingreso familiar	79
Figura 3.6 Varianza explicada por dos variables independientes en la dependiente.....	79
Figura 3.7 Multicolinealidad con r^2	81
Figura 3.8 Distribución bivariada	83
Figura 3.9 Distribución normal bivariada.....	84
Figura 3.10 Normalidad con áreas de rechazo de la hipótesis nula y valores atípicos.....	85
Figura 3.11 Homocedasticidad y heterocedasticidad.....	88
Figura 3.12 Diferentes niveles de confianza en un IC	89
Figura 3.13 Distribuciones de muestras asimétricas del r de Pearson	92
Figura 3.14 IC de los tres tamaños de efecto propuestos por Cohen (1988)	94
Figura 3.15 Relación entre la asimetría del IC y el tamaño de la mues- tra.....	96
Figura 3.16 Distribución de muestras e IC del 95%.....	100
Figura 3.17 Intervalo de confianza asimétrico del 95%.....	103
Figura 3.18 Distribución de muestras.....	104

Figura 3.19 Distribución de muestras asimétricas y aproximadamente normal del r de Pearson	105
Figura 3.20 Distribución de muestras con áreas de rechazo y no-rechazo de la hipótesis nula para la r de Pearson para una y dos colas para Test t y z	110
Figura 3.21 Test t de dos colas.....	120
Figura 3.22 Distribución z para el IC del 95%.....	122
Figura 3.23 Distribución z	125
Figura 3.24 Test z de una hipótesis nula con una sola cola.....	127
Figura 3.25 Distribución z para el IC del 95%.....	128
Figura 3.26 Valores críticos z y sus correspondientes niveles alfa.....	131
Figura 4.1 Correlación r de Spearman.....	137

Introducción

Resumen

LA OBRA CONTIENE CUATRO CAPÍTULOS CON UNA SECCIÓN de *Preguntas para reflexionar* (interrogantes que tratan de ir más *allá* del contenido propio de este libro) y otra con la *Opinión del autor* (que trata de explicar y asociar el contenido de cada capítulo con la experiencia y perspectiva de este profesor-investigador). En el primer capítulo, se habla de la probabilidad para conceptualizarla como un área, junto con las *propiedades de la sumatoria y la multiplicación*. Para continuar con la probabilidad, se explican los conceptos de *permutación y combinación*. Los temas de combinatorias y probabilidad se entrelazan en el segundo capítulo para luego crear la distribución binomial, la cual puede ser dividida en porciones cuyas dimensiones varían, y, con ello, también cambian las probabilidades de dichas porciones. Asimismo, en este segundo capítulo también se explica cómo la distribución binomial es la base de la distribución normal, así como de la distribución estandarizada. En el tercer capítulo, se muestra la correlación de Pearson principalmente, pero también se habla de las correlaciones *poli-córicas*, que son más apropiadas para escalas tipo Likert. Más en detalle, se explica cómo calcular el r de Pearson con diferentes fórmulas, intervalos de confianza, planteamiento de preguntas de investigación e hipótesis, y varianza explicada, así como el proceso de poner la hipótesis nula

a prueba. En el cuarto y último capítulo, se trata la correlación r de Spearman, que se puede usar cuando *no* se cumplen algunos de los supuestos de la r de Pearson. Todos estos capítulos, en mayor o menor medida, contienen una serie de ejemplos para ilustrar los conceptos.

Descripción de la obra y su relación con la investigación educativa

EN ALGUNAS ÁREAS DE LA INVESTIGACIÓN EDUCATIVA, SE HACE uso de la *medición* del aprendizaje, inteligencia, deserción, motivación, autorregulación y ausentismo, entre muchos otros *fenómenos*, para luego ser analizados por medio de estadísticas. De estos análisis, pueden surgir resultados que pueden ser empleados para crear argumentos que puedan o no apoyar a teorías del aprendizaje, como la *Teoría Cognoscitiva Social*, *Teoría del Constructivismo* y *Teoría del Procesamiento de la Información*, entre otras. Con el presente volumen, se trata de poder observar hasta dónde las mediciones del aprendizaje y otros fenómenos definidos por teorías educativas poseen una distribución normal (campana de Gauss), para poder aplicarles los principios de probabilidad, estadística y correlación que son explicados en este libro. Además, algunas de las preguntas que se tratan de contestar son:

- ¿Qué es la probabilidad?, ¿cómo se representa geoméricamente y para qué se utiliza?
- ¿Qué es la distribución binomial?, ¿qué son las permutaciones y combinatorias?, ¿cómo se asocian las combinatorias y la probabilidad con la distribución binomial?
- ¿Qué es la distribución normal y de dónde surge?
- ¿Cuáles son las ventajas de la distribución estandarizada?
- ¿Qué es la correlación de Pearson?, ¿cómo se obtiene?, ¿cuáles son sus representaciones gráficas?
- ¿Cuál es la correlación de Spearman?, ¿cuándo se usa esta correlación en lugar de la de Pearson?

Al dar respuesta a estas preguntas, se espera que el panorama de las y los lectores se expanda, para poder aplicar los conceptos y análisis mostrados en la obra para llevar a cabo una investigación educativa cuantitativa.

Asimismo, el uso del presente libro también puede ser un inicio para investigadoras e investigadores educativos para ayudar a entender y analizar la enorme cantidad de datos que se generan en las redes sociales, como Facebook, Instagram y Twitter, entre otras fuentes. En otras palabras, esta obra y el libro *Conceptos básicos de estadísticas inferenciales aplicadas a la investigación educativa* (2019), de este servidor, podrían ser un posible principio para hacer análisis de lo que se conoce como *Big Data*, concepto que engloba bases de datos *my* grandes como las que generan las ya mencionadas redes sociales, que pueden ser analizados por medio de *softwares* que ayuden a encontrar patrones, tendencias y asociaciones, especialmente relacionados con algunos comportamientos humanos e interacciones. Este concepto de *Big Data* ligado a otro llamado minería de datos (*Data Mining*, que es la práctica de examinar grandes bases de datos para generar nueva información), pueden ser de gran ayuda en la investigación educativa para resolver preguntas como los siguientes dos ejemplos: ¿existe una relación entre el aprendizaje y ciertos factores (*e.g.*, horas invertidas en las redes sociales)?; y ¿existe una diferencia en el aprendizaje entre el grupo A y el grupo B?

Más precisamente, un ejemplo de *Big Data* y minería de datos en la presente obra sería examinar si el aprendizaje de estudiantes se relaciona con las horas que pasan en las redes sociales. Se podría esperar que entre más tiempo se esté en estas menor sea el aprendizaje. Sin embargo, se podría pensar que cierta cantidad de contacto a través de las redes sociales sea beneficioso para el aprendizaje. Esta es una mera especulación, pero más allá de lo que se encuentre en un solo estudio uno de los aspectos recomendables para hacer ciencia y obtener conocimiento es un *metaanálisis*, el cual reúne los resultados de múltiples estudios para ser analizados estadísticamente. Estos resultados podrían mostrar que existe un patrón como una correlación que se repite a través de ellos en varios estudios (*e.g.*, una correlación entre el aprendizaje y las horas invertidas en las redes sociales).

Dirigido a lectoras y lectores

ESTE LIBRO ESTÁ DIRIGIDO A LAS Y LOS ALUMNOS, INVESTIGADORES y profesores que deseen entender la relación entre la probabilidad y la estadística, las distribuciones binomial, normal y estandarizada, así como las correlaciones de Pearson y de Spearman. Aunque la obra podría ser utilizada de manera independiente por aquellas personas que estén familiarizadas con ciertos principios de probabilidad y estadística, sería aún mejor que se empleara como un complemento de *Conceptos básicos de estadísticas inferenciales aplicadas a la investigación educativa* (2019).

Conocimientos previos

SE ASUME QUE LAS Y LOS USUARIOS DEL PRESENTE LIBRO TIENEN ciertos conocimientos de estadística descriptiva, probabilidad a nivel elemental, álgebra y geometría analítica. Aunque se cubre en la obra la formulación de preguntas de investigación, el poder formularlas es un pilar para derivar hipótesis y ponerlas a prueba por medio de análisis estadísticos. Se recomienda consultar a Ercikan y Roth (2006) y a Ponce, Domínguez y Arriaga (2016), para más información acerca de la creación de preguntas de investigación en la investigación educativa.

La mayoría de los cálculos del presente libro fueron ejecutados en forma manual, pero en algunos de los casos se emplearon calculadoras en línea para la probabilidad calculada, así como tablas. El cálculo manual fue para mostrar lo que está detrás de *softwares* estadísticos, como SPSS, Minitab, SAS y R, entre otros.

Aporte principal de la obra

LA APORTACIÓN PRINCIPAL DE LA OBRA ES TRATAR DE AUMENTAR la base probabilística y estadística de *Conceptos básicos de estadísticas inferenciales aplicadas a la investigación educativa* (2019), para servir como un marco metodológico para la investigación educativa cuantitativa. También, la mayoría de las fuentes fueron publicadas

originalmente en idioma inglés y en esta obra, se tradujeron para hacerlas más accesibles a aquellas personas que prefieren el idioma español.

Más al respecto, los conceptos y análisis del presente libro (probabilidad, distribuciones binomial, normal y estandarizada, y correlación), junto con los de Ponce (2019) (temas de intervalos de confianza, estadísticas paramétricas y no-paramétricas, variables dependientes e independientes, hipótesis nula y alternativa, error tipo I y II, interpretación de resultados significativos, poder estadístico y d de Cohen), forman una parte de la base de la probabilidad y estadística para entender análisis más sofisticados, como las regresiones lineales, correlaciones canónicas y regresiones logísticas, así como el análisis de la varianza, covarianza y multivariante. Asimismo, esta parte de la base probabilística y estadística sirve para entender análisis multivariados, como ecuaciones estructurales y modelos lineales jerárquicos, entre muchos otros. En resumen, es fundamental tener el conocimiento de este material para hacer análisis básicos, así como para llegar a entender y realizar otros más sofisticados.

Razón para la creación y publicación de la obra

ESTA OBRA SURGIÓ COMO LA SEGUNDA PARTE DE *CONCEPTOS BÁSICOS de estadísticas inferenciales aplicadas a la investigación educativa* (2019). Ambas comenzaron como materiales didácticos de materias de métodos para la investigación educativa y las ciencias sociales. Además, dichos materiales didácticos también fueron empleados en seminarios de tesis de educación a nivel pre y posgrado. Durante casi una década, estos materiales didácticos se han ido perfeccionando por las contribuciones de estudiantes y colegas para llegar a las versiones actuales. Ahora, *Estadística elemental para investigación educativa: probabilidad, distribuciones y correlación*, se ha convertido en un libro de divulgación, ya que existen muchos programas educativos que lo podrían usar como referencia para sus tesis y clases.

Objetivos de aprendizaje de la obra

Al terminar el libro, la lectora o el lector serán capaces de:

- Comprender la probabilidad representada por áreas de figuras geométricas.
- Calcular permutaciones y combinatorias para su uso en la investigación educativa.
- Crear distribuciones binomiales a partir de combinatorias y probabilidades.
- Entender los principios básicos de la distribución normal.
- Comprender cómo la distribución de los datos afecta la distribución normal.
- Identificar cómo los coeficientes de la simetría y curtosis indican la posible normalidad de los datos.
- Calcular los valores estandarizados (zeta) z y las probabilidades que les corresponden en la distribución normal estandarizada.
- Comprender los usos de tres tipos de correlaciones: policóricas, r de Pearson y r de Spearman.
- Diferenciar las características de los coeficientes: r de Pearson y r de Spearman.
- Distinguir las diversas fórmulas r de Pearson.
- Formular hipótesis nulas y alternativas para poblaciones y muestras usando la r de Pearson.
- Entender los errores tipo I y II para las correlaciones.
- Interpretar los resultados significativos y no-significativos de correlación.
- Identificar la diferencia de usar Test de Una Cola con respecto a dos colas y el poder estadístico para la correlación.
- Entender y calcular el tamaño del efecto llamado *Coefficiente de Determinación* (r^2).
- Calcular intervalos de confianza para la r de Pearson.
- Relacionar el Test de Significancia Estadística y un intervalo de confianza para la r de Pearson.
- Identificar casos apropiados para la r de Spearman y estimar este coeficiente.

Términos clave

- Área
- Combinatorias
- Correlaciones: policóricas; r de Pearson; y r de Spearman
- Curtosis
- Distribución binomial
- Distribución de muestras de r de Pearson
- Distribución normal
- Distribución normal estandarizada
- Error estándar
- Error tipo I y II
- Hipótesis nula y alternativa para el r de Pearson
- Homocedasticidad vs. heterocedasticidad
- Intervalos de confianza para el r de Pearson
- Leptocúrtica
- Monotónico
- Multicolinealidad
- Permutaciones
- Platicúrtica
- Poder estadístico
- Probabilidad
- Operación factorial
- Simetría
- Tamaño del efecto: Coeficiente de Determinación (r^2)
- Test de Una Cola y Dos Colas
- Variables dependiente e independiente con el r de Pearson
- Valores ξ

CAPÍTULO 1

El área y la probabilidad

A muchas personas nos gustaría vivir eternamente. Para ser una vida eterna deseable para algunos, esta tendría ciertas condiciones en las cuales se pudieran satisfacer varias necesidades (*e.g.*, básicas y profesionales, entre otras). Dada la imposibilidad de la inmortalidad, algunas y algunos tratan de seguir el adagio de tener descendencia, plantar un árbol y escribir un libro para seguir vivas o vivos de alguna manera, porque la probabilidad de morir es del cien por ciento. Es decir, es seguro que esto suceda a menos que alguna tecnología terrestre —o extraterrestre— nos salve de ello. Asumiendo que no hay salvación por medio de la tecnología, ¿qué probabilidad tenemos de vivir hasta los ochenta o noventa años? Hay empresas privadas —seguros de vida— que analizan una serie de factores (*e.g.*, edad, estado de salud, hábitos, etcétera) para estimar la duración de vida de un individuo. Basándose en esta estimación de vida, las compañías deciden si una persona constituye un riesgo aceptable: tener una vida relativamente larga para que pague una póliza de seguro por el mayor tiempo posible.

Otro deseo que muchas y muchos tenemos es sacarnos algún premio en la lotería. Aunque la probabilidad de obtener el premio sea cercana a cero en muchos de los casos, hay que conseguir un boleto para tener

ciertas probabilidades de ganar. Sin embargo, a veces nos sentimos con suerte y decidimos comprarlo, a pesar de ser conscientes de las remotas posibilidades.

También, las y los investigadores educativos usan las probabilidades para estimar la aparición de algún tipo de evento: probabilidad de aprobar un examen; probabilidad de graduarse a tiempo; probabilidad de conseguir un empleo, etcétera. Una manera elegante de representar una probabilidad es el área de una figura geométrica: distribuciones binomiales, normales y estandarizadas, entre otras.

a. Probabilidad

Generalidades

“EN TÉRMINOS GENERALES, LA PROBABILIDAD DE UN EVENTO es un valor numérico que mide la posibilidad de que ocurra el evento” (Johnson, & Bhattacharyya, 1987, p. 85). Estos autores explicaron que las probabilidades expresan la factibilidad de eventos que no pueden ser predichos con certidumbre. El concepto de probabilidad se conoce como la proporción de ocasiones en las que se espera que un evento suceda (Johnson, & Bhattacharyya, 1987). Antes de adentrarse más en las probabilidades, sería necesario definir algunos conceptos, así como utilizar los símbolos adecuados:

- Espacio muestral (*sample space*) denotado por la letra griega omega (Ω) o la S de muestra en inglés (*sample*): es la recolección de todos los posibles resultados de un experimento cuando se puede entender este en un sentido amplio de la palabra (*e.g.*, lanzar una moneda al aire o jugar a los dados). Por ejemplo, el espacio muestral de lanzar una moneda, que sería el experimento, es de dos elementos, porque puede caer cara o cruz: $\Omega = 1 + 1 = 2$.
- Evento simple (*simple event*) o elemento del espacio muestral denotado por e_1, e_2, e_3, \dots : es cada uno de los resultados de un experimento. Volviendo al experimento

anterior, este tendría el evento simple de cara (e_{cara}) y el evento simple de cruz (e_{cruz}). Cuando se suman todos los elementos simples de un experimento, dan como resultado un espacio muestral.

- Evento denotado por letras mayúsculas (A, B, C , etcétera): es un set (*i.e.*, grupo) de resultados que poseen ciertas características. Por ejemplo, si en el experimento se selecciona la cara como el evento de interés ($A = \{e_{\text{cara}}\}$), se lee: el evento A es igual al conjunto del evento simple de cara; los símbolos de llaves $\{\dots\}$ significan un grupo de elementos de un conjunto.
- Probabilidad de un evento denotado por $P(A)$: es un valor numérico que representa la proporción de ocasiones que se espera encontrar un experimento en un evento dado, el cual es repetido en condiciones idénticas. Para colocar valores en el ejemplo anterior, sería: la probabilidad de cara cuando se lanza la moneda en una ocasión: $P(\text{caer cara}) = 1/2$.

A continuación, se presenta la *Teoría Clásica de la Probabilidad*, que está relacionada con el número de frecuencias. En conciso, las estadísticas están relacionadas con la probabilidad, que es una medida de que suceda algún evento de interés: en la Teoría de la Probabilidad, un evento puede ser uno o un set de resultados de algún experimento, a los cuales se les asignó una probabilidad. De esta, se puede obtener una fórmula que generalice otro tipo de eventos (Ecuación 1.1):

$$P(A) = \frac{F}{N}, \quad (\text{Ecuación 1.1})$$

donde:

- $P(A)$ = Probabilidad de un evento;
- F = Frecuencia de un evento;
- N = Número total de posibles resultados o espacio muestral.

Es decir, la probabilidad de que suceda algo [$P(A)$] en un experimento, es el resultado (cociente) de la frecuencia con la que ocurre un evento (F) dividido entre el número de posibilidades de que suceda o espacio muestral (N). Un experimento es:

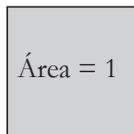
¿cuál es la probabilidad de que al lanzar un dado (seis caras) aparezca el número cuatro (evento), aplicando la Ecuación 1.1? La respuesta es uno de seis (un sexto):

$$P(\text{del número cuatro}) = \frac{1}{6}.$$

Área y propiedades de la probabilidad

EL SIGUIENTE CUADRO TIENE UN ÁREA DE 1 (FIGURA 1.1). SI SE VA A la fórmula del área = Lado₁ x Lado₂, de este modo se obtiene: 1 x 1 = 1. Si un investigador(a) está más familiarizado(a) con unidades cuadradas, se pueden usar los centímetros: 1 cm x 1 cm = 1 cm². Este número 1 representa el 100% (1 x 100) del área de esta figura.

Figura 1.1 Área de un cuadrado



Para el siguiente ejemplo, este cuadrado (Figura 1.1) representa una probabilidad de un evento. Esto es, el área expresaría la probabilidad de un evento que es 1. Es decir, la probabilidad de un evento es del 100%; *i.e.*, existe una probabilidad del 100% de que un lápiz caiga al piso cuando es soltado por una mano, sin considerar a alguna otra persona o variable que evite la caída. A la vez, esta área sería el espacio muestral.

Al lanzar una moneda común y corriente al aire, el número de posibilidades es de dos: cara (e_{cara} ; 50%) o cruz (e_{cruz} ; 50%). Algunos podrían pensar que este 50% y 50% es solamente en teoría, pero ha habido varios investigadores que han mostrado que cuando se lanza una moneda en un gran número de ocasiones, los porcentajes de cara y cruz se aproximan a dichos porcentajes teóricos. A esto también se le conoce como *converger* hacia el número teórico, de acuerdo con la *Ley de los Grandes Números* (también se le conoce como Teorema de Bernoulli; véase: Salkind, 2007). Un evento es algo que pasa, como la lluvia, y, por otro lado, está la probabilidad de un evento que es una predicción basada en la frecuencia de ocurrencia bajo ciertas circunstancias (véase:

Salkind, 2007, con teorías de la probabilidad para más información). Más al respecto, un evento como lanzar una moneda puede ser calculado como la probabilidad de que esta pueda caer con la cara hacia arriba; *i.e.*, probabilidad del evento = $1/2$:

$$P(\text{probabilidad del evento de cara}) = \frac{1}{2}.$$

El complemento de este evento de cara es que caiga cruz, porque se excluyen mutuamente, *i.e.*, o cae cara o cae cruz, pero no ambas. La palabra clave es la *o*, la cual indica que es una u otra de las posibilidades. Sumando ambas probabilidades darían 1, que es un 100% interpretado como la probabilidad de que caiga cara o cruz: *propiedad de la sumatoria de la probabilidad*, que dice que se pueden sumar las probabilidades de dos eventos que se excluyen mutuamente (Lane *et al.*, 2014; Figura 1.2).

Expresando lo anterior de una manera sintetizada:

$$\text{Probabilidad de cara o cruz} = 1/2 + 1/2 = 1.$$

Figura 1.2 Probabilidad de cara o cruz

$1/2$ Cara	$1/2$ Cruz
---------------	---------------

También, cabe la posibilidad de que se quiera computar la probabilidad en más de un lanzamiento de la moneda. Como cada tirada de la moneda es independiente de las demás (*i.e.*, no tiene influencia una tirada en otra), se utiliza la *propiedad multiplicativa de la probabilidad* (Lane *et al.*, 2014). Por ejemplo, la probabilidad de que caiga dos veces seguidas cara. También, esto se puede expresar como la probabilidad de que caiga cara en la primera ocasión *y* cara en la segunda, siendo la palabra *y* la clave para utilizar la multiplicación:

$$P(\text{cara en dos ocasiones consecutivas}) = 1/2 \times 1/2 = 1/4.$$

En concreto, la propiedad multiplicativa de la probabilidad se emplea cuando se están calculando las probabilidades en conjunto de dos o más sucesos independientes.

b. Permutaciones

SON LAS DIFERENTES MANERAS EN QUE EVENTOS, PERSONAS U OBJETOS se pueden organizar cuando el orden importa (Chandra, 2014). Un ejemplo intuitivo de lo anterior sería que hay tres estudiantes (María, Luisa y Sebastián) en un aula, que ocuparán tres puestos importantes, presidente, secretario y tesorero, en un comité para mejorar las condiciones de una escuela. Cada uno de los puestos es diferente y requiere diferentes habilidades y responsabilidades. Más allá de eso, se desea saber el número de maneras en las que se pueden distribuir los alumnos en estos tres puestos, *i.e.*, el número de permutaciones, lo cual se puede organizar de acuerdo con la Tabla 1.1:

Tabla 1.1 Permutaciones con tres personas y tres puestos

Candidata(o)	1.ª Per			2.ª Per			3.ª Per			4.ª Per			5.ª Per			6.ª Per		
	P	S	T	P	S	T	P	S	T	P	S	T	P	S	T	P	S	T
María	X			X				X				X		X				X
Luisa		X				X	X				X				X		X	
Sebastián			X		X				X	X		X	X				X	

Nota: Per = Permutación; P = Presidente; S = Secretario; y T = Tesorero.

En la Tabla 1.1, se muestran las seis posibles permutaciones de los tres candidatos para los tres puestos. La primera representa a María como presidenta, a Luisa como secretaria y a Sebastián como tesorero. En resumen, cada uno de ellos puede ser presidente en dos permutaciones mientras los otros dos alternan puestos (secretario y tesorero). Este sería un escenario en donde un cuadro (Tabla 1.1) sería suficiente para explicarlo, pero hay otras situaciones en donde la elaboración de cuadros sería una actividad bastante laboriosa. Para solucionar este problema está la siguiente fórmula (Ecuación 1.2):

$${}_n Per_k = \frac{n!}{(n-k)!}, \quad (\text{Ecuación 1.2})$$

donde:

- ${}_n Per_k$ = Número de permutaciones;
- n = Número de participantes;
- k = Número de puestos $\leq n$.

Otra manera de expresar la fórmula anterior es:

$$Per(n, k) = n! / (n - k)!$$

Son las posibles permutaciones de n , dado k (*i.e.*, el número de permutaciones de los participantes dado el número de puestos). El signo de exclamación significa una operación factorial. Por ejemplo, el cinco factorial (5!) da como resultado:

$$5 \times 4 \times 3 \times 2 \times 1 = 120.$$

Se tiene entonces una equidad:

$$5! = 120.$$

Para seguir una secuencia de ejemplos:

$$4! = 4 \times 3 \times 2 \times 1 = 24;$$

$$3! = 3 \times 2 \times 1 = 6;$$

$$2! = 2 \times 1 = 2;$$

$$\text{y } 1! = 1.$$

Un caso aparte es el cero factorial (0!), que es: $0! = 1$. Esto parece retar la intuición, pero en las permutaciones se utiliza de este modo (*cf.* Chandra, 2014). Usando las tres personas para los tres puestos del ejemplo anterior, resulta la siguiente operación con seis permutaciones posibles:

$$\begin{aligned}
 {}_3P_3 &= 3! / (3 - 3)! \\
 {}_3P_3 &= 3 \times 2 \times 1 / (0)! \\
 {}_3P_3 &= 6 / 1 = 6.
 \end{aligned}$$

Esto es, se tienen tres personas para tres puestos distintos y, por lo tanto, existen seis posibles permutaciones.

c. Combinatorias

UN ASPECTO DE LAS COMBINATORIAS ES QUE EL ORDEN NO ES IMPORTANTE. Es decir, en un set de eventos, participantes u objetos estos se organizan formando grupos sin que una secuencia importe. Tomando el ejemplo de los candidatos (María, Luisa y Sebastián), se quiere establecer un comité con tres miembros. Entonces, solo existiría una combinación para este comité: María, Luisa y Sebastián. No habría puestos jerárquicos, así que el orden *no* importaría en este caso. Por otro lado, un caso hipotético sería formar un comité con solo dos miembros de los tres candidatos. En pocas palabras, habría que ver cuántos comités se podrían crear con solamente dos miembros.

Tabla 1.2 Combinatorias para el comité

Combinatorias	María	Luisa	Sebastián
1	x	x	No participaría
2	x	No participaría	x
3	No participaría	x	x

En este caso solo habría tres combinatorias (Tabla 1.2). En la primera combinación, María y Luisa serían miembros sin importar ninguna jerarquía, y Sebastián no participaría. En la segunda, Luisa quedaría fuera; y en la tercera, María sería quien quedaría fuera. Entonces, se podrían establecer tres comités de dos miembros cada uno con dos candidatos. La fórmula de las combinatorias es la siguiente (Ecuación 1.3):

$${}_n C_k = \frac{n!}{k!(n-k)!}, \quad (\text{Ecuación 1.3})$$

donde:

- ${}_n C_k$ = Número de combinatorias;
- n = Número de participantes;
- k = Número de puestos.

Otra manera de expresar la fórmula anterior es:

$$C(n, k) = n! / k!(n - k)!$$

Son las posibles combinatorias de n , dado k (*i.e.*, las combinatorias posibles dado el número de participantes y de puestos). Si utilizamos el ejemplo anterior para ilustrar la fórmula, tendremos el resultado siguiente que se traduce en tres posibles combinatorias:

$$\begin{aligned} {}_3 C_2 &= 3! / 2!(3 - 2)! \\ {}_3 C_2 &= 3 \times 2 \times 1 / 2 \times 1 \times (1)! \\ {}_3 C_2 &= 6 / 2 \times 1! \\ {}_3 C_2 &= 6 / 2 = 3. \end{aligned}$$

Las combinatorias se podrían utilizar en varios aspectos de la investigación educativa. Por ejemplo, a la hora de tomar una muestra de participantes para un estudio. Supongamos que se tiene acceso a 100 estudiantes, pero solo hay recursos para efectuar una sola muestra de 30 integrantes. Entonces, la pregunta sería: ¿cuántos grupos/combinatorias de 30 se podrían formar de 100 alumnos? Como esta operación es difícil de efectuar manualmente, se recomienda consultar la siguiente página (<http://www.calculatorsoup.com/calculators/discretemathematics/combinations.php>) con una calculadora de combinatorias (CalculatorSoup®):

$${}_n C_k = 100! / 30!(100 - 30)!$$

$${}_n C_k = 29, 372, 339, 821, 610, 944, 823, 963, 760.$$

Este es un número bastante grande y es el que resultó para las combinatorias, lo que puede dar una idea del gran número de posibles combinatorias al tomar una muestra.

En los ejemplos de este primer capítulo, se usaron números discretos positivos, los cuales son unidades completas, *i.e.*, solo pueden tomar valores enteros. Un ejemplo es el número de estudiantes en un grupo: este va a ser un número discreto, *e.g.*, 31. Por otro lado, están los números continuos: 2.1, 2.2, etcétera, que pueden tomar valores con decimales, como la calificación de una materia: 8.3.

Preguntas para reflexionar

1. ¿Existen otras formas de calcular las probabilidades de algún evento?
2. ¿Cómo se podría representar una probabilidad sin emplear una figura geométrica bidimensional?
3. ¿En qué ocasiones en la investigación educativa sería adecuado utilizar las permutaciones en lugar de las combinatorias?
4. ¿Cuál es la fórmula (permutaciones vs. combinatorias) que produce números mayores?
5. ¿Qué otros usos se le pueden dar a estas fórmulas en la investigación educativa?
6. ¿Qué otros usos prácticos tienen estas fórmulas?

Opinión del autor

UN ÁREA ES UN MODELO PARA REPRESENTAR LA POSIBILIDAD DE un evento en la probabilidad y la estadística. La probabilidad y la estadística son parte de las matemáticas a las que se les conoce como aplicadas o prácticas. En este primer capítulo, se presentaron figuras bidimensionales para modelar probabilidades. Los modelos pueden llegar a ser más complejos donde se empleen tres dimensiones: una para la probabilidad y las dos restantes para otras dos variables, *e.g.*, x y y . Sin embargo, podría ser que los modelos tengan más de dos variables y su representación gráfica sería un tanto complicada. Cuando se exceden las tres dimensiones y se desea ahondar en este tema de la representación geométrica de las variables, recomiendo el tópico de los hiperplanos (*hyperplanes*).

Para algunas y algunos estudiantes es complicado entender a la primera explicación las permutaciones y las combinatorias. Lo que recomiendo es tratar de ligar los ejemplos con estos dos anteriores conceptos con ciertos conocimientos que posean los alumnos. Esto sería para crear *aprendizaje significativo*. No me preocuparía mucho si los estudiantes no logran entender estos conceptos en algunas sesiones, porque este proceso suele llevar tiempo.

CAPÍTULO 2

Distribución binomial y su relación con la distribución normal y estandarizada

Una distribución binomial puede representar una serie de eventos, como lanzar monedas o dados, o contestar un examen de matemáticas. La constante es que existen dos alternativas o eventos posibles que son clasificados como *éxito* o *fracaso*, que se excluyen mutuamente, *i.e.*, o es uno de ellos o es el otro. Por ejemplo, si están dos individuos jugando a lanzar una moneda y adivinar el resultado, la persona uno elige que caerá cara en la primera tirada, por lo que cara sería clasificada como éxito en este caso. En la siguiente ocasión, la persona dos elige que caerá cruz, así que cruz sería el éxito, entonces. Es decir, el éxito o el fracaso de un evento puede alternar. Donde no podría alternar esta definición de éxito o de fracaso sería en un examen de matemáticas u otra materia, asumiendo que solo exista una respuesta correcta. Un ejemplo sería un examen de matemáticas de opción múltiple de cuatro alternativas. Una alumna(o) que no estudió, se enfrenta a contestar cada pregunta o ítem. Simplemente podría responder al azar y, con ello, tendría un $1/4$ de probabilidad de contestar correctamente un ítem. Si decide usar solo una letra (*e.g.*, letra *C*) porque no sabe y el profesor(a) distribuyó uniformemente las respuestas correctas, tiene la misma probabilidad ($1/4$) de

lograr correctamente una respuesta. La distribución binomial es la base de la distribución normal, que representa muchos fenómenos como diferentes niveles de aprendizaje. Sin embargo, es más práctico y simple utilizar la distribución estandarizada que la normal a la hora de hacer investigación educativa.

a. Ejemplo para construir la distribución binomial

SE DEFINE A X COMO UNA VARIABLE QUE OCURRE AL AZAR. Por ejemplo, si se lanza una moneda al aire en una ocasión existen dos alternativas (cara o cruz) que se excluyen mutuamente: cae una u otra, pero no pueden aparecer ambas a la vez. Estas dos alternativas son dos eventos y son considerados como éxito o fracaso. La probabilidad de que caiga cara es de $1/2$ y de que sea cruz es de $1/2$ también. En otras palabras, la probabilidad de que aparezca cara es de $1/2$, expresada en decimal por .50 y en porcentaje por 50%. ¿Cuál sería la probabilidad de éxito? Eso dependería de qué es el éxito. En este caso, puede ser cara o cruz. Al definir una alternativa como éxito, la otra se convertiría en fracaso y viceversa. Un ejemplo puede ser: ¿cuál es la probabilidad de que el evento sea cara (éxito)? La respuesta es:

Probabilidad de que caiga cara es igual a $1/2$. Esto se puede expresar en lenguaje matemático como:

$$P(\text{cara}) = p;$$

$$P = \text{Probabilidad de un evento};$$

$$p = \text{Probabilidad de que caiga cara} = 1/2.$$

Un ejemplo podría ser que la moneda se lanzara al aire en múltiples ocasiones. Cada una es independiente de la anterior (dada esta independencia, se puede aplicar la antes mencionada *propiedad multiplicativa de la probabilidad* para calcular la probabilidad en más de una ocasión). En concreto, la fórmula para calcular la probabilidad de un

set de ocasiones independientes es la siguiente con el ejemplo de lanzar una moneda en dos ocasiones:

$$P(x) = p \times p, \quad (\text{Ecuación 2.1})$$

donde:

- $P(x)$ = Probabilidad del set de eventos (en este caso, son dos caras);
- p = Probabilidad de un solo evento, *i.e.*, caiga cara = $1/2$.

En el caso de que se quiera usar la cruz: q = Probabilidad de que caiga cruz = $1/2$. El otorgarle la p o la q al éxito y el fracaso, respectivamente, es por convencionalismo. El set de eventos puede contener un número grande de multiplicaciones: $P(x) = p \times p \times p \times \dots \times p$.

Otro ejemplo es cuando una moneda es lanzada en cinco ocasiones. Podría pasar que, en esas cinco tiradas, *no* cayera *cruz* (éxito) en ninguna ocasión o que apareciera en una, en dos, en tres, en cuatro o en cinco ocasiones. Para empezar desde lo más simple, uno se podría preguntar: ¿cuál es la probabilidad de que *no* caiga cruz? Es decir, la moneda se lanza en cinco ocasiones consecutivas y no aparece cruz en ninguna. Esta misma pregunta se puede hacer de otra manera: ¿cuál es la probabilidad de que caiga *cara* (fracaso) en cinco ocasiones consecutivas? Para resolver estas dos preguntas que significan lo mismo, se necesita de la probabilidad con su propiedad multiplicativa junto con las combinatorias correspondientes:

$$P(A) = \frac{F}{N}, \quad (\text{Ecuación 1.1})$$

$${}_n C_k = \frac{n!}{k!(n-k)!}, \quad (\text{Ecuación 1.3})$$

Paso uno: sería definir el éxito que, en este caso, es que caiga *cruz* (se representa con la letra p). En cambio, cara (q) es el fracaso. Recapitulando, caben las probabilidades de que *no* aparezca cara en ninguna de las cinco tiradas o de que caiga en una ocasión, en dos, en tres, en cuatro y en cinco. Dada esta información, se puede desig-

nar como evento el lanzar una moneda en cinco ocasiones. Entonces, se calcula esta probabilidad con la propiedad multiplicativa (Ecuación 2.1).

Tabla 2.1 Probabilidades

De cinco tiradas	Multiplicación	En números	Resultado
No cae cruz	$q \text{ X } q \text{ X } q \text{ X } q \text{ X } q$	$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$.031
Cae una cruz	$p \text{ X } q \text{ X } q \text{ X } q \text{ X } q$	$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$.031
Caen dos...	$p \text{ X } p \text{ X } q \text{ X } q \text{ X } q$	$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$.031
Caen tres...	$p \text{ X } p \text{ X } p \text{ X } q \text{ X } q$	$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$.031
Caen cuatro...	$p \text{ X } p \text{ X } p \text{ X } p \text{ X } q$	$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$.031
Caen cinco...	$p \text{ X } p \text{ X } p \text{ X } p \text{ X } p$	$1/2 \times 1/2 \times 1/2 \times 1/2 \times 1/2$.031

Nota: p = Probabilidad de que caiga cruz; q = Probabilidad de que caiga cara.

Posiblemente suena muy simplista que el resultado de la probabilidad en este momento sea el mismo para todos los eventos, pero el ejemplo no está completo.

Paso dos: hay que considerar las combinatorias, porque esto aumenta la probabilidad de cada evento. Es decir, un resultado que tenga más combinatorias tendrá una probabilidad más alta de aparecer. Habría que estimar estas combinatorias para completar la estimación de la probabilidad de que *no* caiga cruz en ninguna ocasión o de que aparezca en una, en dos, en tres, en cuatro o en cinco ocasiones.

Antes de usar la fórmula de las combinatorias, se muestran las posibilidades de que caiga una *sola* cruz (Tabla 2.2). Esta ilustración es para dar a conocer cómo con una tabla se pueden hacer las combinatorias. Es decir, puede que aparezca una sola cruz en alguno de los cinco lanzamientos (puede caer en la primera, segunda, tercera, cuarta o quinta tiradas). Existen cinco combinatorias diferentes de que aparezca una sola cruz.

Tabla 2.2 Combinatorias con una sola cara

Cara	Tiradas				
	1. ^a	2. ^a	3. ^a	4. ^a	5. ^a
1	Sí	No	No	No	No
1	No	Sí	No	No	No
1	No	No	Sí	No	No
1	No	No	No	Sí	No
1	No	No	No	No	Sí

En la Tabla 2.3, se muestran los cálculos de las combinatorias y las probabilidades. Dadas las combinatorias, la probabilidad de que no caiga cruz o de que aparezca cara en cinco eventos consecutivos es igual a .031. De la misma manera, lo es para que caiga cara una vez o en cuatro ocasiones (.156), y para que aparezca dos veces o en tres ocasiones (.313).

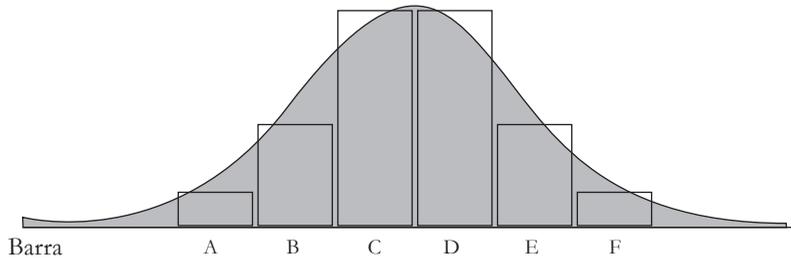
Tabla 2.3 Probabilidades y combinatorias

(n) Tira- das	Número de la cruz (k)	Fórmula de combinación	Combinato- rias	Combinato- rias posibles	Probabili- dad (1/32)	Resultado (Combinatorias x probabilidad)
5	0	$5! / [0!(5 - 0)!]$	1	$P(q, q, q, q, q)$	0.031	.031*
5	1	$5! / [1!(5 - 1)!]$	5	$P(q, q, q, q, p) \dots$ $P(p, q, q, q, q)$	0.031	.156
5	2	$5! / [2!(5 - 2)!]$	10	$P(q, q, q, p, p) \dots$ $P(p, p, q, q, q)$	0.031	.313
5	3	$5! / [3!(5 - 3)!]$	10	$P(q, q, p, p, p) \dots$ $P(p, p, p, q, q)$	0.031	.313
5	4	$5! / [4!(5 - 4)!]$	5	$P(q, p, p, p, p) \dots$ $P(p, p, p, p, q)$	0.031	.156
5	5	$5! / [5!(5 - 5)!]$	1	$P(p, p, p, p, p)$	0.031	.031*
Suma de las probabilidades dadas las combinatorias						1

Nota: *Esta probabilidad calculada es menor a .05, así que sería estadísticamente significativa (véase: Ponce, 2019).

Este patrón es simétrico y se puede expresar en una gráfica de barras delineadas por una curva similar a la curva normal (Figura 2.1). Esta es la distribución binomial donde es más probable obtener un valor en el área del centro de la curva que en los extremos, por ser esta más grande.

Figura 2.1 Distribución binomial de la moneda



Nota: dadas las combinatorias en cinco tiradas: A (probabilidad de 0 cruces) o F (probabilidad de 5 caras) = .031; B (probabilidad de 1 cruz) o E (probabilidad de 4 caras) = .156; y C (probabilidad de 2 cruces) o D (probabilidad de 3 caras) = .313.

La suma de todas las probabilidades, dadas las combinatorias, es de 1, que representa el 100% del área en la distribución binomial. Esta unidad obtenida responde a la pregunta: ¿cuál es la suma de las probabilidades, dadas las combinatorias, de que caiga cruz una vez ... caiga en cinco ocasiones? La respuesta es 1. Esta propiedad de la suma, se podría utilizar para responder otras interrogantes como: ¿cuál es la probabilidad de que en cinco tiradas, y dadas las combinatorias, aparezca cruz en cuatro o más ocasiones? La respuesta es $.156 + .031 = .187$. Para más usos de las combinatorias y probabilidades en ejemplos educativos, véase a Schumacker y Tomek (2013). En síntesis, las áreas de C y D son iguales y son las más grandes de esta distribución binomial, lo cual las hace más probables que las demás áreas.

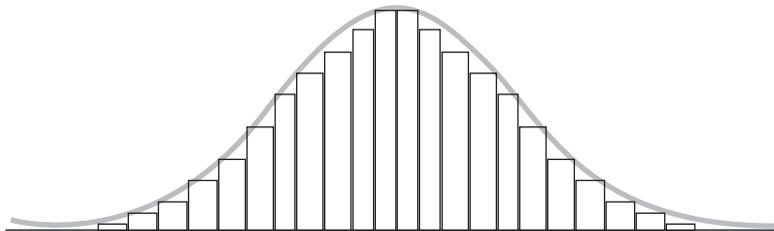
b. Estadísticas de la distribución binomial

SE CALCULAN DE UNA FORMA DIFERENTE A OTRAS ESTADÍSTICAS (*i.e.*, promedio, varianza, desviación estándar, etcétera). Un ejemplo podría ser lanzar una moneda en cien ocasiones, lo cual produciría una distribución binomial semejante a una distribución normal (véase la Figura 2.2 como una ilustración). Más al respecto, el promedio, la varianza y la desviación estándar (*SD*) de la distribución binomial son:

- Promedio de una distribución binomial (*i.e.*, \bar{x}) = np o también puede ser nq , porque dará el mismo resultado cuando $p = q$;
 - p = Probabilidad de obtener *crúz* en un lanzamiento ($1/2$);
 - q = Probabilidad de obtener *cara* en un lanzamiento ($1/2$);
 - n = Número de tiradas.

- Promedio para el ejemplo de 100 tiradas en la distribución binomial:
 - $\bar{x}_{\text{distribución binomial}} = 100 \times 1/2 = 50$.
- Varianza $_{\text{distribución binomial}} = npq$;
 - Varianza $_{\text{distribución binomial}} = 100 \times 1/2 \times 1/2 = 25$.
- $SD_{\text{distribución binomial}} = \sqrt{npq}$;
 - $SD_{\text{distribución binomial}} = \sqrt{25} = 5$.

Figura 2.2 Distribución binomial de la moneda con 100 tiradas



Nota: la gráfica es solo un modelo de aproximación, porque si fuera una copia fiel necesitaría 100 barras.

c. Ejemplo de combinatorias y probabilidades con dados

OTRO EJEMPLO DE PROBABILIDADES, DADAS CIERTAS COMBINATORIAS, es mostrado a continuación con *dados*. La probabilidad de que salga cualquier cara de un dado (seis caras) es de $1/6$. La probabilidad de éxito (adivinar la unidad que aparece) es de $1/6$ y la probabilidad de fracaso es de $5/6$. La suma de la probabilidad de éxito y de fracaso es: $1/6 + 5/6 = 6/6 = 1$ (este 1 es la probabilidad de que caiga cualquiera de las caras del dado: 100%).

Al arrojar dos dados a la vez, la probabilidad de que salgan dos números pre-determinados es: P (de dos eventos) $= 1/6 \times 1/6 = 1/36$ (*propiedad multiplicativa de la probabilidad*). La probabilidad de éxito es de $1/36 = .028$. Más específicamente, ¿cuál sería la probabilidad de que al lanzar dos dados la suma de estos sea de dos unidades?: $1/6 \times 1/6 = 1/36$. Ahora, ¿de cuántas maneras se podrían sumar los resultados para que nos dieran dos unidades? Este evento se resuelve de la siguiente manera:

$${}_n C_k = \frac{n!}{k!(n-k)!}, \quad (\text{Ecuación 1.3})$$

donde:

- ${}_n C_k$ = Número de combinatorias;

- n = Número de dados: 2 (cada uno se arroja una sola vez);
- k = Número de puestos: 2;

$${}_2C_2 = 2! / 2! (2 - 2)! = 1.$$

Solo hay una combinación para obtener una suma de dos unidades en el lanzamiento de dos dados. Esto es, un dado muestra una unidad y el otro también, para que ambos sumen dos unidades, por lo que no existe otra combinación de unidades que pudiera dar esta suma. Dado que solo hay una combinación, la probabilidad de que caiga una unidad en ambos dados es de $1/36$ (i.e., $1/6 \times 1/6$) multiplicada por la combinación 1, la cual da $1/36$. En resumen, lo anterior puede ser expresado como *probabilidad de éxito del evento uno multiplicada por probabilidad de éxito del evento dos multiplicada por el número de combinatorias*. En lenguaje matemático, se tiene un evento de esta manera:

$$P(\text{dos dados sumen dos unidades}) = 1/6 \times 1/6 \times 1 = 1/36.$$

Ejemplo hipotético con dados

UN EJEMPLO PARA MOSTRAR LO ANTERIOR: DOS PERSONAS (JOSEFINA y Mario) están jugando a lanzar dos dados y adivinar la suma de las unidades. Para simplificar y usar el ejemplo anterior, Mario dice que sus dados sumarán dos unidades. Por otro lado, Josefina dice que los de ella sumarán seis unidades. ¿Quién ganaría? Probablemente Josefina lo haría. La razón es que Josefina tendría más probabilidad de ganar, porque la suma de seis unidades tiene más combinatorias que las de Mario. En la Tabla 2.4, se muestran las combinatorias posibles para Josefina y Mario.

Tabla 2.4 Dos lanzamientos de dados en combinatorias

Mario								
Dado/puntos	1	2	3	4	5	6	Suma	Combinatorias
1.º	x						2	1
2.º	x							
Total								1
Josefina								
Dado/puntos	1	2	3	4	5	6	Suma	Combinatorias
1.º			x				6	1
2.º			x					
1.º		x					6	1
2.º				x				
1.º					x		6	1
2.º	x							
Total								3

Hay tres combinatorias por las cuales Josefina podría obtener una suma de seis unidades: (tres en un dado y tres en el otro); (cuatro y dos); y (cinco y una). Como al lanzar los dados una tirada es independiente de la otra: $1/6 \times 1/6 = 1/36$. Este $1/36$ también es la probabilidad de Mario. Considerando las combinatorias, la probabilidad de que los dados sumen seis unidades para Josefina es: P (dos dados sumen seis unidades) $= 1/6 \times 1/6 \times 3 = 3/36$. Por lo tanto, Josefina probablemente le ganaría a Mario, porque su probabilidad ($3/36$) es mayor que la de él ($1/36$). En conclusión, cuando existen más combinatorias en las que se puede obtener un resultado como la suma de puntos de un dado, mayor será la probabilidad que se tendrá:

$$P(\text{dos dados sumen dos unidades}): 1/6 \times 1/6 \times 1 < P(\text{dos dados sumen seis unidades}): 1/6 \times 1/6 \times 3.$$

d. Características de la distribución binomial

En un experimento

EN RESUMEN, CUANDO SE UTILIZA LA DISTRIBUCIÓN BINOMIAL EN un experimento debe tener las siguientes características:

- Debe tener un número determinado de tiradas, lanzamientos o ítems, que se denotan por la letra n .
- Cada tirada, lanzamiento o ítem debe tener solamente dos posibles resultados: éxito (p) o fracaso (q).
- El resultado de cada tirada, lanzamiento o ítem debe ser independiente de los demás: no hay efecto entre unos y otros.
- La probabilidad de éxito debe ser la misma en cada tirada, lanzamiento o ítem.

En la distribución de la curva

UNA VEZ QUE SE OBTIENEN LAS PROBABILIDADES DE CADA EVENTO, se pueden representar por una serie de barras como las de las figuras 2.1 y 2.2. Asimismo, se puede dibujar una curva semejante a la curva normal que contenga dichas barras. Esta curva de la distribución binomial es simétrica en este caso en particular donde los valores cercanos al promedio son los más probables y los cercanos a las colas son los menos probables. Sin embargo, no siempre es simétrica, porque depende del número de tiradas, lanzamientos o ítems para obtener la simetría. Otra característica es que *no* es asintótica como la curva normal. Esto quiere decir que sus colas *no* se prolongan hacia el infinito como las de la curva normal, que se muestra más adelante. También, se usan los ya antes mencionados números discretos para el éxito o el fracaso.

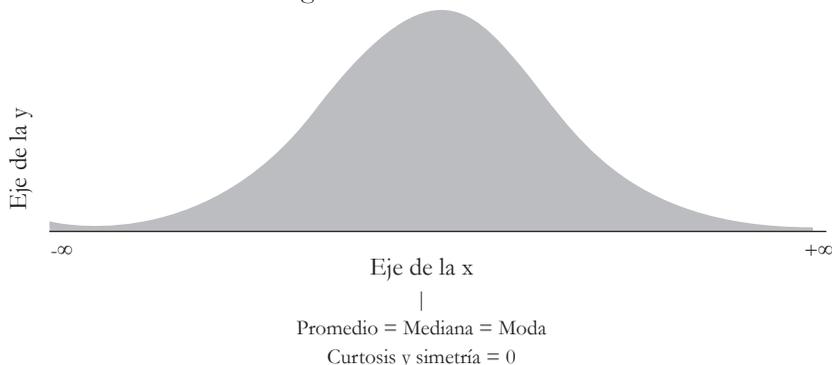
e. Curva de la distribución normal

LA CURVA NORMAL ES DERIVADA DE LA DISTRIBUCIÓN TEORÉTICA binomial (Lane *et al.*, 2014). La curva normal expresa la distribución al azar de un fenómeno de una población (Hurley, 2014, p. 597). “La normalidad en estadísticas es una condición en la cual un set de datos presenta una distribución normal de valores” (APA, 2015, p. 716). Se considera que puede representar varios aspectos de la naturaleza, como la estatura o el peso de las personas o su inteligencia, entre otros (Ponce, 2019). También conocida como distribución gaussiana o campana de Gauss, la curva normal o distribución normal fue definida de la siguiente manera:

Una distribución teórica en la cual la mayoría de los valores se apilan en el centro donde está el promedio y los demás de ellos se distribuyen en los dos extremos. Cuando los datos se grafican, dan la conocida curva en forma de campana cuando la variación alrededor del promedio es al azar. La distribución normal tiene varias características: es simétrica, tiene una asíntota baja y un alta, y su promedio, mediana y moda tienen el mismo valor (APA, 2015, pp. 715-716).

Una forma tradicional de representar la curva normal es usando dos ejes (Figura 2.3): a) el eje de la x , que representa una variable como un nivel de aprendizaje; y b) el eje de la y , que representa una probabilidad del aprendizaje. Más al respecto se cubrirá cuando se aborde la *curva de la distribución normal estandarizada*.

Figura 2.3 Curva normal



Ponce (2019) explicó que: “La curva normal tiene la propiedad de que los valores de y son en función (*i.e.*, dependen de) de los de x ”. Esto quiere decir que si se toma un valor de x , se puede calcular el valor correspondiente de y utilizando la *Fórmula de la Distribución Normal*, que se muestra a continuación (Ecuación 2.2):

$$\text{Fórmula de la Distribución Normal: } y_i = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2}, \quad (\text{Ecuación 2.2})$$

donde:

- y_i = Altura de la curva para cada uno de los valores de x_i (en un plano cartesiano);
- σ = Desviación estándar de una población. También, es interpretada como la desviación estándar de los datos: muestra;
- π = Valor de 3.1416... que es una constante y representa la razón de la circunferencia al diámetro de un círculo; *e.g.*, para la longitud de una circunferencia hay 3.1416... diámetros: $1/3.1416$;
- e = Valor constante de 2.7183... que es la base de los logaritmos naturales;
- μ = Promedio de la población. También, es interpretado como el promedio de los datos: muestra, en este tipo de casos;
- x_i = Valor de la variable independiente, *i.e.*, es la longitud.

Curtosis y asimetría

OTRA MANERA DE EVALUAR LA NORMALIDAD DE LA DISTRIBUCIÓN de los datos es emplear los coeficientes que dan las estadísticas de la curtosis y la asimetría. La evaluación de la distribución de los datos sirve para ver si se cumplen algunos de los supuestos de análisis que se utilizan para relacionar variables o comparar promedios de diferentes grupos (véase: Tabachnick y Fidell, 2012). Cuando los coeficientes de curtosis y asimetría son iguales a cero, no hay problemas de curtosis o de asimetría y, hasta cierto punto, se podría asumir que los datos siguen una distribución normal. Algunas y algunos investigadores han dicho que los problemas de estos dos coeficientes empiezan cuando se superan los umbrales de los valores absolutos de 2 y 3 para muestras pequeñas y grandes, respectivamente.

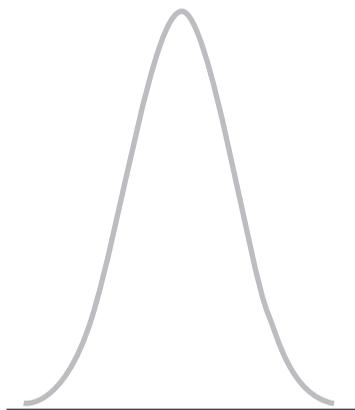
Curtosis

“LA CURTOSIS ES UNA MEDIACIÓN DE LA CANTIDAD DE DATOS QUE se concentran en las colas de una distribución con relación al resto de los datos” (Ponce, 2019, p. 19). Cuando la curtosis es igual a cero (normal o mesocúrtica; Figura 2.4[a]), la distribución de los datos es normal, sin considerar otras cosas (*ceteris paribus*). Puede pasar que los datos, se concentren demasiado alrededor del promedio y, por lo tanto, el Coeficiente de Curtosis será mayor que cero (distribución leptocúrtica (Figura 2.4[b])). Pero cuando este coeficiente es menor que cero, se tiene una distribución **platicúrtica**; Figura 2.4[c]).

Figura 2.4 Curtosis



(a) Curtosis = 0
Normal



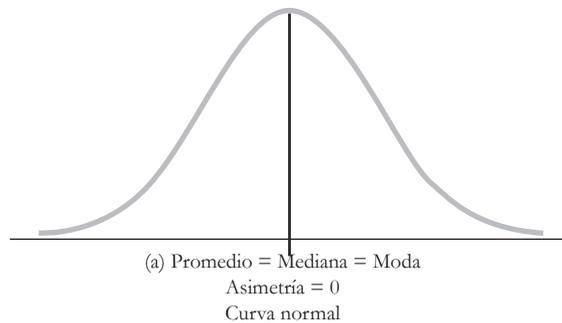
(b) Curtosis > 0
Leptocúrtica

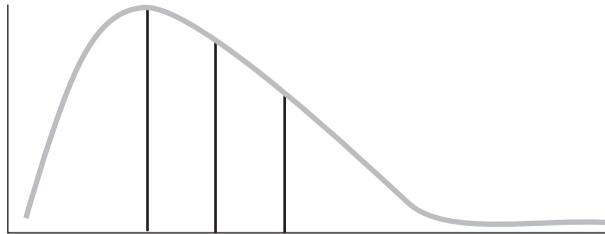


(c) Curtosis < 0
Platicúrtica

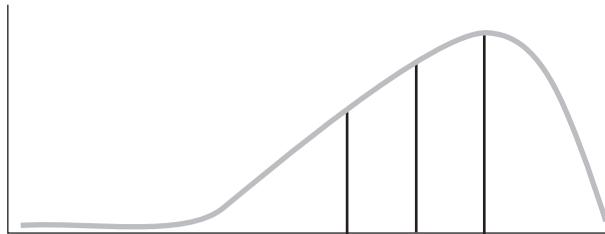
Asimetría (skewness)

TAMBIÉN ES CONOCIDA COMO SESGO. SEGÚN SALKIND (2011, P. 435), es una característica de una distribución que muestra una frecuencia desproporcionada de ciertos datos. Esto es, mide el grado de asimetría de un set de datos. En la Figura 2.5, se muestra una serie de ejemplos de asimetría. El ejemplo (a) representa el caso en el que la asimetría es igual a cero, lo cual indica que la curva es perfectamente simétrica. Por otro lado, la distribución del ejemplo (b) indica que la curva está sesgada positivamente, porque el promedio es mayor que la mediana y esta es mayor que la moda. Por lo tanto, el Coeficiente de Asimetría es mayor que cero, lo cual ocasiona que la cola del lado derecho sea más alargada que la otra. El caso de la distribución del ejemplo (c) indica que la curva está sesgada negativamente, porque el promedio es menor que la mediana y esta es menor que la moda. El resultado es que el Coeficiente de Asimetría es menor que cero, lo que ocasiona que la cola del lado izquierdo sea más alargada que la otra.

Figura 2.5 Asimetría



(b) Moda < Mediana < Promedio
Asimetría > 0
Positivamente sesgada



(c) Promedio < Mediana < Moda
Asimetría < 0
Negativamente sesgada

Asimetría y observaciones atípicas

UN SESGO O DISTRIBUCIÓN ASIMÉTRICA PUDO HABER SIDO OCASIONADO por una sola observación atípica: es un valor en un set de datos que es diferente a los demás en el nivel en que se mide una variable (véase: Hair *et al.*, 2014). Esto es, esta observación atípica tiene un valor extremo: más allá de un valor α (explicado en la siguiente sección) de $|2|$ o $|3|$. A continuación, se muestra un ejemplo para exponer el efecto de una observación atípica en la asimetría de una distribución. Se tiene una situación original en la cual están cinco estudiantes (Tabla 2.5), A-E, con sus respectivas calificaciones (de 3 a 7) en un examen de ciencias naturales. Se calcularon los valores α con la Ecuación 2.3, que se explica en la siguiente sección: la *distribución normal estandarizada*. En este caso, estos valores α indican que la distribución es simétrica: -1.27 y

1.27; -0.63 y 0.63; y 0. Es decir, se despegan del 0 en forma equidistante. Asimismo, el promedio es igual a la mediana y el Coeficiente de Asimetría es igual a cero. La moda no se incluyó en este ejemplo, porque los valores *no se repiten*.

Tabla 2.5 Valores atípicos y asimetría

Alumnos	Situación original	Valores ε	Situación 1	1 Valores ε	Situación 2	2 Valores ε
I			0	-3.16		
A	3	-1.27	3	-1.27	3	-1.27
B	4	-0.63	4	-0.63	4	-0.63
C	5	0	5	0	5	0
D	6	.63	6	.63	6	.63
E	7	1.27	7	1.27	7	1.27
II					10	3.16
Promedio	5		4.17		5.83	
SD	1.58		2.48		2.48	
Mediana	5		4.5		5.5	
Asimetría	0		-0.87		.87	

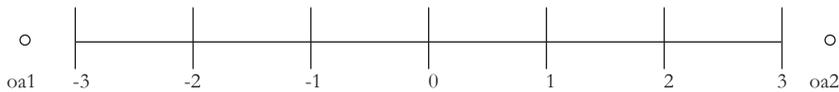
Nota: no hay moda, porque los valores no se repiten.

Por otro lado, la situación 1 muestra que se añadió una calificación más a las cinco originales (Tabla 2.5), *i.e.*, 0 (alumno I). Esto hizo que, al usar el promedio y la *SD* de la situación original, se pudiera detectar esta nueva calificación como una observación atípica: $\varepsilon = -3.16$. Dada esta nueva calificación de cero, la distribución de los datos ya no es simétrica, debido a que el promedio (4.17) es menor que la mediana (4.5): está sesgada negativamente a la izquierda (Figura 2.5c). Asimismo, su Coeficiente de Asimetría = -0.87 y aunque no excede los límites de $|2|$ o $|3|$ que se consideran preocupantes en la investigación educativa, sí cambió el coeficiente original de cero.

La situación 2 implica la inclusión de una calificación más a las cinco originales, *i.e.*, 10 (alumno II). De una manera similar a la situación anterior, esta sexta calificación hace que su valor ε sea atípico (3.16), relativo al escenario original. Por lo tanto, la distribución de los datos se vuelve asimétrica, porque el promedio (5.83) es mayor que la mediana (5.5): está sesgada positivamente a la derecha (Figura 2.5b). De igual manera, su Coeficiente de Asimetría = .87, pero sin exceder los límites.

Estas observaciones atípicas (Figura 2.6), se colocan en una recta donde se tienen los límites que se suelen emplear en la investigación educativa ($|3|$), las cuales se salen de los límites antes descritos. Como resultado, estas observaciones sesgan los datos.

Figura 2.6. Observaciones atípicas



Nota: oa = Observación atípica.

Normalidad y el Test de Kolmogorov-Smirnov

OTRA MANERA DE PONER A PRUEBA LA NORMALIDAD DE LOS DATOS con los que se va a trabajar, es usando el Test de Kolmogorov-Smirnov para una muestra, el cual se aplica a nivel univariable (una sola variable). Con este test, se pone a prueba la hipótesis nula de que dos distribuciones son idénticas (Salkind, 2007, p. 512). En corto, se pone a prueba la hipótesis nula de que los datos de una muestra siguen una distribución normal:

$$H_0: \text{distribución normal} = \text{distribución observada.}$$

Si la probabilidad calculada es igual o mayor al alfa seleccionado como criterio, no se rechaza la hipótesis nula y se asume que la distribución observada tiene una distribución normal (véase: Ponce [2019], para más detalles al respecto de la significancia estadística y la prueba de la hipótesis nula). En contraparte, si se rechaza la hipótesis nula, se puede concluir que la distribución de los datos es estadística y significativamente diferente de la distribución normal. Entonces, esta evidencia apoyaría la hipótesis alternativa:

H_A : distribución normal \neq distribución observada.

El Test de Kolmogorov-Smirnov ha sido incorporado en el *software* SPSS, entre otros. Sin embargo, está más allá de los objetivos del presente libro explicarlo a detalle, así como su ejecución en dicho *software*, pero Salkind (2007) hace una discusión bastante profunda sobre el tema.

f. Distribución normal estandarizada

PARA ANALIZAR LOS DATOS DE UNA INVESTIGACIÓN, HINKLE, Wiersma y Jurs (2003) explicaron que, para poder emplear el modelo de la distribución normal, es necesario utilizar la *distribución normal estandarizada*. El uso de la distribución estandarizada se debe a que habría una infinidad de distribuciones normales producidas por cada promedio y desviación estándar (Ponce, 2019). Además, al emplear la *distribución normal estandarizada* se pueden utilizar las tablas de valores z que aparecen en muchos libros de estadística, para calcular las probabilidades de los valores. Un valor z se obtiene mediante la siguiente fórmula (Ecuación 2.3):

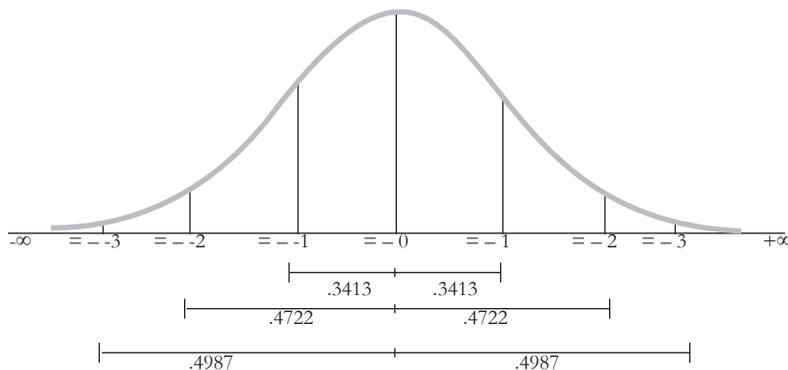
$$z_i = \frac{(x_i - \bar{x})}{SD}, \quad (\text{Ecuación 2.3})$$

donde:

- z_i = Coeficiente que se obtiene de la operación. Puede ser localizado en la *distribución normal estandarizada*;
- x_i = Valor conocido como crudo o bruto en la recolección de datos. Se encuentra en la unidad en la que se recolectaron los datos; *e.g.*, calificaciones, puntajes, edad, etcétera;
 - \bar{x} = Promedio de los datos crudos o brutos en su unidad;
 - SD = Desviación estándar de los datos crudos o brutos.

En otras palabras, si se tiene un valor individual en bruto, el promedio de los datos en bruto y su SD , se puede calcular su correspondiente valor estandarizado (z). Con este valor z , se puede saber su ubicación en la *distribución normal estandarizada* (Figura 2.7). Los valores ($|1|$, $|2|$ y $|3|$), que se muestran en la Figura 2.7, también coinciden con la distancia en desviaciones estándar con el promedio. Ejemplos: un valor $z = |1|$ es igual a una $SD = |1|$ y un área del promedio (0) a un valor $z = |1|$ es de .3413, que en porcentaje es el 34.13%; un valor $z = |2|$ es igual a dos $SD = |2|$ y un área del promedio (0) a un valor $z = |2|$ es de .4772, que en porcentaje es el 47.72%; y un valor $z = |3|$ es igual a tres $SD = |3|$ y un área del promedio (0) a un valor $z = |3|$ es de .4987, que en porcentaje es el 49.87%.

Figura 2.7 Distribución estandarizada y sus proporciones de áreas y de probabilidad



Un resumen de la distribución estandarizada es que sirve para estimar la probabilidad de que un valor o una serie de valores estén entre ciertos criterios, que son los valores z que se seleccionen. Un valor z *no* necesariamente tiene que ser un número discreto. Por simplicidad, en la Tabla 2.6 se muestra cuál es la probabilidad de que un número tomado de una población esté entre ciertos criterios. La característica más importante son las proporciones fijas que caen dentro de ciertas secciones.

Tabla 2.6 Valores z_i y su interpretación

Valores z_i	Densidad del área	Porcentaje	Interpretación
-1 a 1	$.3413 + .3413 = .6826$	68.26%	Existe una probabilidad del 68.26% de que un valor tomado al azar, se encuentre entre los valores z_i de -1 a 1
-2 a 2	$.4772 + .4772 = .9544$	95.44%	Existe una probabilidad del 95.44% de que un valor tomado al azar, se encuentre entre los valores z_i de -2 a 2
-3 a 3	$.4987 + .4987 = .9974$	99.74%	Existe una probabilidad del 99.74% de que un valor tomado al azar, se encuentre entre los valores z_i de -3 a 3

Un ejemplo de lo anterior, es la ubicación de la calificación de un estudiante en la distribución estandarizada. Si el alumno obtuvo una calificación de 9 en una materia donde el promedio fue de 6, con una $SD = 1.5$, usando la Ecuación (2.3) se tiene:

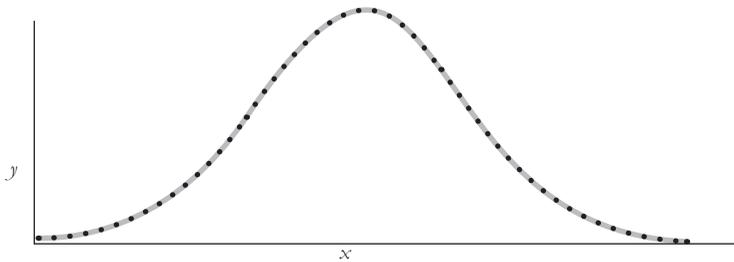
$$z_i = \frac{(9-6)}{1.5} = 3 / 1.5 = 2.$$

En este caso, la calificación del estudiante se encuentra por encima del promedio a dos desviaciones estándar. Otra manera de ver esto es que .9772 (.50 + .4772; *i.e.*, 97.72%) de las calificaciones de las y los alumnos está por debajo de la calificación de 9. Por lo tanto, el 2.28% (100 – 97.72%) está por encima de la calificación de 9. Todo lo anterior es bajo el supuesto de una distribución normal de los datos.

Fórmulas de la distribución normal

CADA PUNTO EN LA LÍNEA DE LA *DISTRIBUCIÓN NORMAL ESTAN-*
darizada, puede ser calculado de la misma manera que los puntos en la curva normal (en la Figura 2.8, se muestran algunos de los puntos que pueden ser calculados mediante la Ecuación 2.4). Similarmente, la *distribución normal estandarizada* tiene la propiedad de que los valores de y son en función (*i.e.*, dependen de) de los de x .

Figura 2.8 Puntos en la curva estandarizada



En otras palabras, si se toma un valor de x , se puede calcular el valor correspondiente de y usando la fórmula de la *distribución normal estandarizada*, que se muestra a continuación (Ecuación 2.4 obtenida de Excel):

$$\text{Fórmula de la distribución normal estandarizada: } y_i = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{z_i^2}{2}\right)}, \text{ (Ecuación 2.4)}$$

donde:

- y_i = Altura de la curva para cada uno de los valores de x_i (en un plano cartesiano);
- π = Valor de 3.1416... que es una constante y representa la razón de la circunferencia al diámetro de un círculo. Esto es, existe un diámetro por cada 3.1416... diámetros: 1 : 3.1416;
- e = Valor constante de 2.7183... que es la base de los logaritmos naturales;
- z_i = Valor de la variable independiente, *i.e.*, es la longitud (el eje horizontal).

El uso de esta distribución estandarizada economiza el tener que calcular manualmente los valores de las áreas de probabilidad correspondientes y, también, evita la integración de la función para calcular el área bajo la curva (*i.e.*, cálculo integral; Ecuación 2.5). La siguiente fórmula de cálculo integral es para la *distribución normal estandarizada* (véase: Spiegel, Schiller, & Srinivasan [2000], para ejemplos al respecto):

$$P(-\infty, +\infty) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 1, \quad \text{(Ecuación 2.5)}$$

donde:

- $P(-\infty, +\infty)$ = Tamaño del área desde el infinito negativo hasta el infinito positivo representado por este intervalo;
- z_c = Valor z que sirve como criterio o valor calculado;
- $\pi \approx 3.1416\dots$
- $e \approx 2.72\dots$

En síntesis, esta fórmula indica que la probabilidad desde el infinito negativo hasta el infinito positivo, es de 1 (100%). En otras palabras, esta probabilidad del 100% indica toda el área de la curva normal estandarizada. Con la Ecuación 2.5, se podría calcular cualquier área dentro de la distribución estandarizada al especificar dos puntos, e.g., del 0.5 al 1.5. Haciendo una observación, el cálculo de un área bajo una curva requiere de dos puntos o límites para poderse llevar a cabo. Si solo se tiene un punto, no se puede calcular.

Preguntas para reflexionar

1. ¿Qué otros fenómenos se pueden representar con la distribución binomial?
2. ¿Por qué se usa la distribución normal en lugar de la binomial, en algunos casos?
3. ¿Se pueden estandarizar todos los valores de un set de datos?
4. ¿De qué otras maneras se podrían organizar los datos para estimar sus probabilidades?
5. ¿En qué tipo de investigación educativa serían más útiles los valores z ?
6. ¿Qué se podría hacer si los datos aparentemente no siguen una distribución normal?
7. ¿Cómo se podría sacar el área bajo la curva sin utilizar la tabla de valores z ?

Opinión del autor

EN LA INVESTIGACIÓN EDUCATIVA, EL USAR LAS PROBABILIDADES, combinatorias y distribuciones binomiales puede llegar a ser muy útil. Esto es, porque se estaría utilizando una teoría de la probabilidad. Por ejemplo, se podría calcular la probabilidad de un ítem de un examen de matemáticas, que podría ser una afirmación de cierto o falso. Entonces, la o el estudiante tendría un 50% de probabilidad de acertarlo. Por otro lado, si es un ítem de opción múltiple con cuatro opciones, la probabilidad de acertarlo sería de un 25%, bajo el supuesto de no saber la respuesta correcta. Puede ser que un(a) docente quiera minimizar la parte de adivinar por parte de los que no saben tanto y vaya por el ítem de opción múltiple.

Lo importante para la o el lector es que la distribución binomial se construyó de estas probabilidades y combinatorias. Esta distribución binomial es la base para construir la distribución normal, que sirve como modelo para una serie de fenómenos educativos, *e.g.*, aprendizaje, coeficiente intelectual, etcétera.

En muchos casos, el modelo de la distribución normal puede ser utilizado para la investigación educativa. Sin embargo, hay que cerciorarse antes de usarlo, ya que los datos siguen una distribución aproximadamente normal, como parte de ser una buena o un buen investigador educativo. Esto se puede hacer cuando nuestros datos tienen ciertas características de medidas de tendencia central similares, cierta dispersión de los datos y coeficientes aceptables de curtosis y simetría. Asimismo, el Test de Kolmogorov-Smirnov es una respuesta más robusta a la hora de inferir la distribución normal de los datos. Dada la situación de que estos no sigan una distribución normal, hay maneras para corregirlos.

La distribución normal estandarizada es un modelo que se emplea en lugar de la distribución normal, porque esta puede variar su forma dependiendo del valor del promedio y la desviación estándar. Lo que surge de la variación en los promedios y la desviación estándar es una familia de distribuciones normales, pero que no resultan prácticas a la hora de calcular las probabilidades, porque involucran el uso del cálculo integral. Por estas razones y bajo el supuesto de que los datos siguen más o menos una distribución normal, se utiliza la distribución normal estandarizada.

CAPÍTULO 3

Correlación

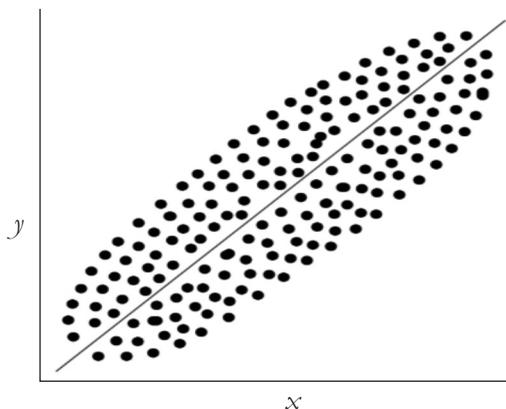
En la investigación educativa, suele pasar que la investigadora o el investigador tengan una pregunta acerca del posible vínculo entre dos cosas: ¿existe una relación entre la variable x y la variable y ? Estas dos variables pueden ser el aprendizaje y la edad de un set de estudiantes. De entrada, se podría especular que entre más joven sea el alumno aprenderá más, asumiendo que los estudiantes tienen más o menos el mismo coeficiente intelectual, nivel socioeconómico, motivación, autoestima, etcétera. En este tipo de situación, un análisis de correlación con dos variables es una opción para ver qué tan relacionadas están ambas. Por otro lado, puede haber correlaciones espurias, *i.e.*, falsas, lo que puede pasar especialmente cuando no hay antecedentes o teorías que asocien un par de variables, o no se pueda formar un argumento válido y sólido para explicar dicha relación. Por ejemplo, si se correlaciona el aprendizaje en alguna materia con el tamaño de la mano de un alumno, podría tener una *gran* correlación. Sin embargo, ¿qué teoría o antecedente podría explicar esta gran correlación? Posiblemente no habría ninguno. Entonces, se tendría una correlación espuria que se debió probablemente a la casualidad.

Asimismo, puede surgir la posibilidad de usar una comparación entre un coeficiente de correlación encontrado en una muestra y el coeficiente de su población correspondiente. Para dicha comparación, se puede utilizar un test estadístico y resolver una interrogante de investigación como esta: ¿existe una diferencia estadísticamente significativa entre la estadística de una muestra y un parámetro de la población? Cabe recordar que las muestras contienen estadísticas que son estimaciones de los parámetros de las poblaciones, *e.g.*, promedio, mediana, moda, varianza, desviación estándar, etcétera.

a. Correlaciones r de Pearson y policóricas

CUANDO SE ASOCIAN DOS VARIABLES DE UNA BASE DE DATOS, se les llama datos bivariados. Un coeficiente de correlación en general corresponde a esta descripción de datos bivariados. El Coeficiente r de Pearson (*i.e.*, *Coficiente de Correlación Producto-Momento de Pearson*) y el Coeficiente de Correlaciones Policóricas corresponden a datos bivariados, entre otros coeficientes, que se pueden representar en un plano cartesiano (*i.e.*, el cuadrante donde la x y la y son variables positivamente relacionadas; Figura 3.1) en donde cada punto corresponderá a un par de coordenadas, *i.e.*, x y y .

Figura 3.1 Datos bivariados



Coefficiente r de Pearson

EL COEFICIENTE R DE PEARSON ES “UN ÍNDICE DEL GRADO DE RELACIÓN LINEAL ENTRE DOS VARIABLES” (APA, 2015, p. 841). Para estimar la correlación dentro de poblaciones con este coeficiente, se emplea la letra griega ρ (se pronuncia: *ro*) y para las muestras, la letra r . Más al respecto, se le suelen agregar dos letras p (*i.e.*, producto-Pearson) a la ρ para distinguirla de otras correlaciones de poblaciones: ρ_{pp} . Recordando: las estadísticas (promedio, desviación estándar, r , etcétera) son estimaciones que se calculan de las muestras y contienen un error de muestreo, y de las poblaciones se obtienen parámetros (media, desviación estándar, ρ , etcétera) que son fijos. Se ha dicho que el Coeficiente de Correlación r de Pearson, fue diseñado para calcular la relación entre dos variables continuas, como la estatura y el peso, entre muchas otras similares. Estas variables están en una escala continua, porque siempre hay un punto intermedio entre dos números, *e.g.*, entre 2.00110 y 2.00120 está 2.00115, entre otros más. Además, poseen un cero absoluto que significa sin cantidad alguna (*e.g.*, la velocidad de un auto estacionado es de cero kilómetros por hora). Sin embargo, este coeficiente se utiliza bastante en la investigación educativa para relacionar variables que fueron medidas en una escala tipo Likert, la cual ha sido catalogada en ocasiones

como ordinal o de intervalo. *Recordando*: una escala ordinal no tiene necesariamente el mismo espacio entre los números que la componen, pero, en cambio, una de intervalo sí posee la misma distancia entre sus puntos. Un ejemplo de ambas escalas es un ítem con respuestas de tipo Likert: 1 = Muy en desacuerdo; 2 = En desacuerdo; 3 = Neutral; 4 = De acuerdo; y 5 = Muy de acuerdo. Esta escala solo tiene números discretos, porque toma valores enteros ya predeterminados. En contraparte, una escala ordinal o de intervalo no tiene un cero absoluto (*e.g.*, una escala tipo Likert).

Correlaciones policóricas

APARTE DEL EMPLEO DEL COEFICIENTE R EN ENCUESTAS, JÖRESkog y Sörbom (1996) han recomendado el uso de las correlaciones policóricas (*polychoric correlations*: ρ_{pc}) cuando se utilizan escalas tipo Likert. Para realizar estas correlaciones policóricas, estos autores crearon el *software* estadístico y psicométrico llamado Lisrel (*i.e.*, *Linear Structure Relations*: Relaciones lineales estructurales), que es usado en ecuaciones estructurales (*i.e.*, para medir las relaciones entre varios constructos/variables). *Recordando*: un constructo es una variable que no se puede ver directamente, pero que teóricamente se manifiesta en algún test o encuesta (*e.g.*, la inteligencia es un *constructo* y se expresa en los puntajes de un test de inteligencia). Asimismo, Holgado-Tello *et al.* (2010) compararon el Coeficiente *r* de Pearson y las correlaciones policóricas en el contexto de procesos de validación (*i.e.*, análisis de factores exploratorios y confirmatorios). Estos autores concluyeron que las correlaciones policóricas son más precisas al calcular la correlación *real* entre variables. Cuando se habla de valores verdaderos o reales de relación entre variables, se refiere a los vínculos —relaciones— a nivel de poblaciones. Otra manera de decirlo: la correlación real es la que *existe* a nivel de población.

La recomendación para una investigadora o un investigador es que empleen las correlaciones policóricas cuando se tenga una escala tipo Likert, porque son más precisas para estimar los valores verdaderos de la relación entre variables de las poblaciones. Los coeficientes de correlación policórica son un índice que muestra el grado de asociación entre dos variables que fueron medidas con una escala ordinal o de inter-

valo, pero que se asume que son manifestaciones de una variable continua que subyace (*cf.* Jöreskog y Sörbom, 1996). Por ejemplo, el constructo de la autoestima se podría medir con una serie de ítems derivados de alguna de sus definiciones. La autoestima *subyace* a los ítems que se derivaron de su definición y tendría una escala continua, pero se vuelve ordinal o de intervalo con una escala tipo Likert cuando se administra a las personas, asumiendo que no hay alguna tecnología que pueda medir la autoestima por medio de la actividad cerebral en lugar de una simple encuesta, por ejemplo.

El error de muestreo de los coeficientes

COMO SE MANEJAN MUESTRAS, EL COEFICIENTE R DE PEARSON ES una simple estimación de un parámetro (*i.e.*, la correlación a nivel de poblaciones). En resumen, se puede asumir que hay una diferencia de las correlaciones a nivel de poblaciones de las que existen a nivel de una muestra. Por lo tanto, esta diferencia entre este parámetro de la población y la estadística de la muestra es llamada *error de muestreo*. Más al respecto, la APA (2015, pp. 932-933) llamó a este error margen de error predecible, *e.g.*, un margen de error del 5% para un promedio hipotético calculado en un 70% de una muestra, indica que el valor verdadero de la población estaría entre el 65% y el 75%. En pocas palabras, el error se puede resumir de la siguiente manera:

- $\rho_{pp} - r = \text{Error de muestreo de la correlación } r$;
- $\rho_{pp} = \text{Correlación entre dos variables de una población}$;
- $r = \text{Correlación entre dos variables de una muestra}$.

Nota: recordando que los parámetros corresponden a las poblaciones y las estadísticas (*i.e.*, estimaciones), a las muestras.

Asimismo, entre la correlación policórica habría un error entre el parámetro de la población y la estadística de la muestra:

$$\rho_{pc} - \text{Correlación policórica de una muestra} = \text{Error de correlación policórica.}$$

Enfatizando, según Holgado-Tello *et al.* (2010) y Jöreskog y Sörbom (1996), las correlaciones policóricas son más precisas al calcular el verdadero parámetro que el r . Por lo tanto, el error de las correlaciones policóricas es más pequeño por su precisión:

$$\text{Error de } r > \text{Error de correlación policórica de la muestra.}$$

Para estimar este error de r , se utilizan intervalos de confianza (IC), como se había mencionado. Más adelante en este capítulo, se explica cómo se obtiene un IC para el r , al igual de cómo se interpreta.

Un *software* gratuito que se puede emplear para calcular las relaciones policóricas es R (*i.e.*, The R Project for Statistical Computing: <https://www.r-project.org/>). Este programa no es tan amigable como Lisrel, pero existen cursos en Coursera (<https://es.coursera.org/>), entre otras páginas, en donde se puede aprender a usar R.

b. Características del r

VOLVIENDO AL R DE PEARSON, ESTE TIENE DOS LÍMITES (-1 Y +1) Y en medio de estos, se le considera una variable continua. Se utiliza con escalas continuas, ordinales y de intervalo (véase la discusión previa). Hay dos condiciones que, según Hinkle *et al.* (2003), se deben considerar para usar el r :

1. Los datos de las dos variables que se estén correlacionando *deben* estar en pares para el mismo set de observaciones, personas u objetos. Un ejemplo de lo anterior es cuando tres estudiantes toman dos test de ciencias naturales (1.º y 2.º) y sus porcentajes son los siguientes (Tabla 3.1):

Tabla 3.1 Datos en pares para usarlos en correlaciones

Estudiante / Exámenes de ciencias naturales	1.º	2.º
Miguel	100%	99%
María	82%	84%
Pablo	79%	76%

En este caso, se pueden correlacionar los dos test, porque están en pares, *i.e.*, cada par corresponde a uno de los alumnos. Un contraejemplo de lo anterior sería intentar correlacionar los porcentajes de Miguel, María y Pablo con los de otros estudiantes que hayan tomado un par de exámenes diferentes. En este caso, los otros porcentajes no corresponden a los pares de Miguel, María y Pablo.

2. Se utiliza el promedio y la varianza para calcular el r , así que las variables deben estar en una escala continua o de intervalo. De nuevo, suele suceder en las investigaciones donde se emplean frecuentemente las escalas tipo Likert, que estas se utilizan como si fueran de intervalo (*i.e.*, la distancia entre el 1 y el 2 es la misma que del 3 al 4).

Escenarios para correlaciones

EN LA FIGURA 3.2, SE MUESTRAN ALGUNOS DE LOS POSIBLES ESCENARIOS cuando se trata de obtener una correlación entre dos variables con el r . En el escenario A del plano cartesiano, se plantea el límite bajo del r de Pearson ($r = -1$); por lo tanto, es una correlación perfectamente negativa (todos los puntos caen en una línea), ya que cuando los valores de la horizontal (x) aumentan, los de la vertical (y) disminuyen, y todos están perfectamente alineados. Esta línea antes mencionada se llama *línea de mejor ajuste (line of best fit)*, con la cual se trata de minimizar la distancia entre ella y los puntos de los pares ordenados, porque entre menor sea la distancia entre esta línea y los puntos, menor será el error del Coeficiente r . Un error implica imprecisión y, por ello, es algo que se busca minimizar en las estadísticas en general. Un tema sobre el error será abordado con más detalle en la sección del Coeficiente de Determinación (r^2). Por otro lado, en el escenario B se muestra una relación perfectamente positiva

donde todos los valores de x y de y aumentan ($r = 1$), por lo que están perfectamente alineados con la línea de mejor ajuste.

Un caso diferente a los dos anteriores es el escenario C, donde los datos forman una nube en la cual se podrían trazar una infinidad de líneas, dando como resultado que el $r = 0$, es decir, la correlación es nula.

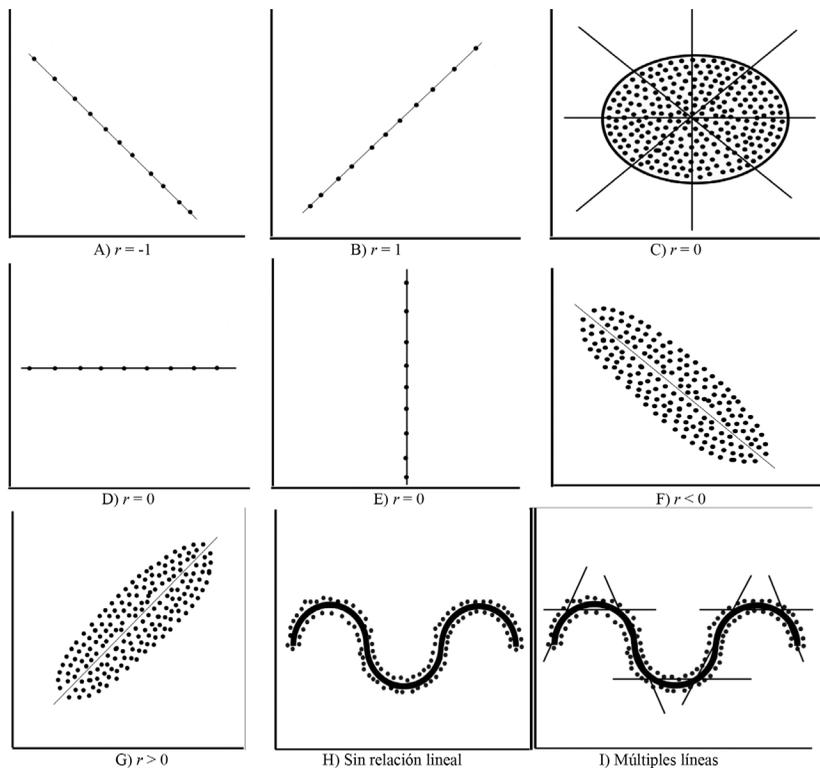
En el escenario D sí hay un patrón, porque los puntos se alinean, pero solo los valores de x aumentan, mientras que los de y (*i.e.*, más bien, el valor de la y) se quedan fijos. Por lo tanto, el $r = 0$, porque ambas variables necesitan variar para que $r \neq 0$. El escenario E es similar al anterior, en el sentido de que hay un patrón porque los puntos se alinean, pero de igual manera solamente los valores de una de las variables cambian (*i.e.*, y) y los de la otra, no (*i.e.*, x); entonces, el r es igual a 0.

En el escenario F, los valores no se alinean perfectamente, pero forman un patrón que es evidenciado con la elipse que los rodea. En este caso, los valores de x aumentan y los de y disminuyen. Por lo tanto, es una relación negativa. Como no todos los valores están en la línea, el r está entre -1 y 0 (*i.e.*, $-1 < r < 0$). Al trazar la línea de mejor ajuste, se trata de minimizar las distancias entre esta y todos los puntos. En análisis más sofisticados, como la regresión, se pormenoriza este error, para encontrar el modelo que mejor explique los datos que se están analizando. Similarmente, en el escenario G se muestra un patrón positivo entre los valores de x y los de y . No es una relación perfecta, porque no todos los valores caen en la línea de mejor ajuste, pero sí es mayor que 0 y menor que 1 (*i.e.*, $0 < r < 1$).

En el escenario H, se muestra una relación no lineal cuando se toman en cuenta todos los puntos alrededor de la curva cíclica (*i.e.*, sube y baja desde el mismo patrón). Aunque estos puntos forman un patrón, no se usaría el r , porque esta estadística requiere una correlación lineal, ya sea positiva o negativa, que no existe en este ejemplo. Ahora, en el escenario I se muestra que se podrían trazar varias líneas a lo largo del ciclo. Por ejemplo, cuando los valores de x y los de y aumentan, se puede trazar una línea positiva en este segmento de los datos y calcular un r . En concreto, los datos que forman un patrón cíclico pueden ser utilizados para trazar varias líneas positivas, negativas (*i.e.*, para obtener coeficientes r) y paralelas al eje de las x , como en este caso. Las líneas paralelas al eje de las x tienen un $r = 0$. Sin embargo, habría que ver otros

aspectos antes de usar uno o varios coeficientes de r para modelar los datos (*i.e.*, modelar es crear líneas que sigan los datos en estos casos). Entre estos están: a) buscar si hay otros modelos *no* lineales donde los datos queden mejor (para más información acerca de modelos no lineales, véase: Berenson *et al.*, 2012); y b) estudiar qué dicen las teorías psicológicas y pedagógicas de las relaciones entre las variables (*e.g.*, lineales durante cierto periodo, como las etapas de desarrollo de una persona empleando la Teoría de Piaget).

Figura 3.2 Algunos posibles escenarios para usar el r de Pearson



c. Fórmulas del Coeficiente r

CUANDO SE USAN LAS CORRELACIONES DEL COEFICIENTE R DE Pearson, no necesariamente se especifica cuál es la variable independiente y cuál es la dependiente. Por simplicidad, las siguientes fórmulas se explican con la letra x (variable independiente) y la letra y (variable dependiente). Esta designación de las variables también corresponde al plano cartesiano, donde la x está en la horizontal y la y , en la vertical. Existen varias fórmulas para calcular el Coeficiente r . Para estas, es necesario tener dos sets de datos en pares (*i.e.*, dos variables). Todas estas fórmulas darían exactamente el mismo resultado:

a) Fórmula de la r de Pearson con valores z (Ecuación 3.1):

$$r_{xy} = \frac{(\sum z_x z_y)}{n - 1}, \quad (\text{Ecuación 3.1})$$

donde:

- \sum = Suma;
- $z_x = (x_i - \bar{x}) / SD_x$; (Ecuación 2.3)
- x_i = Cada valor individual de un set de datos llamado x ;
- \bar{x} = Promedio del set de datos llamado x ;
- SD_x = Desviación estándar del set de datos llamado x ;
- $z_y = (y_i - \bar{y}) / SD_y$;
- y_i = Cada valor individual de un set de datos llamado y ;
- \bar{y} = Promedio del set de datos llamado y ;
- SD_y = Desviación estándar del set de datos llamado y ;
- n = Tamaño de la muestra.

b) Fórmula de la desviación de los escores (*Deviation Score Formula*; Ecuación 3.2):

$$r_{xy} = \frac{(\sum x_{des} \cdot y_{des})}{\sqrt{(\sum x_i^2 \sum y_i^2)}} \quad (\text{Ecuación 3.2})$$

donde:

- \sum = Suma;
- $x_{des} = x_i - \bar{x}$;
- $y_{des} = y_i - \bar{y}$;
- $x_i^2 = x_i \times x_i$;
- $y_i^2 = y_i \times y_i$.

c) Fórmula del Escore Crudo (*Raw Score Formula*; Ecuación 3.3):

$$r_{xy} = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n \sum x_i^2 - (\sum x_i)^2][n \sum y_i^2 - (\sum y_i)^2]}} \quad (\text{Ecuación 3.3})$$

d) Fórmula con el uso de la covarianza (Ecuación 3.4) para encontrar el r (Ecuación 3.5):

$$\text{Covarianza} = S_{xy} = \frac{(\sum x_i - \bar{x})(y_i - \bar{y})}{n - 1}; \quad (\text{Ecuación 3.4})$$

$$r_{xy} = \frac{S_{xy}}{SD_x SD_y} \quad (\text{Ecuación 3.5})$$

Demostración con las fórmulas

PARA DEMOSTRAR QUE LAS FÓRMULAS ANTERIORES DEL R DAN EL mismo resultado, se creó una base de datos. Con este ejemplo, se explica que un grupo de alumnos presentó un test de matemáticas en cierto periodo. Es decir, esta base de datos contiene dos variables en pares (*i.e.*, la variable independiente: x = Tiempo en completar el test; y la variable dependiente: y = Puntajes en un test de matemáticas). En síntesis, la pregunta de investigación para este ejemplo es: ¿existe una correlación

entre la variable independiente del *Tiempo en completar el test* y la variable dependiente *Puntajes en un test de matemáticas*? El resultado del uso de las cuatro fórmulas aparece en las tablas 3.2, 3.3, 3.4 y 3.5, *i.e.*, $r = .98$.

Tabla 3.2 Fórmula del r de Pearson con valores z

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	Tiempo en valor: $z_{xi} = (x_i - \bar{x}) / SD_x$	Matemáticas valor: $z_{yi} = (y_i - \bar{y}) / SD_y$	$z_{xi} \times z_{yi}$
1	13	60	-1.96	-1.62	3.18
2	14	62	-1.47	-1.47	2.16
3	14	62	-1.47	-1.47	2.16
4	15	67	-0.98	-1.08	1.05
5	15	67	-0.98	-1.08	1.05
6	15	67	-0.98	-1.08	1.05
7	16	72	-0.49	-0.68	0.34
8	16	72	-0.49	-0.68	0.34
9	16	72	-0.49	-0.68	0.34
10	16	72	-0.49	-0.68	0.34
11	17	81	0.00	0.02	0.00
12	17	81	0.00	0.02	0.00
13	17	81	0.00	0.02	0.00
14	17	81	0.00	0.02	0.00
15	17	81	0.00	0.02	0.00
16	18	90	0.49	0.72	0.35
17	18	90	0.49	0.72	0.35
18	18	90	0.49	0.72	0.35
19	18	90	0.49	0.72	0.35
20	19	95	0.98	1.11	1.09
21	19	95	0.98	1.11	1.09
22	19	95	0.98	1.11	1.09
23	20	98	1.47	1.35	1.98

Continúa...

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	Tiempo en valor: $z_{xi} = (x_i - \bar{x}) / SD_x$	Matemáticas valor: $z_{yi} = (y_i - \bar{y}) / SD_y$	$\sum z_{xi} X z_{yi}$
24	20	98	1.47	1.35	1.98
25	21	100	1.96	1.50	2.95
\bar{x}	17	80.76		Suma	23.61
Mediana	17	81		$n - 1$	24.00
Moda	17	81		r	0.98
Varianza	4.17	163.52			
sd	2.04	12.79			
Rangos					
Mínimo	13	60			
Máximo	21	100			
Diferencia	8	40			
n	25				

Nota: Estu. = Estudiante; Fórmula del r de Pearson con valores z

$$r_{xy} = (\sum z_{xc} z_{yc}) / n - 1; \quad \text{(Ecuación 3.1)}$$

sustituyendo:

$$r_{\text{Tiempo y Matemáticas}} = 23.61 / (25 - 1) = .98.$$

Tabla 3.3 Fórmula del r de Pearson con la desviación de los scores

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	Desviación de x_i del promedio: $x_{des} = (x_i - \bar{x})$	Desviación de y_i del promedio: $y_{des} = (y_i - \bar{y})$	x_{des} por y_{des}	x^2	y^2
1	13	60	-4	-20.76	83.04	16	430.98
2	14	62	-3	-18.76	56.28	9	351.94
3	14	62	-3	-18.76	56.28	9	351.94
4	15	67	-2	-13.76	27.52	4	189.34

Continúa...

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	Desviación de x_i del promedio:	Desviación de y_i del promedio:	x_{des} por y_{des}	x^2	y^2
			$x_{des} = (x_i - \bar{x})$	$y_{des} = (y_i - \bar{y})$			
5	15	67	-2	-13.76	27.52	4	189.34
6	15	67	-2	-13.76	27.52	4	189.34
7	16	72	-1	-8.76	8.76	1	76.74
8	16	72	-1	-8.76	8.76	1	76.74
9	16	72	-1	-8.76	8.76	1	76.74
10	16	72	-1	-8.76	8.76	1	76.74
11	17	81	0	0.24	0	0	0.06
12	17	81	0	0.24	0	0	0.06
13	17	81	0	0.24	0	0	0.06
14	17	81	0	0.24	0	0	0.06
15	17	81	0	0.24	0	0	0.06
16	18	90	1	9.24	9.24	1	85.38
17	18	90	1	9.24	9.24	1	85.38
18	18	90	1	9.24	9.24	1	85.38
19	18	90	1	9.24	9.24	1	85.38
20	19	95	2	14.24	28.48	4	202.78
21	19	95	2	14.24	28.48	4	202.78
22	19	95	2	14.24	28.48	4	202.78
23	20	98	3	17.24	51.72	9	297.22
24	20	98	3	17.24	51.72	9	297.22
25	21	100	4	19.24	76.96	16	370.18
\bar{x}	17	80.76		Suma	616	100	3924.56
						Multiplicación de x^2 por y^2	392 456
						Raíz cuadrada	626.46
						r	.98

Nota: Estu. = Estudiante; Fórmula del r de Pearson con la desviación de los escores (*Deviation Score Formula*):

$$r_{xy} = (\sum x_{des} y_{des}) / \sqrt{\sum x_i^2 \sum y_i^2}; \quad \text{(Ecuación 3.2)}$$

sustituyendo:

$$r_{\text{Tiempo y Matemáticas}} = 616 / 626.46 = .98.$$

Tabla 3.4 Fórmula del escore crudo

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	x_i por y_i	x^2	y^2
1	13	60	780	169	3600
2	14	62	868	196	3844
3	14	62	868	196	3844
4	15	67	1005	225	4489
5	15	67	1005	225	4489
6	15	67	1005	225	4489
7	16	72	1152	256	5184
8	16	72	1152	256	5184
9	16	72	1152	256	5184
10	16	72	1152	256	5184
11	17	81	1377	289	6561
12	17	81	1377	289	6561
13	17	81	1377	289	6561
14	17	81	1377	289	6561
15	17	81	1377	289	6561
16	18	90	1620	324	8100
17	18	90	1620	324	8100
18	18	90	1620	324	8100
19	18	90	1620	324	8100
20	19	95	1805	361	9025
21	19	95	1805	361	9025
22	19	95	1805	361	9025
23	20	98	1960	400	9604
24	20	98	1960	400	9604
25	21	100	2100	441	10 000
Suma	425	2019	34 939	7325	166 979

Nota: Estu. = Estudiante; Fórmula del escore crudo (*Raw Score Formula*):

$$r_{xy} = n (\sum x_i y_i) - (\sum x_i \sum y_i) / \sqrt{[n \sum x_i^2 - (\sum x_i)^2] [n \sum y_i^2 - (\sum y_i)^2]} ; \text{ (Ecuación 3.3)}$$

sustituyendo:

$$r_{\text{Tiempo y Matemáticas}} = 25(34\ 939) - (425 \times 2019) / \sqrt{[25 (7,325) - (425)^2] [25(166,979) - (2,019)^2]}$$

$$r_{\text{Tiempo y Matemáticas}} = 15\ 400 / \sqrt{2,500 \times 9,814} = .98.$$

Tabla 3.5 Usando la covarianza para encontrar el *r* de Pearson

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	Desviación de xi del promedio: x des. = (xi - \bar{x})	Desviación de yi del promedio: y des. = (yi - \bar{y})	x des. por y des.
1	13	60	-4	-20.76	83.04
2	14	62	-3	-18.76	56.28
3	14	62	-3	-18.76	56.28
4	15	67	-2	-13.76	27.52
5	15	67	-2	-13.76	27.52
6	15	67	-2	-13.76	27.52
7	16	72	-1	-8.76	8.76
8	16	72	-1	-8.76	8.76
9	16	72	-1	-8.76	8.76
10	16	72	-1	-8.76	8.76
11	17	81	0	0.24	0
12	17	81	0	0.24	0
13	17	81	0	0.24	0
14	17	81	0	0.24	0
15	17	81	0	0.24	0
16	18	90	1	9.24	9.24
17	18	90	1	9.24	9.24

Continúa...

Estu.	Tiempo en completar el test	Puntajes en un test de matemáticas	Desviación de xi del promedio: x des. = $(x_i - \bar{x})$	Desviación de yi del promedio: y des. = $(y_i - \bar{y})$	x des. por y des.
18	18	90	1	9.24	9.24
19	18	90	1	9.24	9.24
20	19	95	2	14.24	28.48
21	19	95	2	14.24	28.48
22	19	95	2	14.24	28.48
23	20	98	3	17.24	51.72
24	20	98	3	17.24	51.72
25	21	100	4	19.24	76.96
\bar{x}	17	80.76		Suma	616
sd	2.04	12.79		$n - 1$	24
				Covarianza	25.67
				SD _x	2.04
				SD _y	12.79
				r	.98

Nota: *Estu.* = Estudiante; usando la covarianza para encontrar el *r*, las fórmulas son:

$$Covarianza = S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) / n - 1; \text{ y } r_{xy} = S_{xy} / SD_x SD_y; \text{ (Ecuación 3.5)}$$

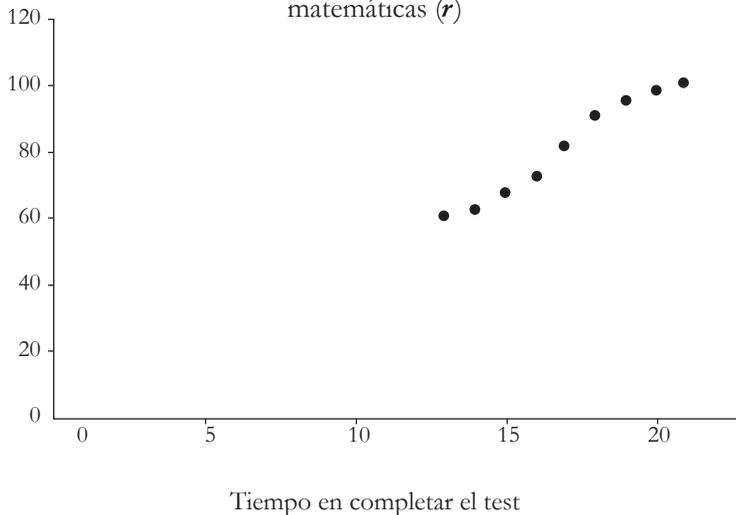
sustituyendo:

$$Covarianza_{Tiempo y Matemáticas} \text{ (Ecuación 3.4)} = 616 / 24 = 25.67;$$

$$\text{ y } r_{Tiempo y Matemáticas} = 25.67 / 2.04 \times 12.79 = .98.$$

En la Figura 3.3, se muestra el histograma de la correlación entre la variable independiente *Tiempo en completar el test* y la variable dependiente *Puntajes en un test de matemáticas*. Se puede observar que dichas variables tienen una relación positiva, ya que ambas aumentan y tienen un Coeficiente $r = .98$.

Figura 3.3 Histograma del Tiempo en completar el test y Puntajes en un test de matemáticas (r)



Más detalles del Coeficiente r de Pearson

EL COEFICIENTE R DE PEARSON INDICA LA DIRECCIÓN Y LA FUERZA de la relación entre dos variables. También, el r es considerado como un *tamaño de efecto*. En este sentido, Cohen (1988) manifestó tres tamaños diferentes para la evaluación de un r cuando se calcula en un estudio: $|.10|$ = Pequeño; $|.30|$ = Mediano; y $|.50|$ = Grande. Por otro lado, Hinkle *et al.* (2003, p. 109) expresaron una clasificación al r que fue diferente, porque incluyó rangos para los tamaños de efecto también como otras categorías para los mismos: 0 al $|.30|$ = De nada a pequeño; $|.30|$ al $|.50|$ = Pequeño; $|.50|$ al $|.70|$ = Moderado; $|.70|$ al $|.90|$ = Efecto alto; y $|.90|$ al $|1.00|$ = Efecto muy alto. La recomendación de lo anterior es que, al efectuar una investigación y evaluar el r , se revise la literatura pertinente para ver cómo sus autores han definido los diferentes tamaños de efecto en ese campo de estudio. Con base en la literatura pertinente, una investigadora o un investigador pueden clasificar sus resultados del r de acuerdo con los tamaños de efecto definidos anteriormente.

La r de Pearson *no* indica relación de causa y efecto en sí misma. De nuevo, este coeficiente indica la fuerza de la relación entre dos variables, *i.e.*, su nivel de asociación. Por un lado, la *causalidad* es una inferencia que se obtiene del diseño del estudio, *i.e.*, experimental. De un experimento o cuasi-experimento, se puede concluir que un tratamiento tuvo cierto efecto en los resultados. Para más información acerca de diseños experimentales, se recomienda consultar a Schneider *et al.* (2007). Por otro lado, otra manera de inferir causalidad es cuando existe una teoría que explique las relaciones entre las variables. Para ello, se usan las ecuaciones estructurales y se recomienda consultar a Byrne (2009) y a Jöreskog y Sörbom (1996), para estos modelos teóricos de causa y efecto cuando se emplean *correlaciones*.

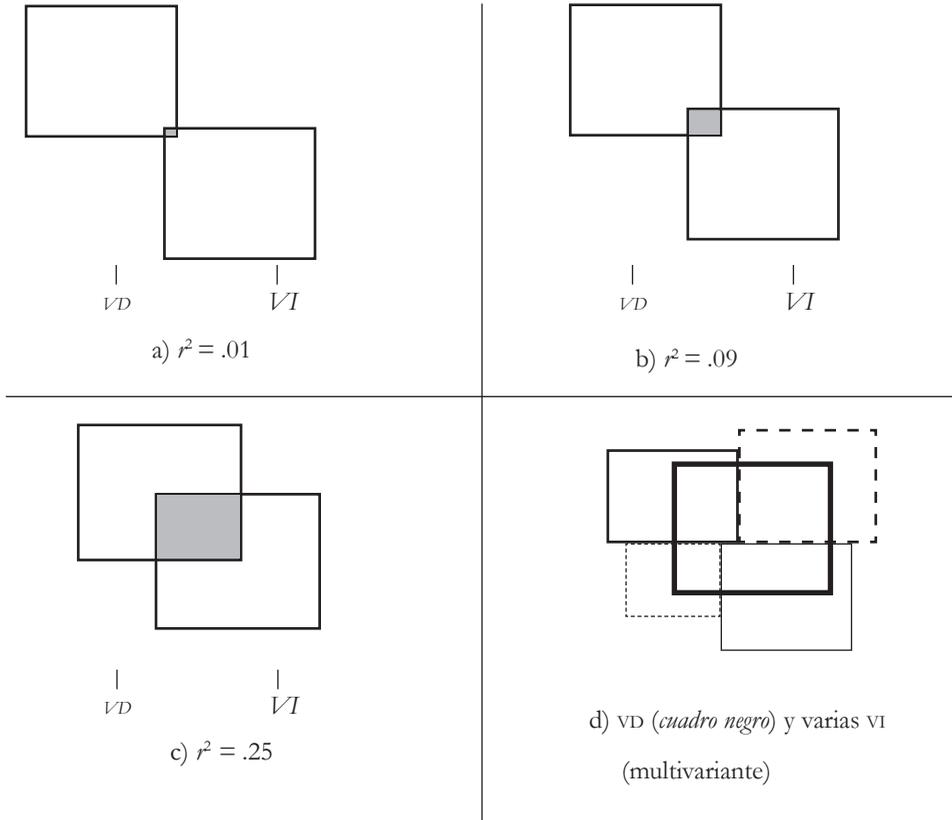
d. Coeficiente de Determinación (r^2) (*Determination Coefficient*) y el r

EL COEFICIENTE R DE PEARSON ES LA BASE PARA CALCULAR EL COEFICIENTE DE DETERMINACIÓN (r^2). Para calcularlo es necesario elevar al cuadrado el r (*i.e.*, $r \times r = r^2$). Algunos ejemplos de lo anterior son los diferentes tamaños del efecto que Cohen (1988) especificó: un efecto pequeño de $r = |.10|$ al cuadrado es: $r^2 = .01$; uno mediano de $r = |.30|$ al cuadrado es: $r^2 = .09$; y uno grande de $r = |.50|$ al cuadrado es: $r^2 = .25$. También, se puede utilizar la ya mencionada clasificación de Hinkle *et al.* (2003, p. 109) para el tamaño del r^2 , y, también en este caso, se recomienda revisar la literatura del tema, para tener antecedentes de los tamaños y clasificación de este coeficiente en un contexto en particular.

El Coeficiente de Determinación es “Un índice que refleja la proporción de variación en una variable dependiente asociada con una variable independiente” (APA, 2015, p. 200). En otras palabras, el Coeficiente de Determinación es el porcentaje de la varianza en la variable dependiente explicada por la variable independiente (*i.e.*, $r^2 =$ Tamaño de un efecto). Esto se puede representar con cuadros, ya que el Coeficiente de Varianza de una variable es cuadrado. En la Figura 3.4, se muestran los coeficientes del r^2 : .01, .09 y .25. Se puede designar como variable dependiente (VD) el cuadro de la izquierda y como variable independiente (VI), el de la derecha. Como se puede apreciar,

entre mayor es el r^2 resulta una mayor área explicada en la variable dependiente por parte de la independiente.

Figura 3.4 Varianza explicada



Asimismo, habría que considerar que los fenómenos son *multivariantes* (*i.e.*, varias variables independientes explican la varianza en la variable dependiente). Para ilustrar lo anterior, en el último cuadro de la Figura 3.4. se muestra la varianza de la variable dependiente (cuadro con bordes negros y sólidos), cuya área está cubierta con otros cuadros con bordes no sólidos que representan las variables independientes. En este caso, estas explican toda la varianza en la variable dependiente, porque cubren toda su

área, *i.e.*, el 100%. Esto sería un caso ideal, pero en la investigación educativa, al igual que en muchas otras disciplinas, habrá muchas porciones de la varianza de una variable dependiente que no tendrán explicación. A estas áreas sin explicación, se les suele llamar errores (véase: Tabachnick, & Fidell [2012], para una discusión más profunda sobre este tema). Para los ejemplos de la Figura 3.4, los errores o varianza no explicada serían: error = .99 (*i.e.*, 99%) cuando el $r^2 = .01$ (*i.e.*, 1%); error = .91 (*i.e.*, 91%) cuando el $r^2 = .09$ (*i.e.*, 90%); y error = .75 (*i.e.*, 75%) cuando el $r^2 = .25$ (*i.e.*, 25%).

Ejemplo de análisis con varias variables

UN EJEMPLO HEURÍSTICO DE VARIANZA EXPLICADA A NIVEL MULTIVARIANTE sería el siguiente: en un salón de clases, un grupo de estudiantes aprenden alguna materia, *e.g.*, *Idioma español*. Este constructo de aprendizaje del idioma español es tomado como la variable dependiente y se supone que se manifiesta en los *puntajes de evaluaciones* del mismo. Para simplicidad de este ejemplo, solo se emplea un examen del idioma español que sería una manera observable en la que se manifestaría dicho constructo. Al administrar este examen hay una variación en los puntajes (*i.e.*, van desde un puntaje del 0% de aciertos hasta uno del 100%). Se consideran solo dos variables independientes: *Ingreso familiar* y *Nivel educativo* de las madres y padres, las cuales también tienen variación; *i.e.*, las familias tienen diferentes ingresos y las madres y padres tienen diferentes niveles de educación.

Las preguntas de investigación para el r de Pearson son: ¿existe una relación entre la *evaluación del idioma español* y el *nivel educativo* de las madres y padres?; y ¿existe una relación entre la *evaluación del idioma español* y el *ingreso familiar*? Para complementar las anteriores, las interrogantes de investigación del r^2 son: ¿cuánta varianza explica el nivel educativo en la *evaluación del idioma español*?; y ¿cuánta varianza explica el ingreso familiar en la *evaluación del idioma español*? Para resolver estas cuatro preguntas, se hacen dos análisis de correlación con el r de Pearson y dos con el r^2 . Los resultados son los siguientes:

$$\text{Evaluación del idioma español (VD)} \leftrightarrow \text{nivel educativo (VI)}: r = .50 \text{ y el } r^2 = .25;$$

Evaluación del idioma español (VD) \leftrightarrow *ingreso familiar* (VI_2): $r = .71$ y el $r^2 = .50$;

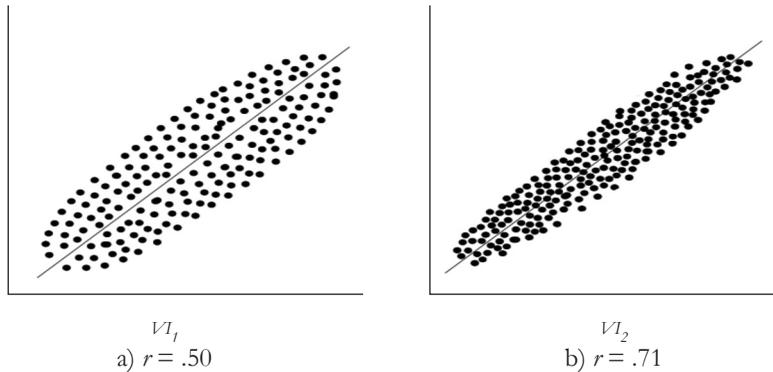
donde:

- VD = Variable dependiente; y
- VI = Variable independiente.

Nota: la flecha con doble cabeza (\leftrightarrow) indica una correlación entre las variables. Una aclaración es que la VI_1 y la VI_2 no tienen correlación: $r = 0$, para hacer el ejemplo más simple.

Las respuestas a las dos primeras preguntas de investigación relacionadas con el r , son contestadas en este párrafo. Según Cohen (1988), estos coeficientes r del ejemplo anterior son grandes, porque son $\geq .50$. Por lo tanto, indican que hay una correlación fuerte entre la variable dependiente y las independientes. También, las correlaciones son positivas, lo cual manifiesta que entre más educación se tenga por parte de las madres y padres mayor será el puntaje en la evaluación. De la misma manera, entre más ingreso familiar se tenga mayor será el puntaje (véase la Figura 3.5 para ambas correlaciones). Aparentemente, como la correlación de la VI_2 con la VD es más fuerte/grande ($r = .71$) que la de la VI_1 con la VD ($r = .50$), los puntos en la Figura 3.5 están más cercanos a la línea de la VI_2 con la VD que los de la VI_1 con la VD . Sin embargo, para comparar estos coeficientes de correlación de una manera *aceptable* en estadística, hay que elevarlos al cuadrado para obtener el r^2 , el cual se utiliza para comparar el grado de varianza explicada por parte de las variables independientes en la dependiente. También, se asumió que la correlación entre el nivel educativo y el ingreso familiar es cero. En otras palabras, cuando se establece que $r = 0$ de manera deliberada, se le llama correlación ortogonal (véase: Salkind [2007]; Thompson [2004], para usos de las relaciones ortogonales en análisis exploratorios de factores).

Figura 3.5 r con evaluación del idioma español con el nivel educativo de las madres y padres, y el ingreso familiar



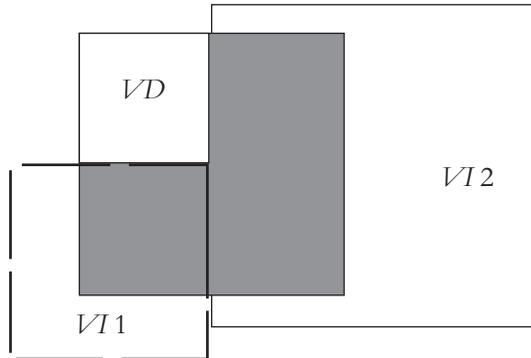
V_D = Evaluación del idioma español;

V_I = Nivel educativo;

V_I = Ingreso familiar.

Las respuestas a las otras dos preguntas de investigación relacionadas con el r^2 , son contestadas a continuación. La parte que complementa el análisis de correlación es el Coeficiente de Determinación (r^2). En la Figura 3.6, se muestra gráficamente la proporción de la varianza explicada en la evaluación del idioma español por nivel educativo e ingreso familiar ($r^2 = .25$ y $r^2 = .50$, respectivamente). Dado que la correlación entre el nivel educativo y el ingreso es de cero, el r^2 es cero también entre ambas variables independientes. Por consiguiente, la varianza que explica cada variable independiente es única (*i.e.*, una varianza explicada no tiene que ver con la otra: correlación *ortogonal*). En otras palabras, los diferentes puntajes en la evaluación del idioma español se asocian positivamente con los diferentes niveles de ingreso, así como con los diferentes niveles de educación de las madres y padres.

Figura 3.6 Varianza explicada por dos variables independientes en la dependiente



Leyenda:

VD con VI_1 ; $r^2 = .25$;

VD con VI_2 ; $r^2 = .50$;

VI_1 con VI_2 ; $r^2 = 0$.

VD = Evaluación del idioma español;

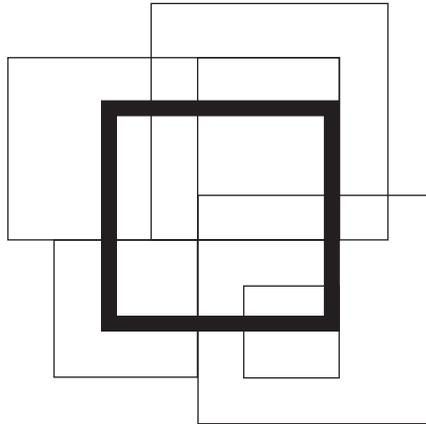
VI_1 = Nivel educativo;

VI_2 = Ingreso familiar

En resumen, el anterior es un caso con una correlación nula ($r = 0$) entre las variables independientes. Por consecuencia, la varianza explicada entre las variables independientes es cero también y la varianza que cada una de estas explica en la dependiente, es única. Por otro lado, las correlaciones entre las variables independientes son comunes en la investigación educativa (*i.e.*, esta correlación entre las variables independientes se llama *multicolinealidad: multicollinearity*). Esto es un problema cuando se usa la regresión, la cual implica la r de Pearson. Sin embargo, está más allá de los propósitos de este libro explicar la *multicolinealidad*, porque aquí no se toca el tema de la regresión múltiple a profundidad, pero Tabachnick y Fidell (2012) hicieron una buena discusión al respecto. En la Figura 3.7, se muestra un ejemplo gráfico con la r^2 de la multicolinealidad, en donde la variable dependiente está representada por el cuadrado con bordes negros en el centro con una multitud de variables independientes representadas por cuadrados con bordes más delgados. Estas variables independientes,

además de explicar el 100% de la varianza de la variable dependiente, se correlacionan entre sí ocasionando una situación de multicolinealidad.

Figura 3.7 Multicolinealidad con r^2



Cuadrados con bordes negros más delgados = Variables independientes;
Cuadrado con borde negro más grueso = Variable dependiente.

e. Requisitos para usar el Coeficiente r de Pearson

DOS CONDICIONES SE HABÍAN DISCUTIDO COMO REQUISITOS PARA utilizar el Coeficiente r : el empleo de una escala continua para el r y la linealidad de los datos en un plano cartesiano. En corto, para usar el r es necesario tener una escala continua, pero, de hecho, se emplean también las escalas tipo Likert, lo cual no es lo óptimo para dicho coeficiente. Según la literatura, sería mejor utilizar las correlaciones *policóricas* con las escalas tipo Likert. La otra condición para usar el r es la linealidad de los datos. En otras palabras, esto quiere decir que los datos forman un patrón en forma de línea en un plano, ya sea de manera positiva o negativa. Para una investigación *descriptiva*, estos dos requisitos serían probablemente suficientes. Por ejemplo, una investigadora o un investigador solo desean mostrar una correlación con el r de Pearson

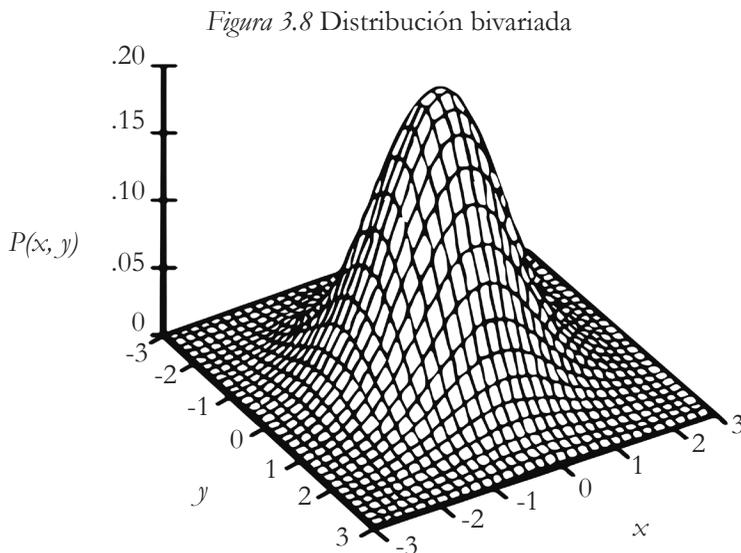
entre exámenes de matemáticas y ciencias, pero sin estimar un intervalo de confianza o someter dicho coeficiente a un *Test de Significancia Estadística*. En este caso, de intervalo de confianza o significancia estadística, Cohen (1988) recomendó que sí es necesario cumplir con los requisitos de normalidad y *homocedasticidad*.

Normalidad para una correlación

UNA DISTRIBUCIÓN NORMAL TIENE UNA SOLA VARIABLE QUE APARECE al azar. *Repasando a nivel de una sola variable*: la distribución normal de los datos, se puede inferir cuando las medidas de tendencia central son iguales o similares (*i.e.*, \bar{x} = Mediana = Moda) cuando sus coeficientes de curtosis y simetría se aproximan a cero. Asimismo, el ya mencionado Test de Kolmogorov-Smirnov es otra opción para evaluar la distribución normal a nivel de una variable a la vez. Como una extensión de esta distribución está la distribución normal bivariada, que fue explicada por Salkind (2007):

La distribución normal bi-variada surge cuando un par de variables (x , y) no solo son normales individualmente, sino que también lo son en forma conjunta. La distribución normal bi-variada puede ser visualizada como una campana tridimensional. Esta distribución puede ser extendida tres o más variables (p. 673).

Una representación gráfica de esta distribución normal bivariada y estandarizada, se muestra en la Figura 3.8. En esta, se representan las dos variables que aparecen al azar (x , y), las cuales tienen una correlación y son normales en forma univariada y bivariada. Cada una de ellas tiene un eje y el tercero es la probabilidad de ambas variables: $P(x, y)$.



Para calcular la probabilidad de esta distribución normal bivariada y estandarizada, se emplea la Ecuación 3.6 (Weisstein, 2020, párr. 1). Para una discusión a fondo acerca de esta ecuación y esta distribución, se recomienda consultar a: Kotz (2006) y Salkind (2007), así como la página de donde fue tomada la Ecuación 3.6: Wolfram MathWorld (véase: Recursos de internet).

$$P(x, y) = \left[\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \right] \left[e^{-\frac{z}{2(1-\rho^2)}} \right] \quad (\text{Ecuación 3.6})$$

donde:

- $P(x, y)$ = Probabilidad de ambas variables.

Las definiciones de sus elementos son:

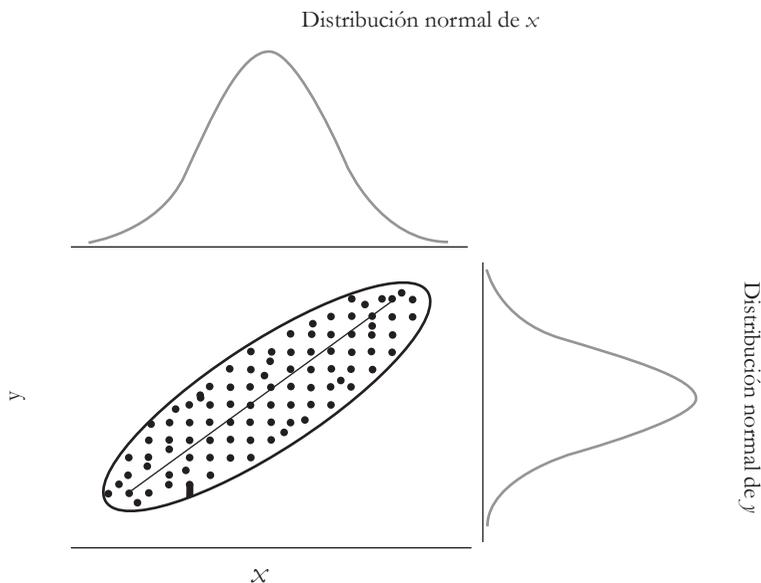
$$z = (x - \mu_x)^2 / \sigma_x^2 - [2\rho (x - \mu_x) (y - \mu_y) / \sigma_x \sigma_y] + (y - \mu_y)^2 / \sigma_y^2$$

donde:

- z = Valor estandarizado de ambas variables;
- μ_x = Promedio de la población de la variable x ;
- μ_y = Promedio de la población de la variable y ;
- σ_x^2 = Varianza de la población de la variable x ;
- σ_y^2 = Varianza de la población de la variable y ;
- ρ = Correlación (x,y) = Covarianza de (x,y) / $\sigma_x \sigma_y$;
- σ_x = Desviación estándar de la variable x ;
- σ_y = Desviación estándar de la variable y ;
- $\pi = 3.1416\dots$
- $e = 2.7183\dots$

Otra manera en la que se manifiesta la normalidad de la distribución bivariada es mediante una elipse, que evidencia que tanto x como y están distribuidas normalmente en forma univariada (Figura 3.9).

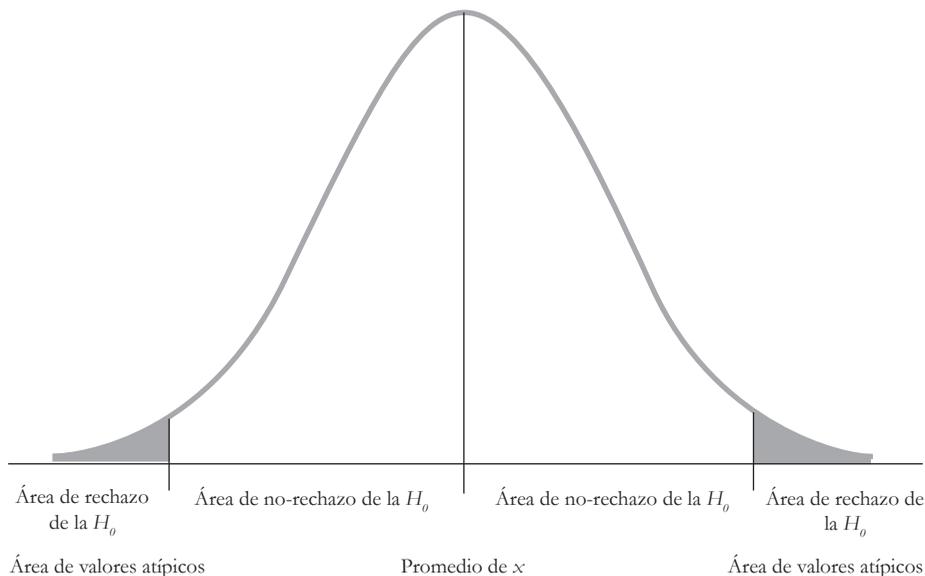
Figura 3.9 Distribución normal bivariada

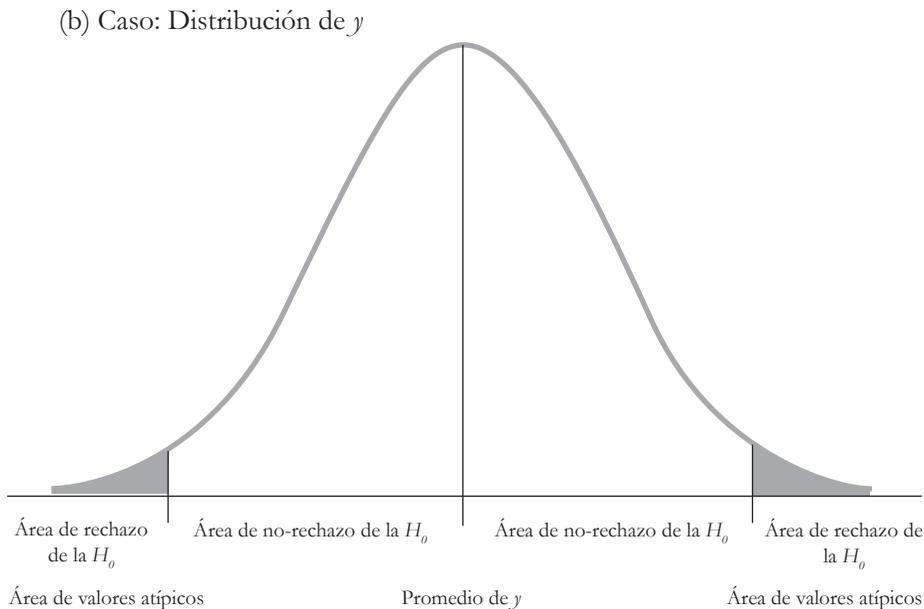


En resumen, para cumplir con parte de los requisitos de Cohen (1988) ambas variables deben estar distribuidas normalmente (casos A y B), sin observaciones atípicas. También, esta normalidad implica la carencia de observaciones atípicas (*i.e.*, datos a nivel *univariable* que estén a $|2|$ o $|3|$ desviaciones estándar del \bar{x}). Cuando existen observaciones atípicas, estas disminuyen el r al cambiar la distancia entre los puntos y la línea de mejor ajuste. La distribución normal también se extiende a la distribución de ambas variables. Sin embargo, está más allá de los propósitos de este libro la estimación de coeficientes para observar si la evidencia sustenta la normalidad bivariada. Pero si se desea ahondar en el tema, se recomienda consultar a Cain, Zhang y Yuan (2017) con su artículo acerca de la curtosis y la asimetría a nivel multivariado.

Figura 3.10 Normalidad con áreas de rechazo de la hipótesis nula y valores atípicos

(a) Caso: Distribución de x





Homocedasticidad

EN LA FIGURA 3.11 (CASO A), SE MUESTRA LA SITUACIÓN DE *HOMOCE-dasticidad* en una correlación que implica no tener observaciones atípicas, así que si los datos lucen así en una investigación cabe la posibilidad de que se cumpla con este supuesto. Al respecto, Salkind (2007, p. 215) definió la homocedasticidad, *i.e.*, homogeneidad de la varianza, como la constancia del error de la varianza a través de todos los valores de x . *Recordando*: la distancia entre la *línea de mejor ajuste* y cada punto de la correlación es un error. Más precisamente, las zonas sombreadas en ambas distribuciones normales (casos A y B de la Figura 3.10) indican con valores críticos dónde estarían los valores atípicos. Un valor atípico cambiaría la distancia entre los puntos y la línea de mejor ajuste (véase: efecto de un valor atípico en la línea de mejor ajuste: caso b; Figura 3.11).

También, la homocedasticidad se refiere a la distancia uniforme a la que están los diferentes puntos en el plano cartesiano respecto de la *línea de mejor ajuste* (caso a;

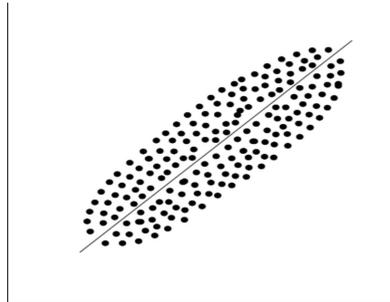
Figura 3.11). En contraparte, la *heterocedasticidad* expresa que la distancia entre la línea de mejor ajuste y los puntos no es uniforme (caso c; Figura 3.11). En este caso de heterocedasticidad, las distancias entre los puntos y la línea de mejor ajuste en un plano cartesiano empiezan siendo uniformes, pero como se incrementan tanto los valores de x como los de y , la distancia entre estos dos se va haciendo más grande. Esta heterocedasticidad sería un impedimento para efectuar una correlación entre estas variables. Habría que transformar los datos de estas variables en caso de desear usar un test de significancia o estimar intervalos de confianza (para las transformaciones, véase a: Tabachnick, & Fidell, 2012). En síntesis, en la Figura 3.11 se muestran ejemplos de:

- Distribución de datos en el plano cartesiano con la línea de mejor ajuste, sin observaciones atípicas y de homocedasticidad (caso a);
- Efecto de un valor atípico en la línea de mejor ajuste (caso b);
- Heterocedasticidad (caso c).

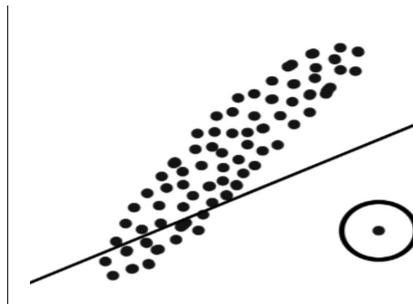
No hay una distancia uniforme (forma de elipse) entre los puntos y la línea de mejor ajuste.

Figura 3.11 Homocedasticidad y heterocedasticidad

(a) Caso: sin observaciones atípicas y de homocedasticidad

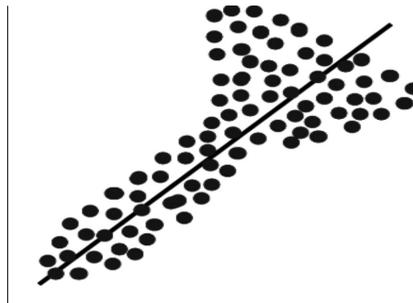


(b) Con una observación atípica
(el círculo indica la observación atípica)



Con efecto en la línea de mejor ajuste

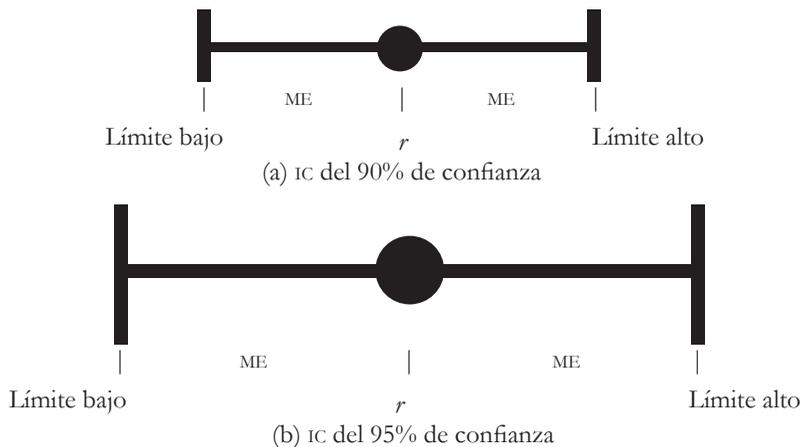
(c) Caso: heterocedasticidad

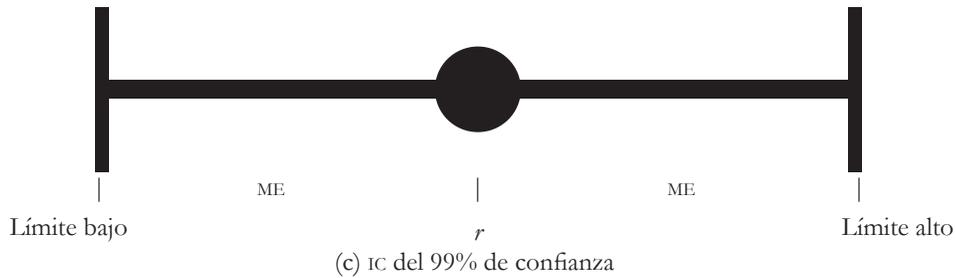


f. Intervalo de confianza para la r de Pearson

COMO LA R DE PEARSON ES OBTENIDA DE UNA MUESTRA AL AZAR para poder generalizar al ρ_{pp} de una población, es recomendable estimar un intervalo de confianza (IC), porque con ellos se calcula un margen de error (ME; *margin of error*; *i.e.*, un grado de imprecisión de una estadística para ilustrar cuál sería un parámetro de interés de una población). Más de esto: los IC permiten estimar el valor verdadero de un parámetro de una población (*e.g.*, promedio, desviación estándar y el coeficiente ρ_{pp} , entre otros). *Recordando*: un IC está constituido por una estadística (*e.g.*, r), que está usualmente en el centro del mismo, y dos brazos que representan el margen de error hacia cada lado del centro (Figura 3.12).

Figura 3.12 Diferentes niveles de confianza en un IC





El ME señala la extensión del IC en donde probablemente se encuentra el valor verdadero de la población. El ME está constituido por dos elementos:

Margen de error = Nivel de confianza deseado x Error estándar de una estadística.

Nivel de confianza deseado: $z = 1.64$ para un nivel de confianza del 90%; $z = 1.96$ para un nivel de confianza del 95%; y $z = 2.58$ para un nivel de confianza del 99%.

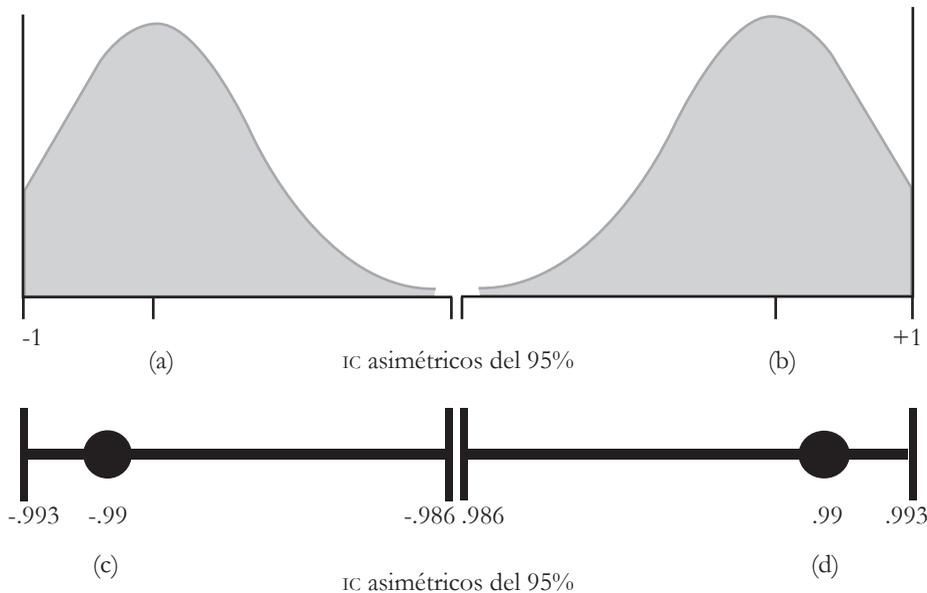
Error estándar de una estadística: $SE_{zr} = \frac{1}{\sqrt{n-3}}$ (Ecuación 3.7)

El nivel de confianza deseado se calcula en valores z (más adelante se ilustran estos procedimientos). Esta simple fórmula muestra que el tamaño del margen de error depende del nivel de confianza y del error estándar de una estadística. En pocas palabras, entre más alto sea el nivel de confianza deseado mayor será el margen de error, *ceteris paribus*. Asimismo, entre más grande sea el error estándar de una estadística (Ecuación 3.7) mayor será el margen de error, *ceteris paribus*. De nuevo: la fórmula del error estándar de una estadística: $SE_{zr} = 1/\sqrt{n-3}$, indica que entre más grande sea la muestra menor será el error estándar, lo que se debe a que el tamaño de la muestra se encuentra en el denominador de la fórmula, que tiene una relación inversa con el cociente de esta (SE_{zr}). De una muestra no se puede obtener el 100% de confianza,

porque existe un error cuando no se tiene toda la población. Este error es igual a la diferencia entre los coeficientes de $\rho_{pp} - r$.

Cumming (2013) definió el margen de error como “La longitud de un brazo de un IC” (p. 442). Esto implica que un IC tiene dos brazos con dos límites: el bajo y el alto, que señalan tanto donde probablemente se encuentra el valor verdadero de la población como donde no es probable que exista dicho valor verdadero: más allá de los límites bajo y alto. Otras cosas siendo iguales: entre más *extenso* sea el IC calculado de una muestra, habrá más posibilidades de capturar el ρ_{pp} .

El valor de una correlación en una población es fijo, *i.e.*, ρ_{pp} = Cierta coeficiente que no cambia, *ceteris paribus*. Difícilmente, se podría saber el valor de dicho coeficiente, porque no se tienen, por lo general, los datos de toda la población. Cuando existe un censo, se supone que, si se tiene toda la población, entonces sí se puede calcular un parámetro sin necesidad de utilizar un IC. Por otro lado, se podrían sacar varias muestras de una población y, de estas, estimar una serie de coeficientes r con los que se puede crear una distribución de muestras (*i.e.*, *sampling distribution of r*), para estimar cuál es el τ_{pp} de su población. En contraparte, esta distribución de muestras no sería normal (simétrica) cuando el valor del r de las muestras se aproxima a -1 o $+1$ (*i.e.*, al valor absoluto de $1: |1|$), porque estos números son los límites del Coeficiente r . Al tener estos límites el r , se sesga la distribución de muestras, porque no se puede ir más allá de -1 o $+1$ (Figura 3.13). Porque la distribución de muestras del r , es asimétrica cuando las muestras se acercan al valor $|1|$, también los IC del r son asimétricos, *i.e.*, las colas ya no son de la misma extensión como cuando hay una distribución normal (Figura 3.13).

Figura 3.13 Distribuciones de muestras asimétricas del r de Pearson

Para mostrar cómo se vuelve asimétrico/sesgado un IC de un Coeficiente r (Figura 3.13), se utilizó la calculadora para obtener los límites del IC (vassarstats.net/rho.html) con $n = 100$ y $r = .99$. Esta asimetría se puede calcular de la siguiente manera:

- $.99 (r) - .986$ (límite bajo del r) = $.004$;
- $.993$ (límite alto del r) - $.99 (r)$ = $.003$;

entonces, $.004 \neq .003$, y no es un IC simétrico.

Más al respecto, se tendría también un IC asimétrico cuando se usara el $r = -.99$ (Figura 3.13).

Siguiendo con el empleo de esta calculadora, se estimaron dos IC del 95% y 99% para cinco tamaños de muestras y los tres tamaños de efecto que clasificó Cohen (1988): pequeño, mediano y grande, los cuales son una recomendación cuando

no existen o son escasas las referencias acerca de algún campo del conocimiento en las ciencias sociales (Tabla 3.6). Como se puede apreciar, los IC son más grandes para tener un 99% de confianza en lugar de un 95%.

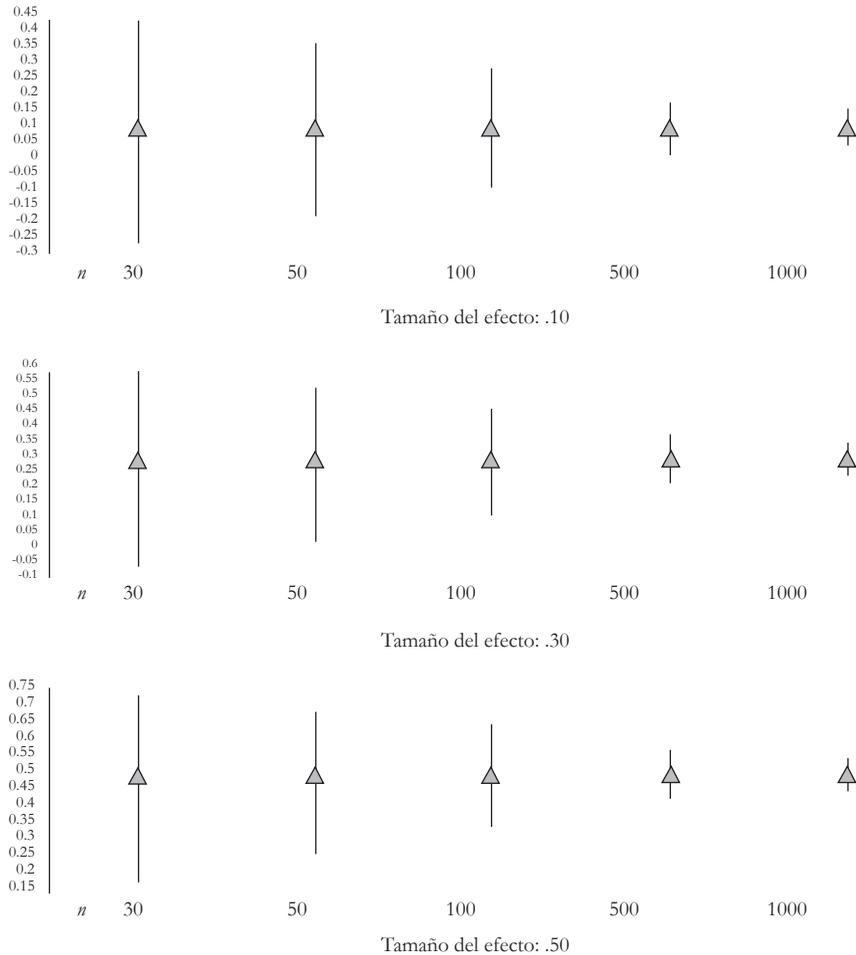
Tabla 3.6 Tamaños del r y sus IC del 95% y 99%

Tamaño/ n	30	50	100	500	1000
.10 Pequeño	[-.269, .444]	[-.183, .368]	[-.098, .29]	[.013, .186]	[.039, .160]
	[-.375, .534]*	[-.268, .443]*	[-.159, .346]*	[-.015, .212]*	[.019, .179]*
.30 Mediano	[-.067, .595]	[.024, .533]	[.111, .468]	[.219, .377]	[.243, .355]
	[-.184, .666]*	[-.066, .594]*	[.048, .516]*	[.192, .401]*	[.225, .372]*
.50 Grande	[.171, .728]	[.258, .683]	[.337, .634]	[.432, .563]	[.453, .545]
	[.054, .779]*	[.175, .728]*	[.281, .670]*	[.409, .581]*	[.437, .558]*

Nota: * = IC del 99%.

En la Figura 3.14, se muestra cómo los IC del 95% de los tres tamaños de efecto propuestos por Cohen (1988), se hacen más pequeños conforme el tamaño de las muestras se hace más grande (de 30 a 1000). En otras palabras, se vuelven más precisos con el aumento del tamaño de las muestras.

Figura 3.14 IC de los tres tamaños de efecto propuestos por Cohen (1988)



También, estos IC se vuelven menos asimétricos conforme el tamaño de la muestra se incrementa desde 30 hasta 1000 (Tabla 3.7). Esta asimetría se calculó igualmente que la anterior:

- $r - \text{límite bajo del } r \text{ (LB)} = a$;
- $\text{Límite alto del } r \text{ (LA)} - r = b$;

- Si son simétricos, $a = b$, o $a - b = 0$;
- Si son asimétricos, $a \neq b$, o $a - b \neq 0$.

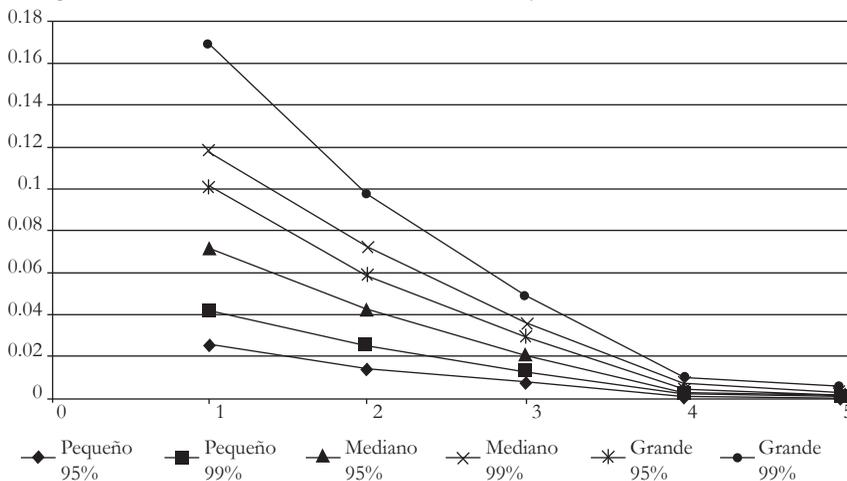
Lo que se encontró es que todos los IC fueron asimétricos.

Tabla 3.7 IC del 95% y 99%

Tamaño	$n = 30$			$n = 50$			$n = 100$			$n = 500$			$n = 1000$		
	$r - LB = a$	$LA - r = b$	$\frac{(r - LB) - (LA - r)}{a - b} =$	$r - LB = a$	$LA - r = b$	$\frac{(r - LB) - (LA - r)}{a - b} =$	$r - LB = a$	$LA - r = b$	$\frac{(r - LB) - (LA - r)}{a - b} =$	$r - LB = a$	$LA - r = b$	$\frac{(r - LB) - (LA - r)}{a - b} =$	$r - LB = a$	$LA - r = b$	$\frac{(r - LB) - (LA - r)}{a - b} =$
.10	.369	.344	.025	.283	.268	.015	.198	.19	.008	.087	.086	.001	.061	.06	.001
.10*	.475	.434	.041	.368	.343	.025	.259	.246	.013	.115	.112	.003	.081	.079	.002
.30	.367	.295	.072	.276	.233	.043	.189	.168	.021	.081	.077	.004	.057	.055	.002
.30*	.484	.366	.118	.366	.294	.072	.252	.216	.036	.108	.101	.007	.075	.072	.003
.50	.329	.228	.101	.242	.183	.059	.163	.134	.029	.068	.063	.005	.047	.045	.002
.50*	.446	.279	.167	.325	.228	.097	.219	.17	.049	.091	.081	.01	.063	.058	.005

En la Figura 3.13, se muestra la diferencia entre los límites del IC con el r : $(r - LB) - (LA - r) = a - b$. Los IC, tanto del 95% como del 99%, comienzan siendo más asimétricos para el efecto grande (*i.e.*, .50). Como se puede apreciar, conforme el tamaño de la muestra aumenta (desde $n = 30$ hasta $n = 1000$), los IC se vuelven menos asimétricos. De hecho, esta asimetría converge hacia el *cero* cuando el tamaño de la muestra es 1000, sin importar el tamaño inicial del efecto: .10, .30 o .50 (Figura 3.15).

Figura 3.15 Relación entre la asimetría del IC y el tamaño de la muestra



Nota: 1 = ($n = 30$); 2 = ($n = 50$); 3 = ($n = 100$); y 4 = ($n = 500$).

La recomendación para tener un IC menos asimétrico, es tratar de obtener la muestra más grande que se pueda para poder capturar los parámetros de una población (*ceteris paribus*).

Conversión de coeficientes r a valores z

ADEMÁS DE UTILIZAR TABLAS Y CALCULADORAS PARA CONVERTIR un Coeficiente r en un valor z (e.g., véase Tabla 3.8), se puede emplear la transformación de Fisher para calcular los IC para el r . En la Tabla 3.8, se muestra en la columna de la izquierda el valor absoluto del r y su correspondiente valor absoluto de z .

Por ejemplo, el valor de $|r| = .01$ corresponde al de $|z| = 0.01$ en el área sombreada. Del mismo modo, se pueden encontrar el resto de las transformaciones del r en z . Por esta razón de los sesgos, Sir Ronald Aylmer Fisher desarrolló una transformación de la r de Pearson a la z de Fisher, para poder usar una distribución de muestras normal: $z_{transformado} = .5 [In(1+r) - In(1-r)]$. Al igual que con los IC, en esta

transformación se utiliza un Test z para algunos Test de Significancia Estadística del r . Para otros test del Coeficiente r , el Test t se emplea también y ambos son ilustrados más adelante.

Tabla 3.8 Transformación de Fisher de la r a la z

$ r $	$ z $	$ r $	$ z $	$ r $	$ z $
.01	0.01	.34	0.35	.67	0.81
.02	0.02	.35	0.37	.68	0.83
.03	0.03	.36	0.38	.69	0.85
.04	0.04	.37	0.39	.70	0.87
.05	0.05	.38	0.40	.71	0.89
.06	0.06	.39	0.41	.72	0.91
.07	0.07	.40	0.42	.73	0.93
.08	0.08	.41	0.44	.74	0.95
.09	0.09	.42	0.45	.75	0.97
.10	0.10	.43	0.46	.76	1.00
.11	0.11	.44	0.47	.77	1.02
.12	0.12	.45	0.48	.78	1.05
.13	0.13	.46	0.50	.79	1.07
.14	0.14	.47	0.51	.80	1.10
.15	0.15	.48	0.52	.81	1.13
.16	0.16	.49	0.54	.82	1.16
.17	0.17	.50	0.55	.83	1.19
.18	0.18	.51	0.56	.84	1.22
.19	0.19	.52	0.58	.85	1.26
.20	0.20	.53	0.59	.86	1.29
.21	0.21	.54	0.60	.87	1.33
.22	0.22	.55	0.62	.88	1.38
.23	0.23	.56	0.63	.89	1.42
.24	0.24	.57	0.65	.90	1.47
.25	0.26	.58	0.66	.91	1.53

Continúa...

$ r $	$ z $	$ r $	$ z $	$ r $	$ z $
.26	0.27	.59	0.68	.92	1.59
.27	0.28	.60	0.69	.93	1.66
.28	0.29	.61	0.71	.94	1.74
.29	0.30	.62	0.73	.95	1.83
.30	0.31	.63	0.74	.96	1.95
.31	0.32	.64	0.76	.97	2.09
.32	0.33	.65	0.78	.98	2.30
.33	0.34	.66	0.79	.99	2.65

Cuando $r = 0$, también $z = 0$. Cuando $r = 1$ o -1 , no se puede calcular el valor z , porque la fórmula no funciona para este tipo de operación, ya que el valor z se va al infinito. En la Tabla 3.8, se muestran los valores absolutos, así que solo hay que agregar el signo negativo a ambos valores para cubrir el lado negativo de la distribución de muestras z . También, los valores de ambas columnas fueron redondeados a dos dígitos, así que con esto se pierde precisión. Por lo tanto, estas transformaciones de los valores r a z (para esta tabla en particular) y viceversa, se deben tomar como aproximaciones solamente.

Pasos para calcular un IC de un Coeficiente r

LOS PASOS PARA CALCULAR UN IC DE UN COEFICIENTE R , ES TRANSFORMAR el $r_{\text{calculado}}$ a un valor z , usando la ya antes mencionada fórmula:

$$z_{\text{transformado}} = .5 [\ln(1 + r) - \ln(1 - r)]. \quad (\text{Ecuación 3.8})$$

El segundo paso es establecer un nivel de confianza, *e.g.*, 90%, 95% o 99%. El nivel de confianza más común es el 95%. El tercer paso es buscar en una Tabla de valores z el valor $z_{\text{crítico}}$, que corresponde al nivel de confianza, pero con la explicación de que los IC tienen dos brazos, así que el nivel de confianza se tiene que dividir entre

dos para crear un IC *simétrico* para el valor z . El cuarto y último paso es calcular los dos brazos del IC.

Ejemplo de los pasos para calcular un IC de un r

DE NUEVO, EL PRIMER PASO (TRANSFORMACIÓN DEL R EN EL Z: Tomando uno de los ejemplos anteriores, para usar el procedimiento de la fórmula de transformación de Fisher). Se tiene un $r = .98$ y $n = 25$:

Sustituyendo el valor de .98 en la Ecuación 3.8:

$$z_{\text{transformado}} = .5 [In(1 + .98) - In(1 - .98)]$$

$$z_{\text{transformado}} = .5 [In(1.98) - In(.02)]$$

$$z_{\text{transformado}} = .5 [.683 - (-3.912)]$$

$$z_{\text{transformado}} = .5 [4.595]$$

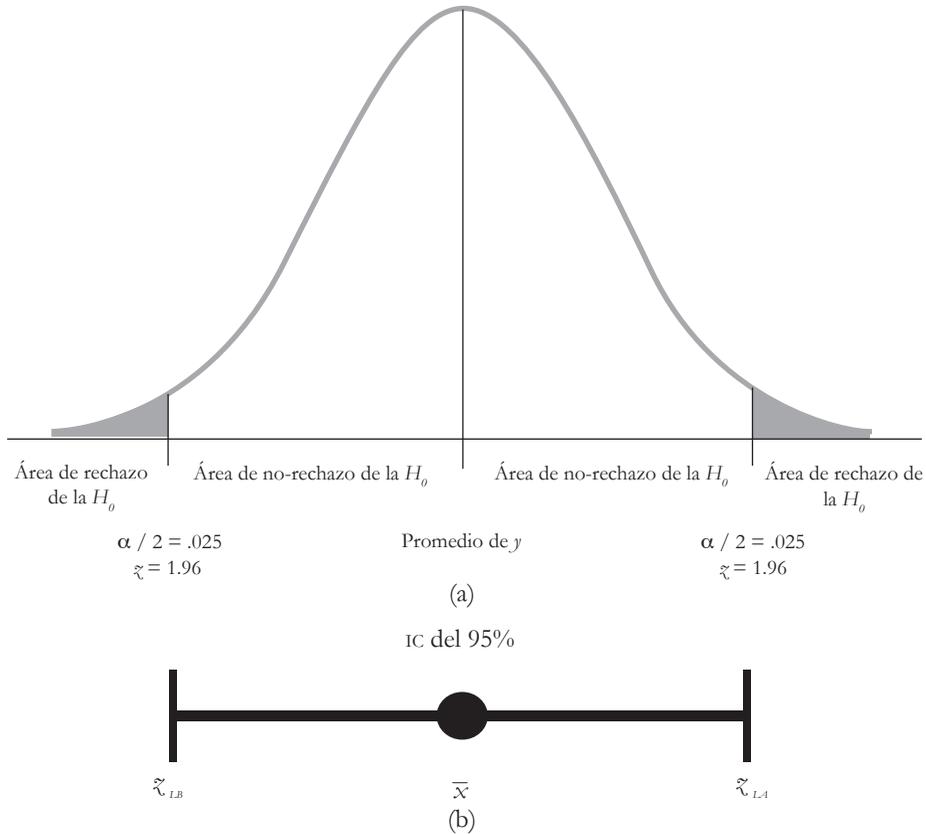
$$z_{\text{transformado}} = 2.297.$$

Para verificar el resultado anterior, se puede utilizar la Tabla 3.8, en la cual el r de .98 es aproximadamente un $z_{\text{transformado}} = 2.30$.

El segundo paso es establecer el nivel de confianza: para este ejemplo, se elige el 95% de confianza, porque es el tradicional en algunas ciencias sociales.

El tercer paso es una búsqueda en una Tabla de valores z del valor correspondiente al nivel de confianza: para el 95% de confianza, le corresponde un alfa de .05 (*i.e.*, este alfa de .05 cubre un área del 5% de una curva normal). Ahora, se divide entre dos: $.05 / 2 = .025$, porque ambos brazos tienen la misma extensión, *i.e.*, es un IC simétrico de un valor z . Asimismo, a este $\alpha / 2 = .025$ le corresponde un valor z de 1.96. Más al respecto, el área que no cubre este nivel de confianza es del 5% (100% - 95%). A esta área que no converge con el IC, le corresponde el área de rechazo de la distribución de muestras z (Figura 3.16).

Figura 3.16 Distribución de muestras e IC del 95%



Nota: z_{LB} = Límite bajo del IC; \bar{x} = Promedio de la muestra; y z_{LA} = Límite alto del IC (estos límites serán usados enseguida).

El cuarto paso es calcular el IC con los dos brazos, que son el límite bajo (LB; *lower limit*) y el límite alto (LA; *upper limit*). El límite bajo del IC, es el valor más pequeño que probablemente tendrá un parámetro de la población. Similarmente, el límite alto es posiblemente el valor más grande que tendrá un parámetro de la población. Los valores más allá de las colas, se consideran poco probables ($\leq .05$). Para calcular los IC, se utilizan las siguientes dos fórmulas (ecuaciones 3.9 y 3.10):

$$\bar{z}_{LB} = \bar{z}_{transformado} - z_{crítico / 100} \times \frac{1}{\sqrt{n-3}}; \quad (\text{Ecuación 3.9})$$

$$\bar{z}_{LA} = \bar{z}_{transformado} + z_{crítico / 100} \times \frac{1}{\sqrt{n-3}}; \quad (\text{Ecuación 3.10})$$

donde:

- \bar{z}_{LB} = Valor estandarizado del límite bajo;
- \bar{z}_{LA} = Valor estandarizado del límite alto;
- $\bar{z}_{transformado}$ = Valor estandarizado de la transformación del r de Pearson, *e.g.*, usando la Tabla 3.8;
- $z_{crítico / 100}$ = Valor estandarizado crítico que se eligió, *e.g.*, $z = 1.64$ para un nivel de confianza del 90%; $z = 1.96$ para un nivel de confianza del 95%; y $z = 2.58$ para un nivel de confianza del 99%;
- n = Tamaño de la muestra.

Usando los datos del ejemplo anterior ($r = .98$, se convirtió en el $\bar{z}_{transformado} = 2.296$):

$$\begin{aligned} \text{Límite bajo: } \bar{z} &= 2.297 - (1.96 \times 1 / \sqrt{25-3}) \\ &= 2.297 - (1.96 \times 1 / \sqrt{22}) \\ &= 2.297 - (1.96 \times 1 / 4.69) \\ &= 2.297 - (1.96 \times .21) \\ &= 2.297 - .4116 \end{aligned}$$

Límite bajo: $\bar{z} = 1.89$ (redondeando).

Nota: para el límite alto son las mismas operaciones del límite bajo (pero se suma en lugar de restar).

$$\text{Límite alto: } \bar{z} = 2.297 + .4116$$

Límite alto: $\bar{z} = 2.71$ (redondeando).

En resumen, con los valores z_{α} el IC del 95% del $r_{\text{calculado}}$ es igual al expresado en valores z_{α} [1.89, 2.71]. *Recordando*: los corchetes se emplean para indicar un IC. También, este IC se puede calcular en valores r al revertir el procedimiento anterior: de valores z_{α} a valores r . Las siguientes fórmulas muestran la forma para transformar valores z_{α} en valores r (ecuaciones 3.11 y 3.12): en estas, se utiliza el número e (2.71828...), que es una constante usada en matemáticas. Por el momento, e es redondeado a 2.72 para simplificar el ejemplo y ser elevado a cierta potencia:

$$r_{LB} = (e^{2 z_{LB}} - 1) / (e^{2 z_{LB}} + 1); \quad \text{(Ecuación 3.11)}$$

$$r_{LA} = (e^{2 z_{LA}} - 1) / (e^{2 z_{LA}} + 1); \quad \text{(Ecuación 3.12)}$$

donde:

- r_{LB} = Valor r del límite bajo;
- r_{LA} = Valor r del límite alto;
- $e = 2.72$ (redondeando);
- z_{LB} = Valor z_{α} del límite bajo;
- z_{LA} = Valor z_{α} del límite alto.

Usando los valores del ejemplo anterior para el límite bajo ($z_{LB} = 1.89$):

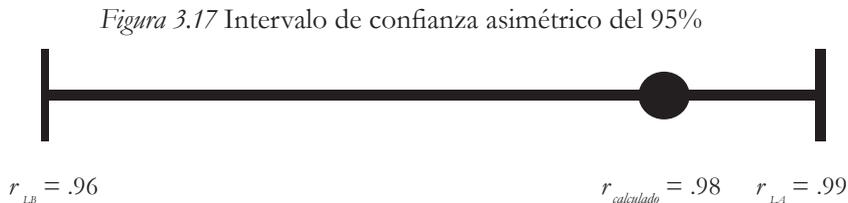
$$\begin{aligned} r_{LB} &= (2.72^{2 \times 1.89} - 1) / (2.72^{2 \times 1.89} + 1) \\ &= (2.72^{3.78} - 1) / (2.72^{3.78} + 1) \\ &= 43.92 - 1 / 43.92 + 1 \\ &= 42.92 / 44.92 \\ r_{LB} &= .96. \end{aligned}$$

Utilizando los valores del ejemplo anterior para el límite alto ($z_{LA} = 2.71$):

$$\begin{aligned} r_{LA} &= (2.72^{2 \times 2.71} - 1) / (2.72^{2 \times 2.71} + 1) \\ &= (2.72^{5.42} - 1) / (2.72^{5.42} + 1) \end{aligned}$$

$$\begin{aligned}
 &= 226.65 - 1 / 226.65 + 1 \\
 &= 225.65 / 227.65 \\
 r_{L,A} &= .99.
 \end{aligned}$$

En concreto, el IC del 95% del $\tilde{r}_{transformado}$ es igual a: [.96, .99]. Dado que el $r_{calculado}$ es .98, su IC es asimétrico (Figura 3.17), porque la diferencia entre el límite bajo y el promedio es mayor que la diferencia entre el límite alto y el promedio: $(.98 - .96 = .02) > (.99 - .98 = .01)$. La simetría se debe a que el $r = .98$ está cercano a 1, el cual es el límite.

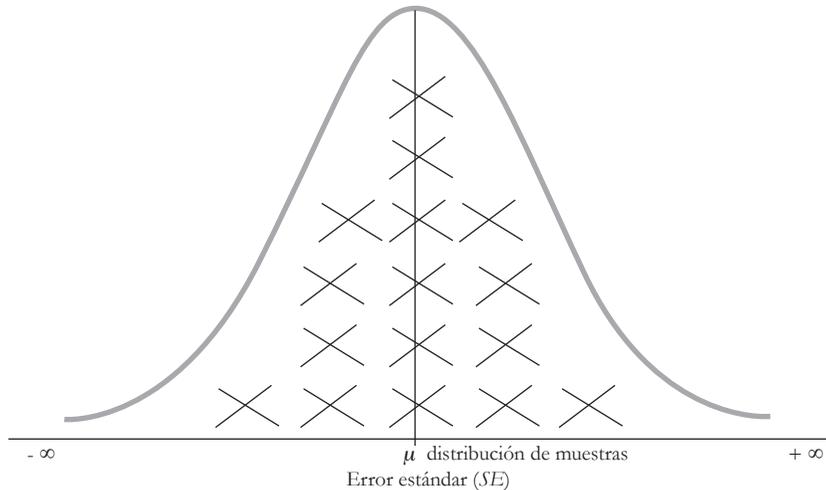


Nota: las distancias de los límites al r calculado son solamente para ilustrar la asimetría del IC.

g. Test de la hipótesis nula

UNO DE LOS SUPUESTOS PARA UTILIZAR ANÁLISIS DE SIGNIFICANCIA, es que la muestra sea representativa para poder hacer la generalización de la población, lo que significa que las características de una o varias muestras son similares a las de la población de donde se extrajeron al *azar*. Se debe enfatizar que debe ser una muestra extraída al azar de una población de interés, de la cual cada individuo, objeto u observación tengan la misma oportunidad de ser seleccionados. Para poner a prueba una hipótesis nula de la r de Pearson, hay que usar una *distribución de muestras* (Figura 3.18), la cual representa a todas las muestras posibles de cierto tamaño de una estadística. Una distribución de muestras tiene una distribución aproximadamente normal, simétrica y asintótica (Ponce, 2019).

Figura 3.18 Distribución de muestras

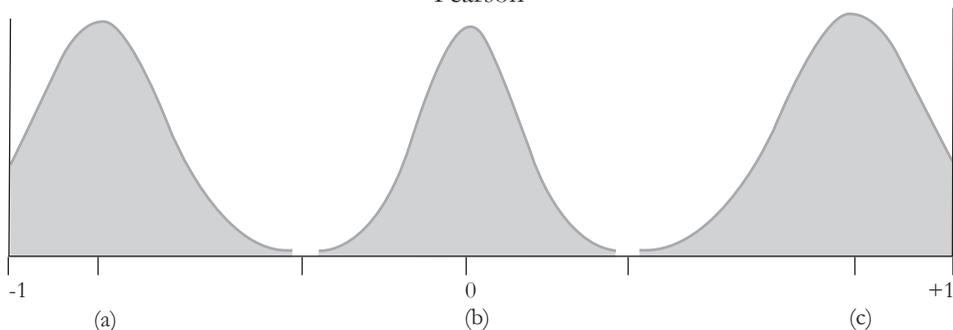


Nota: las cruces representan muestras del mismo tamaño extraídas de una misma población para ser agregadas a una distribución de muestras. El error estándar (SE) es equivalente/análogo a la desviación estándar de una distribución normal.

En particular, para la distribución de muestras del r de Pearson, se hablaría de todas las muestras posibles de cierto tamaño. Sin embargo, no se puede utilizar el Teorema de Tendencia Central (véase: Ponce, 2019) sin considerar, primero, el tamaño y el signo del r , porque la distribución de muestras del r de Pearson no es asintótica (*i.e.*, no se va al infinito sin tocar el eje de las x). De hecho, la r de Pearson tiene dos límites: -1 y $+1$. Por lo tanto, la distribución de muestras de la r de Pearson cambia su forma de acuerdo con el tamaño del coeficiente y su signo. Lo anterior quiere decir que, si el signo del r es negativo y este disminuye su tamaño, la distribución de muestras se sesga hacia el lado negativo, *ceteris paribus* (Figura 3.19). También, cuando el r cambia de signo y aumenta su tamaño, la distribución de muestras se sesga hacia el lado positivo. En resumen, si la estadística del r se aproxima a -1 o a $+1$, la distribución de muestras se sesga, *ceteris paribus*. Otro elemento a considerar para la magnitud de la asimetría en un IC, es el tamaño de la muestra. Como se había mencionado, entre más grande es la muestra menos asimétrico es el IC.

El caso contrario es cuando la distribución de muestras tiende hacia el cero y, entonces, esta tiene una forma aproximadamente normal (Figura 3.19). Sin embargo, se emplean la transformación de Fisher y los Test t y z con sus correspondientes distribuciones de muestras, para poner a prueba las hipótesis nulas del Coeficiente r .

Figura 3.19 Distribución de muestras asimétricas y aproximadamente normal del r de Pearson



Generalidades para poner a prueba una hipótesis nula

EL PONER A PRUEBA UNA HIPÓTESIS NULA, SE PUEDE REALIZAR mediante una serie de pasos. El primero es: *establecer la o las preguntas de investigación*. Posiblemente, la interrogante que podría venir a la mente cuando se habla de correlaciones y del Test de Significancia Estadística sería: ¿existe una relación estadísticamente significativa entre la variable x y la variable y ? Esta pregunta podría ser especialmente apropiada cuando no existen antecedentes en la teoría o la literatura. En contraparte, cuando existen antecedentes (*e.g.*, parámetros hipotéticos de una población), se puede hacer una comparación entre lo que había sido hallado por otros y lo que una investigadora o un investigador han encontrado. Lo anterior se podría expresar así: ¿existe una diferencia estadísticamente significativa entre el Coeficiente de Correlación de la Población y el de la Muestra?

Un segundo paso es: *derivar las hipótesis de la o las preguntas de investigación*. Entonces, se tiene la siguiente interrogante: ¿existe una relación estadísticamente significativa entre la variable x y la variable y ? La derivación de las hipótesis, se puede hacer para una o dos colas.

Dos colas:

- Población:
 - Hipótesis nula (H_0): $\rho = 0$ (*i.e.*, no existe correlación entre dos variables en la población. Este tipo de test suele ser exploratorio, porque no se ha estipulado el signo de la correlación ni la magnitud; sería más apropiado utilizar un test t , porque no se asume una correlación dentro de la población; véase: Hinkle *et al.*, 2003);
 - Hipótesis alternativa (H_a): $\rho \neq 0$ (*i.e.*, existe correlación entre dos variables en la población).
- Muestra:
 - Hipótesis nula (H_0): $r = 0$ (*i.e.*, no existe correlación entre dos variables en la muestra);
 - Hipótesis alternativa (H_a): $r \neq 0$ (*i.e.*, existe correlación entre dos variables en la muestra).

Una cola:

- Población:
 - Hipótesis nula (H_0): $\rho = 0$;
 - Hipótesis alternativa (H_a): $\rho < 0$ (la correlación en la población es menor que cero y, por lo tanto, es negativa); o
 - Hipótesis alternativa (H_a): $\rho > 0$ (la correlación en la población es mayor que cero y, por lo tanto, es positiva).
- Muestra:

- Hipótesis nula (H_0): $r = 0$;
- Hipótesis alternativa (H_A): $r < 0$ (la correlación en la muestra es menor que cero y, por lo tanto, es negativa); o
- Hipótesis alternativa (H_A): $r > 0$ (la correlación en la muestra es mayor que cero y, por lo tanto, es positiva).

Asimismo, si existe información de la literatura o alguna otra fuente, se puede especificar un valor para este parámetro de la población. En este caso, se puede hacer una pregunta de investigación como la siguiente: ¿existe una diferencia estadísticamente significativa entre el Coeficiente de Correlación de la Población y el de la Muestra?

Dos cosas:

- Población:
 - Hipótesis nula (H_0): $\rho = \text{Valor en específico}$ (*i.e.*, para esta hipótesis, los límites de ρ son de -1 a +1; en este tipo de casos, sería más apropiado el uso de un Test χ^2 ; véase: Hinkle *et al.*, 2003);
 - Hipótesis alternativa (H_A): $\rho \neq \text{Valor en específico}$ (*i.e.*, esta hipótesis estipula que el parámetro de la población no es lo que especifica la hipótesis nula).
- Muestra:
 - Hipótesis nula (H_0): $r = \text{Valor en específico}$ (*i.e.*, la estadística de la muestra corresponde con el parámetro de la población);
 - Hipótesis alternativa (H_A): $r \neq \text{Valor en específico}$ (*i.e.*, la estadística de la muestra no corresponde con el parámetro de la población).

Cuando se trabaja con muestras y sus correlaciones, se pueden cometer los errores tipo I y II (véase: Ponce, 2019). También, se debe tomar en cuenta el poder estadístico (para más información al respecto, véase: Cohen [1988] y Ponce [2019], para una discusión comprensiva al respecto):

- El error tipo I es la probabilidad de rechazar una hipótesis nula cuando es cierta: esto pasaría cuando se encuentra una probabilidad $_{calculada} < \text{alfa}$, con lo que se concluiría que hay una diferencia estadísticamente significativa en la muestra y, por ello, se rechazó la hipótesis nula, a pesar de que no existe correlación entre las variables en la población;
- También, el error tipo II es la probabilidad de no rechazar o de fallar en rechazar la hipótesis nula cuando es falsa. Este error sucede cuando en la muestra, se encontró que la probabilidad $_{calculada} \geq \text{alfa}$, y, entonces, se falla en rechazar la hipótesis nula, a pesar de que en la población existe correlación entre las variables;
- Por último, el poder estadístico es la probabilidad de rechazar una hipótesis nula cuando es falsa. Esta probabilidad del poder estadístico debe ser, por lo menos, del 80%, según Cohen (1988).

La nota anterior acerca de los tipos de errores al usar muestras en un test de correlación, se resume en la Tabla 3.9.

Tabla 3.9 Errores tipo I y II en la correlación de Pearson

Tipos de error	Test χ de Significancia Estadística	Realidad: nivel de población
Tipo I: probabilidad de rechazar una H_0 cierta. Este error se comete cuando se rechaza una hipótesis nula que es cierta.	Este error ocurre cuando se encuentra en una muestra donde la probabilidad $_{calculada} < \text{alfa}$; por lo tanto, se rechaza la hipótesis nula. Entonces, los resultados apoyan la hipótesis alternativa, lo que quiere decir que $r > 0$ o $r < 0$ (para una cola) y $r \neq 0$ (para dos colas).	$r = 0$. Por lo tanto, la hipótesis nula es cierta en la población y, sin embargo, se rechazó, cometiendo el error tipo I.
Tipo II: probabilidad de no rechazar una H_0 falsa. Este error se comete cuando no se rechaza una hipótesis nula que es falsa.	Este error sucede cuando se encuentra que la probabilidad $_{calculada} \geq \text{alfa}$; por lo tanto, no se rechaza la hipótesis nula. Entonces, los resultados apoyan la hipótesis nula, lo cual quiere decir que $r = 0$.	$r \neq 0$. Por lo tanto, la hipótesis nula es falsa en la población y, sin embargo, no se rechazó, cometiendo el error tipo II.

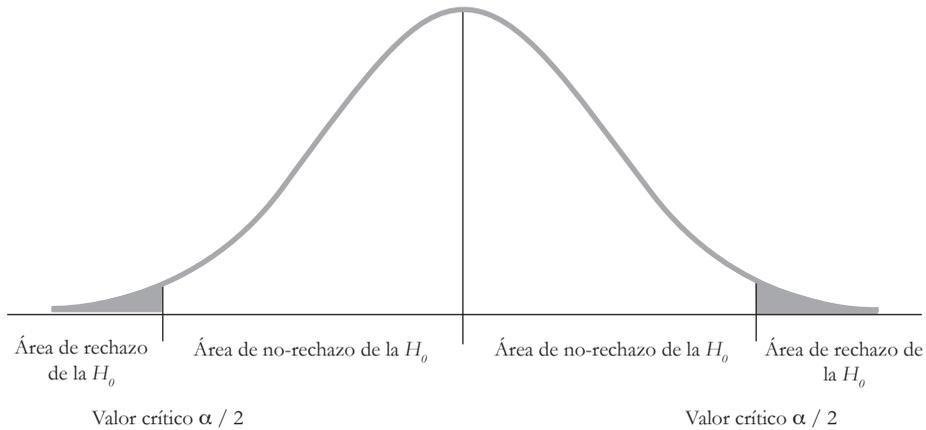
El tercer paso es: *establecer el criterio para rechazar la hipótesis nula (i.e., el nivel del alfa: α)*. Para poner a prueba la hipótesis nula, hay que elegir un nivel del alfa (e.g., $\alpha = .01, .05$ o $.10$). El más común es el $\alpha = .05$ para la investigación educativa. Este alfa o nivel de significancia ayuda a establecer las zonas de rechazo y de no-rechazo de la hipótesis nula (Figura 3.20a: el área sombreada es la zona de rechazo de la hipótesis nula, y la

no-sombreada es la región de no-rechazo o de fallar en rechazar para una y dos colas). Al mismo tiempo que se elige este nivel de significancia, se selecciona si será un test de una o dos colas. Se suele elegir un test de dos colas cuando no se sabe si la correlación (r) va a ser mayor o menor que cero, *e.g.*, en un test exploratorio.

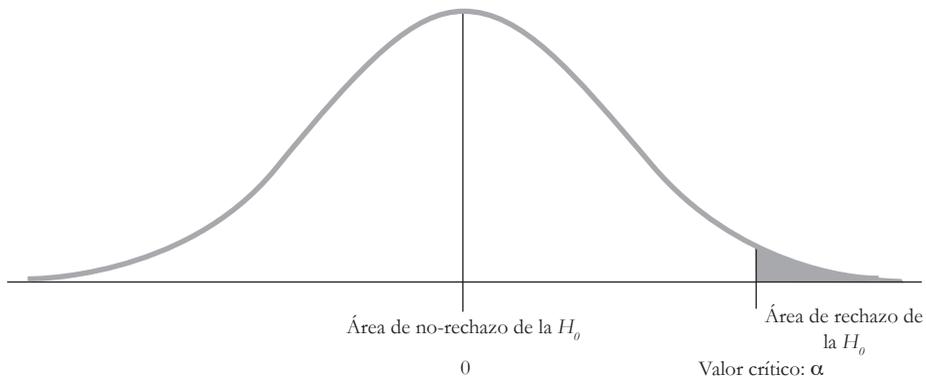
Contrariamente a esto, se elige una cola cuando se anticipa que un valor en específico para el r o una correlación, va a ser positivo ($r > 0$) o negativo ($r < 0$) (figuras 3.20b y 3.20c). Con la elección del alfa de una o dos colas, se puede establecer un valor crítico en una Tabla de valores t o z , o una calculadora en línea. De nuevo, el Test t se suele emplear cuando se especifica el cero como el parámetro de una población, *i.e.*, no existe correlación entre las variables. En este caso, en la Figura 3.20 se representa una serie de distribuciones de muestras con valores t y z que son aproximadamente normales y sirven para poner a prueba el r . Un Test z se usa principalmente para evitar un sesgo en la distribución de muestras cuando el r de la muestra está cercano al valor de -1 o +1. Ambos test son ejemplificados más adelante.

Figura 3.20 Distribución de muestras con áreas de rechazo y no-rechazo de la hipótesis nula para la r de Pearson para una y dos colas para Test t y z

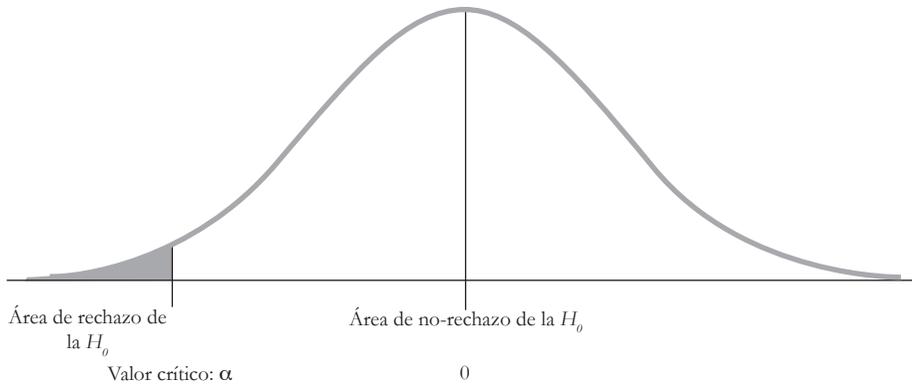
(a) Dos colas



(b) Cola derecha



(c) Cola izquierda



El cuarto paso es: *calcular el valor de la prueba estadística (test statistic)*. Una vez que se ha establecido el criterio, se calcula el r de Pearson de la muestra. Luego, para someterlo a un Test t o z , se cambia: de ser un Coeficiente r pasa a ser un Coeficiente t o z . Para el Test t , se utiliza la siguiente fórmula para convertir un Coeficiente r a un valor t , *i.e.*, $t = r \left[\sqrt{(n-2) / (1 - r^2)} \right]$. Este valor t también es el $t_{\text{calculado}}$, el cual se compara con el $t_{\text{crítico}}$ (*i.e.*, este procedimiento es explicado más adelante en detalle). Un ejemplo heurístico para transformar el r en el $t_{\text{calculado}}$, es el siguiente (Ecuación 3.12):

Se tiene: $r = .05$ y $n = 20$:

$$t_{\text{calculado}} = r \left[\sqrt{\frac{(n-2)}{(1-r^2)}} \right]; \quad (\text{Ecuación 3.12})$$

sustituyendo:

$$t_{\text{calculado}} = .05 \left[\sqrt{(20-2) / (1-.05^2)} \right]$$

$$t_{\text{calculado}} = .05 \left[\sqrt{(18) / (1-.0025)} \right]$$

$$t_{\text{calculado}} = .05 \left[\sqrt{(18) / (.9975)} \right]$$

$$t_{\text{calculado}} = .05 \left[\sqrt{(18.0451)} \right]$$

$$t_{\text{calculado}} = .05 [4.2479]$$

$$t_{\text{calculado}} = 0.2123.$$

También, se puede emplear la calculadora en línea: <https://www.psychometrica.de/correlation.html>. En este caso, esta calculadora comparó el $r = .05$ con $n = 20$ y $df = n - 2$ ($df =$ Grados de libertad: para una discusión más profunda sobre este tema, se recomienda consultar a: Helen [1940]) contra una correlación $= 0$, bajo la categoría *Testing Linear Independence (Testing against 0)*, la cual se podría traducir como: Poner a prueba la independencia lineal (Prueba en contra de 0). En definitiva, el resultado fue: $t_{\text{calculado}} = 0.212$ con una probabilidad $p_{\text{calculada}}$ de .417 (para una cola) y .834 (para dos colas). Con esta información, se puede pasar a los criterios para comparar dicho $t_{\text{calculado}}$ con el $t_{\text{crítico}}$, que son:

- $t_{\text{calculado}} \leq t_{\text{crítico}}$: no rechazar/fallar en rechazar la hipótesis nula. La conclusión es que probablemente no existe correlación en la muestra: $r = 0$; por lo tanto, $\rho = 0$. Lo anterior implica que el $p_{\text{calculado}} \geq \alpha$. En otras palabras, los resultados apoyan la hipótesis nula;
- $t_{\text{calculado}} > t_{\text{crítico}}$: rechazar la hipótesis nula. La conclusión es que probablemente existe correlación en la muestra: $r \neq 0$; por lo tanto, $\rho \neq 0$. Lo anterior implica que la probabilidad $p_{\text{calculada}} < \alpha$. En otras palabras, los resultados apoyan la hipótesis alternativa;
- En este caso, y como se usó la calculadora que no proveyó un $t_{\text{crítico}}$, se compara un alfa tradicional de .05 con dos colas ($.05 / 2 = .025$) con la correspondiente probabilidad calculada de dos colas (.834), *i.e.*, $\alpha .025 < p = .834$, así que no se rechaza la hipótesis nula, porque la probabilidad es más alta que el alfa y se asume que la correlación de $r = .05$ no es diferente de 0.

Para el Test χ^2 , se usa la transformación de Fisher. Más al respecto, APA (2015) explicó que:

Transformación de Fisher de la r a la χ^2 es un procedimiento estadístico para convertir el Coeficiente de Correlación Producto Momento de Pearson a un score estandarizado χ^2 con el propósito de evaluar si la correlación es significativamente diferente de cero. El test es útil en

proveer una estadística/estimación distribución (llamada el valor transformado de Fisher o el ζ de Fisher) que puede ser utilizado en poner a prueba la hipótesis o en crear intervalos de confianza (p. 423).

Antes de entrar en las transformaciones, los logaritmos son explicados brevemente para la transformación del r en ζ con su fórmula. Un logaritmo es un exponente al que se eleva un número, que sirve como base para la obtención de otro número determinado. Un ejemplo es que el logaritmo de 100 con el número 10 como base, es 2 y se escribe: $\text{Log}_{10} 100 = 2$. Para comprobar lo anterior, se realiza lo siguiente: $10^2 = 100$. Un logaritmo responde a la pregunta: ¿a qué potencia se debe elevar la base para obtener el número en cuestión? Otro número que se emplea como base en los logaritmos, es $2.71828183\dots = e$, que es irracional, porque sus decimales se extienden al infinito y también no hay manera de expresarlo con una razón/fracción. Un ejemplo de razón es el número .50, que sí puede ser expresado como una fracción: $1/2 = .50$. Para los ejemplos de esta sección, se redondea el e a 2.72 por simplicidad. Asimismo, cuando se utiliza el número e , se abrevia como \ln en lugar de Log . Por ejemplo, $\ln 10 = 2.30$. Lo anterior, se demuestra como: $2.72^{2.30} = 10$. En palabras, el logaritmo natural de 10 es 2.30, ya que al elevar el número e a la 2.30 potencia, dará como resultado: 10.

Habiendo explicado previamente un logaritmo, la fórmula de Fisher para la transformación del r en ζ , mencionada anteriormente, es la siguiente (Ecuación 3.8):

$$\zeta_{\text{transformado}} = .5 [\ln (1 + r) - \ln (1 - r)]; \quad (\text{Ecuación 3.8})$$

donde:

- \ln = Logaritmo natural cuando se usa de base el número $e = 2.71828183\dots$;
- r = Coeficiente de correlación que se calculó en el análisis y se está transformando;

$$\begin{aligned}
 &= .5 [\ln (1 + .05) - \ln (1 - .05)] \\
 &= .5 [\ln (1.05) - \ln (.95)] \\
 &= .5 [0.0488 - (-0.0513)] \\
 &= .5 [0.1001] \\
 &= 0.05005.
 \end{aligned}$$

Una vez que se ha transformado el r al z_r , es necesario calcular el error estándar de la estadística. En este caso, la estadística es la r de Pearson, pero el error estándar (SE_{z_r}) se obtiene en términos de un valor z_r . En corto, la fórmula es: $SE_{z_r} = 1 / \sqrt{n-3}$ (n = Tamaño de la muestra). Este error estándar es sensible al tamaño de la muestra y tiene una relación inversa con esta última.

Sustituyendo:

$$SE_{z_r} = 1 / \sqrt{20-3}$$

$$SE_{z_r} = 1 / 4.123$$

$$SE_{z_r} = 0.2425.$$

Luego, se estima el $z_{\text{calculado}}$ de la siguiente manera:

$$z_{\text{calculado}} = \frac{(\text{Z transformado } S - \text{Z parámetro de población})}{SE_{z_r}}, \quad (\text{Ecuación 3.13})$$

donde:

- $z_{\text{de la población}}$ = puede ser algún parámetro hipotético de una población con la que se está comparando la estadística de la muestra, *e.g.*, cero o cualquier otro valor que tenga algún antecedente teórico o en la literatura;
- SE_{z_r} = Error estándar de la distribución de muestras del r de Pearson.

Sustituyendo:

$$z_{\text{calculado}} = (0.05005 - 0) / 0.2425 = .2063.$$

Para confirmar este resultado, se puede emplear la calculadora en línea: <https://www.psychometrica.de/correlation.html>. En este caso, esta calculadora comparó el $r = .05$ con $n = 20$ contra una correlación = 0, bajo la categoría *Testing Correlations against a Fixed Value*, la cual se podría traducir como: Poner a prueba las correlaciones en contra de un valor fijo. En definitiva, el resultado fue: $z_{\text{calculado}} = 0.206$ con una probabilidad calculada de .418 (para dos colas). Con esta información, se puede pasar a los criterios para comparar el $z_{\text{calculado}}$ con el $z_{\text{crítico}}$, que son:

- $\chi_{\text{calculado}} \leq \chi_{\text{crítico}}$: no rechazar/fallar en rechazar la hipótesis nula. La conclusión es que no existe correlación entre las variables: $r = 0$;
- $\chi_{\text{calculado}} > \chi_{\text{crítico}}$: rechazar la hipótesis nula. La conclusión es que existe correlación entre las variables: $r \neq 0$ para un test de dos colas; para uno de una cola: $r < 0$ o $r > 0$;
- En este caso, y como se utilizó la calculadora que no proveyó un $\chi_{\text{crítico}}$, se compara un alfa tradicional de .05 con dos colas ($.05 / 2 = .025$) con la correspondiente probabilidad calculada de dos colas (.418), *i.e.*, $\alpha .025 < p = .418$, así que no se rechaza la hipótesis nula, porque la probabilidad es más alta que el alfa y se asume que la correlación de $r = .05$ no es diferente de 0.

El quinto paso es: *llevar a cabo un intervalo de confianza*, que se explicó previamente en la sección *Pasos para calcular un IC de un Coeficiente r*. Bajo la <https://www.psychometrica.de/correlation.html> con *Calculation of Confidence Intervals of Correlations*, se obtuvo un IC_{95%} de [-0.401, .482]. Con este IC, se puede concluir que el r de la población tiene un 95% de probabilidades de aparecer en este rango. Este resultado del IC apoya el Test de Significancia Estadística, porque el rango pasa por el cero, así que cabe la posibilidad de que el r sea cero a nivel de la población.

El sexto paso es: *interpretar los resultados del r de Pearson*. La interpretación se suele hacer para los resultados que arrojen los análisis y lo que algunos teóricos han explicado con respecto al tamaño de algún efecto, como el r , *e.g.*, Cohen (1988). Luego, se pueden comparar dichos resultados con lo que han encontrado otras y otros investigadores. La recomendación a las y los investigadores, es que utilicen las referencias del marco teórico de su investigación, para darle significado a lo que hallaron. En la Tabla 3.10, se muestra un resumen de la serie de pasos antes explicados para poner a prueba la hipótesis nula.

Tabla 3.10 Pasos para poner a prueba una hipótesis nula

Pasos	Test t	Test z
1.º Establecer la o las preguntas de investigación	¿Existe una relación estadísticamente significativa entre la variable x y la variable y ?	¿Existe una diferencia estadísticamente significativa entre el Coeficiente de Correlación de la Población y el de la Muestra?
2.º Derivar las hipótesis de la o las preguntas de investigación	Población: $H_0: r = 0$ $H_A: r \neq 0$ Muestra: $H_0: r = 0$ $H_A: r \neq 0$	Población: $H_0: r =$ Cierta valor $H_A: r \neq$ Cierta valor Muestra: $H_0: r =$ Cierta valor $H_A: r \neq$ Cierta valor
3.º Establecer el criterio para rechazar la hipótesis nula: el nivel del alfa (α)	Las opciones tradicionales son: .01, .05 y .10, y con los df identificar en una tabla el valor $t_{\text{crítico}}$.	Similarmente, las opciones tradicionales son: .01, .05 y .10, y le corresponden los valores z de 2.58, 1.96 y 1.69
4.º Calcular el valor de la prueba estadística (test estadístico)	$* t_{\text{calculado}} = \frac{r}{\sqrt{(n-2)/(1-r^2)}}$	$* z_{\text{calculado}} = \frac{(z_{\text{transformado}} - z_0)}{SE_{zr}}$
5.º Llevar a cabo un intervalo de confianza	El nivel de confianza tradicional es del 95%, lo que quiere decir que de 100 muestras que se tomen del mismo valor, 95 de ellas tendrán el valor real de la población (ver ecuaciones 3.9 y 3.10).	
6.º Interpretar los resultados del r de Pearson	La interpretación se suele hacer, en primera instancia, para los resultados que arrojen los análisis y lo que algunos teóricos han explicado con respecto al tamaño de algún efecto, como el r , e.g., Cohen (1988). Luego, se pueden comparar estos resultados con lo que han encontrado otras y otros investigadores	

g. Un Test t de la hipótesis nula para la r de Pearson

EL EJEMPLO HEURÍSTICO CON LA DISTRIBUCIÓN DE MUESTRAS Y valores t es el siguiente: una investigadora educativa o un investigador educativo tienen la siguiente pregunta: ¿existe una correlación estadísticamente significativa entre el ausentismo escolar y la edad de los estudiantes universitarios? La investigadora o el investigador no encuentran teorías o antecedentes en la literatura, que expliquen y correlacionen las variables del ausentismo y la edad. Por esta razón, su estudio es exploratorio y no tiene una *expectativa* clara del signo de la posible correlación. La investigadora o el investigador logran recolectar los datos de ambas variables de una muestra

($n = 24$) elegida al *azar* de una población de alumnos universitarios. Siguiendo la lista de los pasos, llenan el siguiente procedimiento:

- 1.º *Establecer la o las preguntas de investigación, e.g., una interrogante se puede estipular sin mencionar la significancia estadística: ¿existe una correlación entre el ausentismo escolar y la edad de los estudiantes universitarios? Por otro lado, sería más precisa si se plantea desde la perspectiva de la significancia estadística, i.e., ¿existe una correlación estadísticamente significativa entre el ausentismo escolar y la edad de los alumnos universitarios? Asimismo, es necesario establecer una pregunta de investigación relacionada con el tamaño del efecto (cf. Cumming, 2013). Esta interrogante sería contestada después de la pregunta de significancia estadística: ¿cuál es el tamaño del r ?*
- 2.º *Derivar las hipótesis de la o las preguntas de investigación de significancia estadística. En los Test de Significancia Estadística de este tipo, se asume que no hay correlación a nivel de población, i.e., $\rho = 0$. Como no hay antecedentes, es un estudio exploratorio, y, por lo tanto, no se asume que de existir la correlación sea positiva o negativa. Entonces, el test de dos colas es el apropiado, porque la hipótesis alternativa no especifica el signo de la correlación debido a que es un estudio exploratorio:*

Las hipótesis a nivel de población:

- $H_0: \rho = 0$;
- $H_A: \rho \neq 0$.

Las hipótesis a nivel de muestra:

- $H_0: r = 0$;
- $H_A: r \neq 0$.

- 3.º *Establecer el criterio para rechazar la hipótesis nula: el nivel del alfa (α). Se opta por la probabilidad tradicional de un alfa de .05. Como se tienen dos colas, el alfa se divide*

entre dos: $.05 / 2 = .025$. Asimismo, los grados de libertad (df) se calculan así: $n - 2$, dando como resultado $df = 24 - 2 = 22$. Con estos datos, se va a la Tabla de distribución t para identificar el valor crítico: $t_{\text{crítico}}$. En la Tabla 3.11, se muestra una parte de una Tabla de distribución t , donde se identifica en la columna de la derecha el alfa que corresponde a $.025$ ($.05 / 2$), en color gris, y los $df = 22$, en el renglón que también aparece en color gris. Dadas estas coordenadas, el $t_{\text{crítico}}$ es: 2.074.

Tabla 3.11 Extracto de una Tabla de distribución t

df	Coeficiente Alfa		
	0.10	0.05	0.025
1	3.078	6.314	12.076
2	1.886	2.920	4.303
3	1.638	2.353	3.182
4	1.533	2.132	2.776
5	1.476	2.015	2.571
6	1.440	1.943	2.447
7	1.415	1.895	2.365
8	1.397	1.860	2.306
9	1.383	1.833	2.262
10	1.372	1.812	2.228
11	1.363	1.796	2.201
12	1.356	1.782	2.179
13	1.350	1.771	2.160
14	1.345	1.761	2.145
15	1.341	1.753	2.131
16	1.337	1.746	2.120
17	1.333	1.740	2.110
18	1.330	1.734	2.101
19	1.328	1.729	2.093
20	1.325	1.725	2.086
21	1.323	1.721	2.080
22	1.321	1.717	2.074

4.º *Calcular el valor de la prueba estadística.* Para ello, se usa la siguiente fórmula (Ecuación 3.14) y el procedimiento:

Se tiene: $r = .30$ y $n = 24$:

$$t_{\text{calculado}} = r \left[\sqrt{\frac{(n-2)}{(1-r^2)}} \right]; \quad (\text{Ecuación 3.14})$$

sustituyendo:

$$t_{\text{calculado}} = .30 \left[\sqrt{(24-2) / (1-.30^2)} \right]$$

$$t_{\text{calculado}} = .30 \left[\sqrt{(22) / (1-.09)} \right]$$

$$t_{\text{calculado}} = .30 \left[\sqrt{(22) / (.91)} \right]$$

$$t_{\text{calculado}} = .30 \left[\sqrt{24.18} \right]$$

$$t_{\text{calculado}} = .30 [4.92]$$

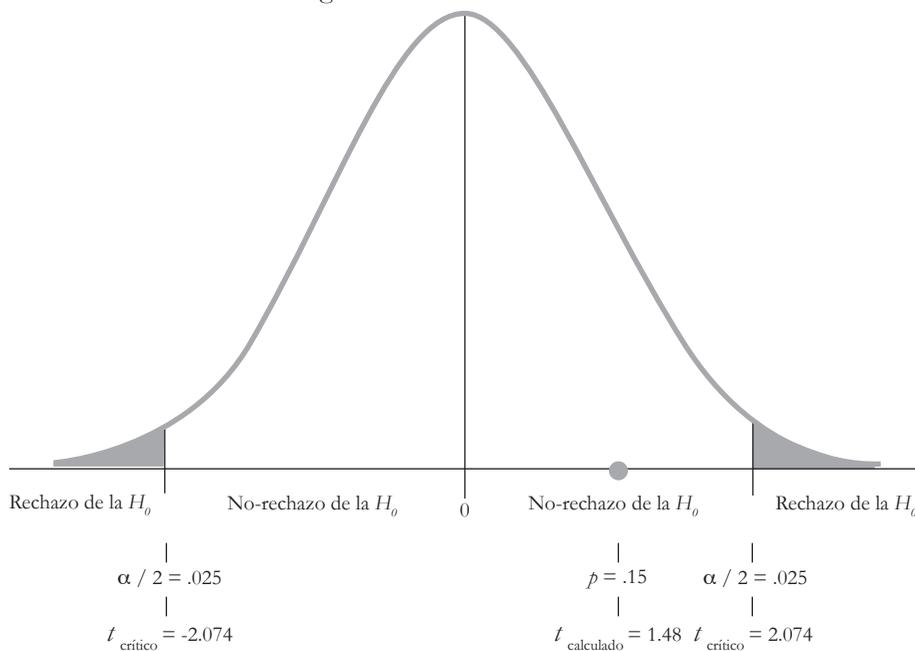
$$t_{\text{calculado}} = 1.48.$$

Nota: un $t_{\text{calculado}}$ se reporta con los grados de libertad: $t_{\text{calculado}} = 1.48$ y $df = 22$.

Más al respecto, este $t_{\text{calculado}}$ (1.48) se puede transformar en una probabilidad $_{\text{calculada}}$, usando una calculadora en línea: <http://www.socscistatistics.com/pvalues/tdistribution.aspx>

Esto se hace al ingresar este $t_{\text{calculado}}$, los $df = 22$, el alfa = .05 y seleccionando dos colas, porque este $t_{\text{calculado}}$ corresponde a una hipótesis alternativa de dos colas: $H_A: \rho \neq 0$. El resultado es una probabilidad $_{\text{calculada}}$: .153 (redondeada es .15). En la Figura 3.21, se muestra la ubicación del $t_{\text{calculado}}$ y su correspondiente probabilidad $_{\text{calculada}}$ en una distribución de muestras con respecto a los valores críticos.

Figura 3.21 Test t de dos colas



5. *Llevar a cabo un intervalo de confianza.* El nivel de confianza tradicional es del 95% ($z = 1.96$), el cual corresponde a la distribución z que se encuentra en la Tabla z . Primero, y usando la Tabla 3.8, se convierte el r (.30) original a un valor z (0.31) con $n = 24$. Las fórmulas para el IC son:

$$IC_{95\%} = z_r + / - [z_{\text{crítico}} (SE_{z_r})],$$

donde:

$$SE_{z_r} = \frac{1}{\sqrt{n-3}} \tag{Ecuación 3.7}$$

En forma extendida:

$$IC_{95\%} = z_{\text{transformado}} + / - z_{\text{crítico}} / 100 \times \frac{1}{\sqrt{n-3}}; \tag{Ecuaciones 3.9 y 3.10}$$

sustituyendo para el error estándar:

$$SE_{\xi} = 1 / \sqrt{(24 - 3)}$$

$$SE_{\xi} = 1 / \sqrt{21}$$

$$SE_{\xi} = 1 / 4.58$$

$$SE_{\xi} = 0.22.$$

Sustituyendo para el IC:

$$IC_{95\%} = 0.31 + / - [(1.96) (0.22)]$$

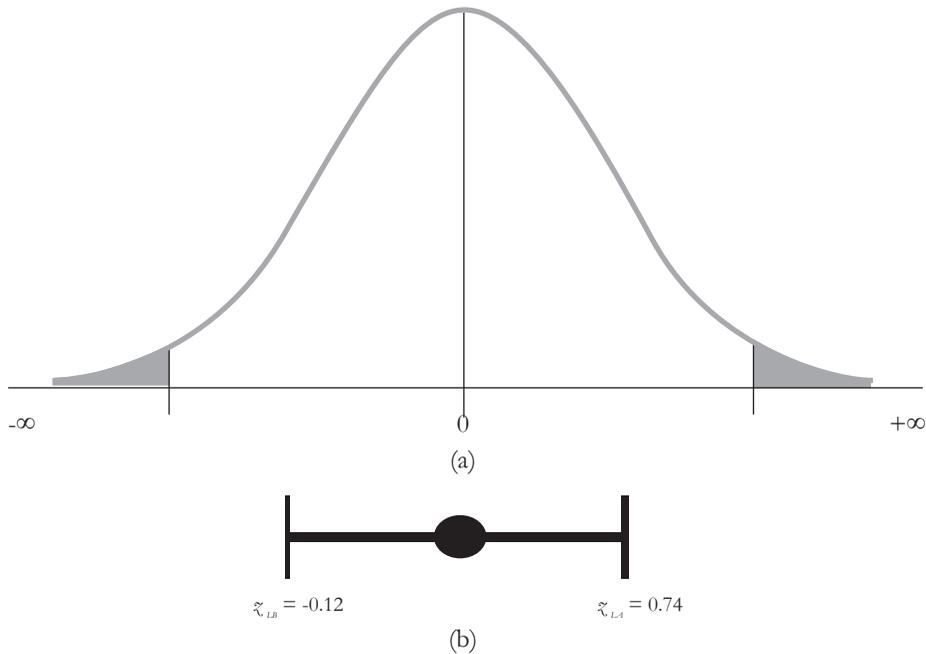
$$IC_{95\%} = 0.31 + / - 0.43: [-0.12, 0.74].$$

Transformando este IC de valores ξ a coeficientes r con la Tabla 3.8:

$$IC_{95\%} \text{ en valores } \xi: [-0.12, 0.74]$$

$$IC_{95\%} \text{ en coeficientes } r: [-.12, .63].$$

En la Figura 3.22, se muestra este IC en una distribución de muestras para los valores ξ únicamente.

Figura 3.22 Distribución z para el IC del 95%

6.º *Interpretar los resultados.* Los resultados muestran que el $t_{\text{calculado}}$ (1.48) < $t_{\text{crítico}}$ (2.07) y la probabilidad calculada (.15) > alfa (.05 / 2 = .025; Figura 3.21). Dados estos resultados, *no* se rechaza la hipótesis nula y la relación entre el ausentismo y la edad de los estudiantes es estadísticamente *no* significativa. Es decir, la evidencia apoya la hipótesis nula porque el $t_{\text{calculado}}$ no cae en la zona de rechazo. Bajo el supuesto de una muestra representativa, otra inferencia es que si la población contiene un $\rho = 0$, entonces la probabilidad de la investigadora o el investigador de haber encontrado por casualidad un Coeficiente $r = .30$, es de más del .05 de probabilidad (*i.e.*, probabilidad $\text{calculada} = .153$). En otras palabras y dados los criterios de rechazo, la probabilidad calculada era alta de encontrar una correlación de este tamaño al azar, dados los $df = 22$.

Con respecto al tamaño del efecto, este es mediano cuando es .30 (véase: Cohen 1988), y del $IC_{95\%}$ para el r [-.12, .63], se puede concluir que el valor de la correlación de la población se encuentra entre estos límites con un 95% de confianza, lo que quiere decir que de 100 muestras de ese tamaño que se tomen, 95 de ellas capturarán la correlación verdadera de la población (ρ). Se puede observar la magnitud de este $IC_{95\%}$, que incluye el cero, así que la correlación puede ser nula.

Dadas estas evidencias de significancia estadística e IC , la investigadora o el investigador podrían concluir que la correlación no es diferente de cero, aunque el efecto fue mediano; es decir, probablemente no exista una correlación en la población. Algo que se podría hacer para remediar la situación es aumentar el tamaño de la muestra, porque los análisis de significancia estadística son sensibles al tamaño de esta, lo cual implica que si se tiene una muestra pequeña ($n = 20$) con un efecto mediano (e.g., $r = .35$), probablemente no sea posible rechazar la hipótesis nula, *ceteris paribus*. En contraparte, si se tuviera este mismo efecto ($r = .35$), pero con $n = 1000$, sí sería posible rechazar la hipótesis nula, *ceteris paribus*.

Un Test z de la hipótesis nula para la r de Pearson

OTRO EJEMPLO DE UN TEST DE SIGNIFICANCIA ESTADÍSTICA, COMO el Test z , es cuando existen antecedentes en la literatura (e.g., $\rho = .50$). Con este antecedente, se podría hacer una comparación con un coeficiente que se hubiera calculado de una muestra de dicha población en cuestión. Para el siguiente ejemplo heurístico, se toman algunos datos de uno de los ejemplos anteriores. Esto es, la pregunta de investigación del ejemplo original es: ¿existe una correlación entre el *Tiempo en completar el test* y la variable dependiente *Puntajes en un test de matemáticas*? La respuesta a esta interrogante, en síntesis, es: sí ($r = .98$). Para ajustar esta pregunta al Test z y comparar el valor de la literatura ($\rho = .50$) con el de la muestra ($r = .98$), se redacta de esta manera: ¿existe una diferencia *estadísticamente significativa* entre el parámetro de la población y la estadística de la muestra con la correlación del *Tiempo en completar el test* y los *Puntajes en un test de matemáticas*? De la respuesta a esta interrogante, se podría observar hasta dón-

de es sostenible una posible generalización. En otras palabras, los resultados indicarían hasta dónde la correlación de la muestra corresponde a la de la población.

A continuación, se lleva a cabo un Test de Significancia Estadística de esta transformación con el $r = .98$ de uno de los ejemplos anteriores, usando un $\alpha = .05$ y $n = 25$ con dos colas.

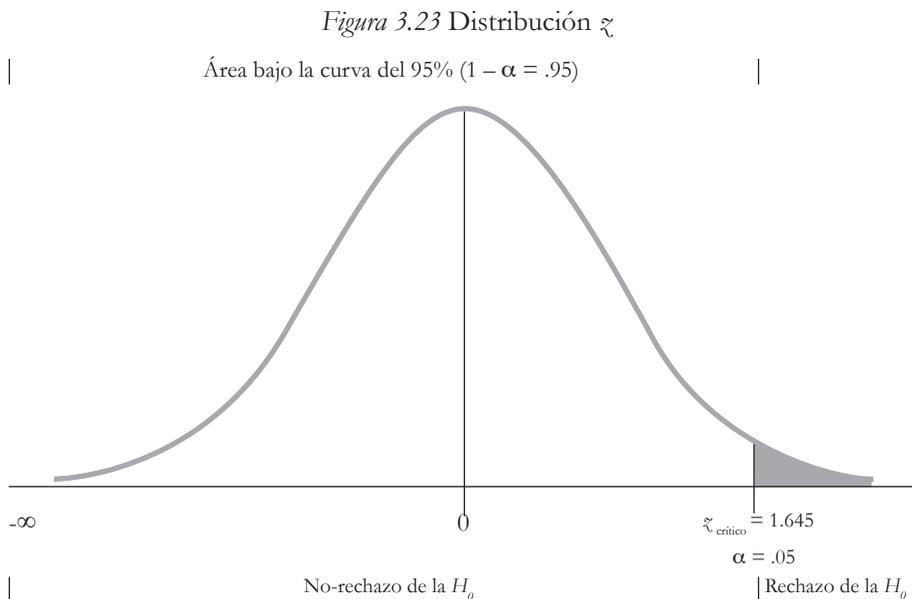
- 1° *Establecer la o las preguntas de investigación.* ¿Existe una diferencia estadísticamente significativa entre el parámetro de la población y la estadística de la muestra con la correlación del *Tiempo en completar el test* y los *Puntajes en un test de matemáticas*?
- 2° *Derivar las hipótesis de la o las preguntas de investigación.* Hipotéticamente hablando, hay un antecedente en la literatura que precisa una correlación en la población a la cual se quiere generalizar: $\rho = .50$. Por esta razón, este coeficiente hipotético se convierte en la hipótesis nula y la alternativa:

- Población:
 - Hipótesis nula (H_0): $\rho = .50$;
 - Hipótesis alternativa (H_a): $\rho \neq .50$.

Nota: se utiliza un test de una cola positiva, porque existe el antecedente de un Coeficiente r en la población.

- Muestra:
 - Hipótesis nula (H_0): $r = .50$;
 - Hipótesis alternativa (H_a): $r \neq .50$.

- 3° *Establecer el criterio para rechazar la hipótesis nula: el nivel del alfa (α).* Se opta por la probabilidad tradicional alfa de .05. Esta área de no-rechazo de la hipótesis nula es: $.50 + .45 = .95$ y le corresponde un valor z de 1.645 (Tabla 3.23).



4.º *Calcular el valor de la prueba estadística.* Hay que transformar el Coeficiente r en un valor z . Empleando la mencionada fórmula de Fisher de transformación de la r a la z :

Sustitución del $r_{\text{calculado}} = .98$ en la Ecuación 3.8:

$$z_{\text{transformado}} = .5 [In(1 + .98) - In(1 - .98)]$$

$$z_{\text{transformado}} = .5 [In(1.98) - In(.02)]$$

$$z_{\text{transformado}} = .5 [.68 - (-3.9120)]$$

$$z_{\text{transformado}} = .5 [.68 + 3.912]$$

$$z_{\text{transformado}} = .5 [4.592]$$

$$z_{\text{transformado}} = 2.296.$$

Nota: In = Logaritmo natural.

También, se tiene que transformar el Coeficiente de Correlación de la Población $\rho = .50$ en un valor z . Para simplificar el procedimiento, se va a la Tabla 3.8 y se ubica el Coeficiente de Correlación ($\rho = .50$), al cual le corresponde un valor $z = .55$.

Ahora es necesario calcular el valor del error estándar de la estadística con la Ecuación 3.7:

$$SE_{z'} = 1 / \sqrt{n - 3};$$

sustituyendo:

$$SE_{z'} = 1 / \sqrt{25 - 3}$$

$$SE_{z'} = 1 / \sqrt{22}$$

$$SE_{z'} = 1 / 4.69$$

$$SE_{z'} = 0.21.$$

Calculando el *valor de prueba estadística* ($z_{\text{calculado}}$):

$$z_{\text{calculado}} = (z_{\text{transformado}} - z_{\rho}) / SE_{z'}; \quad (\text{Ecuación 3.13})$$

donde:

- $z_{\text{transformado}} = r$ de Pearson transformado en valor z 2.296;
- z_{ρ} = Valor hipotético de este parámetro de la población transformado en valor z : .55;
- $SE_{z'} =$ Error estándar: .21.

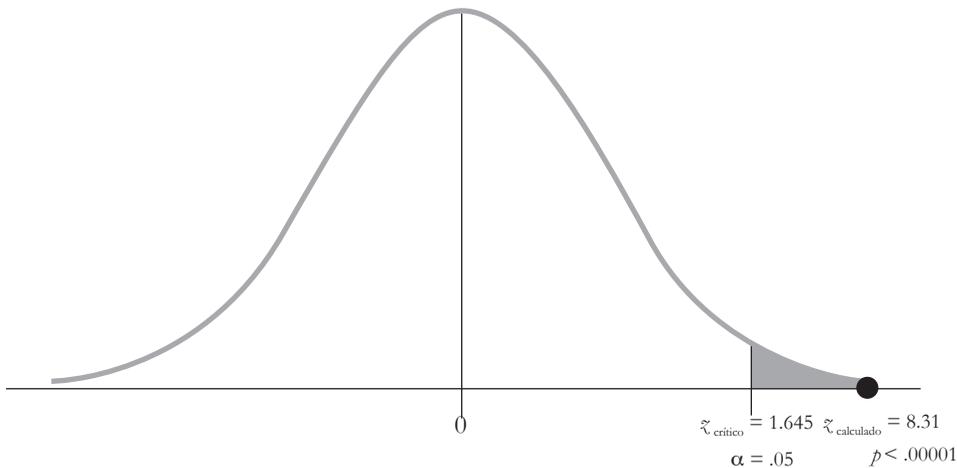
Sustituyendo:

$$z_{\text{calculado}} = (2.296 - .55) / 0.21 = 8.31.$$

Para convertir el $z_{\text{calculado}} = 8.31$ en una probabilidad $p_{\text{calculada}}$, se usa la calculadora <http://www.socscistatistics.com/pvalues/normaldistribution.aspx>, donde se especifica $\alpha = .05$ y dos colas. De esta calculadora, se obtiene: probabilidad $p_{\text{calculada}} < .00001$.

Entonces, se tiene el Test z con el criterio de un $\alpha = .05$, que corresponde al $z_{\text{crítico}} = 1.645$ para una cola. Se hace la comparación: $z_{\text{crítico}} = 1.645 < z_{\text{calculado}} = 8.31$ y $\alpha = .05 >$ probabilidad calculada $< .00001$. Entonces, se rechaza la hipótesis nula, porque el valor $z_{\text{calculado}}$ está en el área sombreada de rechazo de la hipótesis nula (Figura 3.24).

Figura 3.24 Test z de una hipótesis nula con una sola cola



5.º Llevar a cabo un intervalo de confianza. Asimismo, el nivel de confianza tradicional es del 95%, el cual corresponde a la distribución z que se encuentra en la Tabla z : 1.96. El $r (.98)$ se convirtió en $z_{\text{transformado}} = 2.296$. La fórmula es ($n = 25$):

$$IC_{95\%} = z_{\text{transformado}} \pm [(z_{\text{crítico}}) (SE_{z})]; \quad (\text{Ecuaciones 3.9 y 3.10})$$

Los datos para llenar la fórmula son:

$$z_{\text{transformado}} = 2.296;$$

$$z_{\text{crítico}} = 1.96;$$

$$SE_{z} = .21.$$

Sustituyendo para el IC:

$$IC_{95\%} = 2.296 + / - [(1.96) (0.21)]$$

$$IC_{95\%} = 2.296 + / - .41: [1.89, 2.71].$$

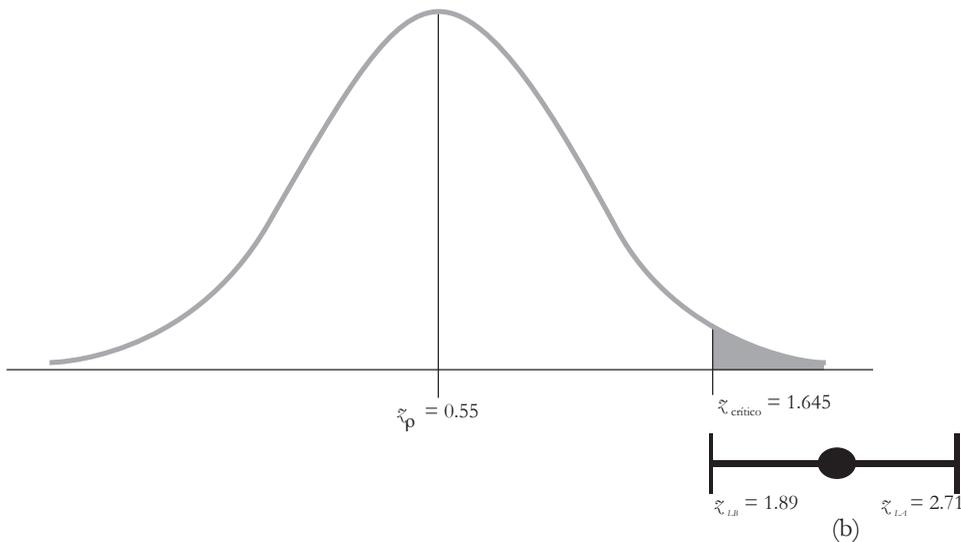
Transformando este IC de valores z a coeficientes r con la Tabla 3.8:

IC_{95%} en valores z : [1.89, 2.71];

IC_{95%} en coeficientes r : [.95, .99].

Nota: la conversión de valores z a r es una aproximación. En la Figura 3.25, se muestra el IC del 95% en valores z .

Figura 3.25 Distribución z para el IC del 95%



6.º *Interpretar los resultados.* Los resultados muestran que el $z_{\text{calculado}} = 8.31 > z_{\text{crítico}}$ (1.645) y la probabilidad $_{\text{calculada}} < (.00001) < \text{alfa } (.05)$. Dados estos resultados, se rechaza la hipótesis nula y la diferencia entre la correlación de la muestra ($r =$

.98) y la población ($\rho = .50$) es estadísticamente significativa, lo que indica que la probabilidad calculada es menor al criterio usado (alfa) cuando se asume que la correlación en la población es *nula* (i.e., $\rho = 0$). Es decir, la evidencia *no* apoya la hipótesis nula, porque el $z_{\text{calculado}}$ cae en la zona de rechazo. Otra inferencia es que si la población contiene un $\rho = .50$ y una muestra ($r = .98$), entonces la probabilidad de la investigadora o el investigador de haber encontrado por casualidad esta diferencia es de menos de .00001 ($p_{\text{calculada}}$). En otras palabras, y dados los criterios de rechazo, la probabilidad era baja de encontrar esta diferencia al azar entre las correlaciones de la población y una muestra cuando se asume que la hipótesis nula era cierta ($\rho = 0$).

Respecto al IC_{95%} [1.89, 2.71], se puede concluir que el valor de la correlación se encuentra entre estos límites con un 95% de confianza (Figura 3.25), lo que quiere decir que de 100 muestras de ese tamaño que se tomen, 95 de ellas capturarán la correlación verdadera de la población (ρ).

Dadas estas evidencias de significancia estadística e IC, la investigadora o el investigador podrían concluir que *probablemente* exista una diferencia entre la correlación de la muestra y la de la población.

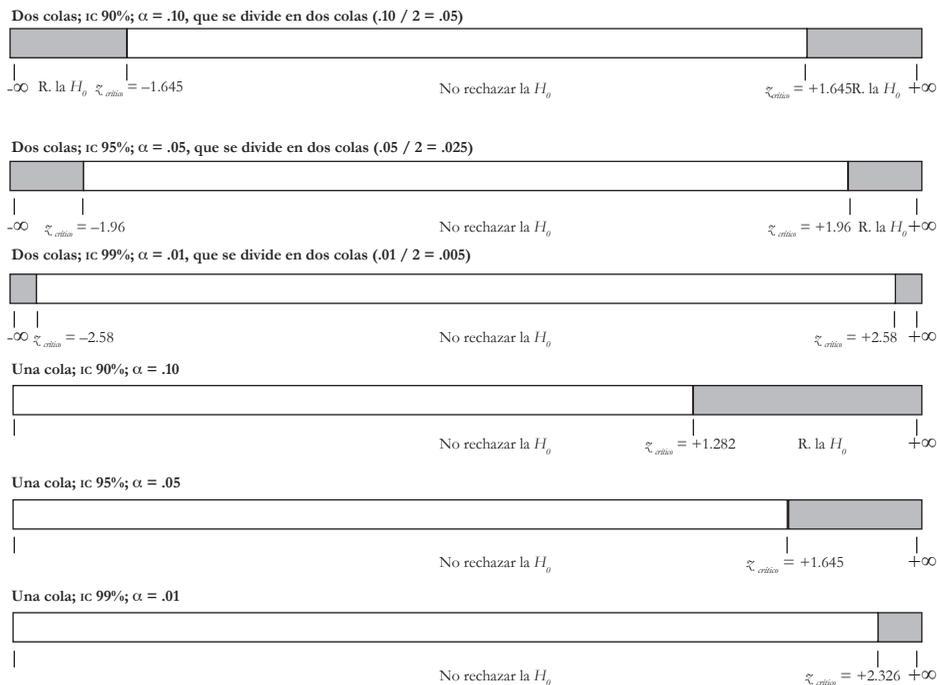
i. Resumen de Test de Significancia Estadística para una y dos colas

ASIMISMO, EN LA FIGURA 3.26 SE MUESTRAN VARIOS TEST Z: UNA serie de valores críticos z con el objetivo de identificar los puntos donde se rechazaría la hipótesis nula y donde no se rechazaría (i.e., se fallaría en rechazar, dados diferentes niveles de significancia estadística [α]). Un ejemplo para dos colas es el valor crítico en término absoluto de $|1.645|$, al que le corresponde un $\alpha = .10$. *Recordando*: si se está sometiendo a prueba una hipótesis nula ($H_0: \rho = 0; H_A: \rho \neq 0$), para poder rechazar la H_0 , se tendría que calcular de los datos un r para transformarlo en z y hacer la comparación que corresponda. Para la cola derecha ($\alpha = .10$), es la misma hipótesis nula ($H_0: \rho = 0$), pero la hipótesis alternativa es positiva ($H_A: \rho > 0$). El valor crítico z es de

1.282, así que para poder rechazar la H_0 sería necesario obtener un $z > 1.28$. Como se puede observar, para poder rechazar la H_0 , se necesita un valor $z_{\text{calculado}}$ menor para una cola que para dos colas, *ceteris paribus*. La recomendación para las y los investigadores, es tener una H_A que indique el sentido de la correlación cuando haya antecedentes en la literatura, porque esto aumenta las posibilidades de rechazar una hipótesis nula y, además, ayuda a llenar vacíos en la literatura de ese campo del conocimiento.

La cola izquierda no está incluida en la Figura 3.26 para evitar redundancias. Sin embargo, los niveles del alfa corresponderían a los mismos valores $z_{\text{críticos}}$ con signo negativo, *i.e.*, -1.282 ($\alpha = .10$); -1.645 ($\alpha = .05$); y -2.326 ($\alpha = .01$). La hipótesis nula seguiría igual: $H_0: \rho = 0$. La hipótesis alternativa sí cambiaría a $H_A: \rho < 0$. Para poder rechazar la hipótesis nula, el valor $z_{\text{calculado}}$ tendría que ser menor al $z_{\text{crítico}}$. Un ejemplo para rechazar la hipótesis nula es: $z_{\text{crítico}} = -1.282 > z_{\text{calculado}} = -5.25$ y, por lo tanto, caería en el área de rechazo de la hipótesis nula.

Figura 3.26 Valores críticos z y sus correspondientes niveles alfa



Nota: la zona gris para una y dos colas indica el área de rechazo de la H_0 (hipótesis nula). Estos valores críticos, se pueden obtener en la calculadora en línea de Mathcracker: https://mathcracker.com/z_critical_values.php

Preguntas para reflexionar

1. ¿Qué correlaciones son más apropiadas para ser utilizadas con escalas tipo Likert?
2. ¿Cuáles son las características del r^2 ?
3. ¿Qué ventajas tienen algunas de las fórmulas del r^2 ?
4. ¿Qué significa el r^2 ?
5. ¿Qué ejemplos de correlación se podrían dar en la investigación educativa?

6. ¿Cuál es el significado de rechazar la hipótesis nula?
7. ¿Qué implica un intervalo de confianza para el r^2 ?
8. ¿Qué se debe hacer cuando no se cumplen los requisitos del r^2 ?
9. ¿Qué tipo de preguntas de investigación se pueden resolver con el uso del r^2 ?

Opinión del autor

¿QUÉ ES UNA CORRELACIÓN, CÓMO SE USA; CÓMO SE PONE A PRUEBA una hipótesis y cómo se interpreta? En este capítulo, se trata de responder esta pregunta compuesta. Al darle respuesta, se espera que el panorama se abra para la lectora o el lector al poder aplicar una serie de conceptos que se desarrollan en el mismo capítulo para los análisis estadísticos con correlaciones. Asimismo, los principios de correlación, aquí empleados, se aplican a análisis más sofisticados, como análisis de la varianza, análisis de la covarianza, regresión, análisis discriminante, ecuaciones estructurales, etcétera.

Lo más relevante de esta sección, es que el Coeficiente r evidencia la fortaleza de la relación entre variables, las cuales pueden ser en la investigación educativa: aprendizajes en ciencias naturales, matemáticas, ciencias sociales, español, inglés, entre muchos otros. Dichos aprendizajes pudieron haberse medido operacionalmente por calificaciones y se esperaría que estas, que representan los aprendizajes, estuvieran positivamente correlacionadas unas con otras. Pero si no, habría que indagar las posibles razones por las que no lo estuvieran: exámenes que no correspondieran a las lecciones en las clases; estudiantes que solo tuvieran interés en algunas materias; diferentes formas de calificar a los alumnos, etcétera. Existen varias teorías de las inteligencias múltiples, que sustentan que la inteligencia y los aprendizajes en muchas de las áreas escolares están altamente correlacionados. Por ello, habría que poner atención en cómo se correlacionan las calificaciones de los estudiantes.

Otro aspecto que encuentro de suma importancia, es el r^2 (varianza explicada). Este coeficiente ha resultado difícil de entender para muchos de los que han sido mis pupilos. Un pequeño ejemplo de lo anterior, es que una variable como el aprendizaje de algo (*e.g.*, las matemáticas) depende de muchas otras variables. En primer

lugar, los conocimientos difieren de personas egresadas de algún pregrado, es decir, probablemente las y los egresados de ingeniería y ciencias naturales obtendrían más altos puntajes en exámenes de matemáticas que el resto de sus contrapartes de arte y ciencias sociales. Uno de los supuestos es que los puntajes reflejan los conocimientos. Otra variable relacionada con la literatura en los aprendizajes de las personas, es el nivel educativo de los padres, *i.e.*, entre más educación tienen los padres más educación tendrán sus hijas e hijos. En concreto, si se relacionaran los conocimientos de matemáticas, como la variable dependiente, y como variables independientes, la carrera y la educación de los padres, habría posiblemente una correlación alta entre estas. De esta correlación, se puede calcular el r^2 y con este, estimar el porcentaje de variación de los conocimientos de matemáticas, que son explicados por la carrera y la educación de los padres.

Algo que hay que tener en mente, es que el r se obtiene de una muestra que representa una población. Puede pasar que al rechazar o fallar en rechazar una hipótesis nula, se cometa un error, es decir, se puede encontrar una correlación en la muestra que *no* exista en la población. También, puede pasar lo contrario: que no se encuentre una correlación en la muestra que *sí* exista en la población. Para minimizar errores, es necesario ver si el tamaño de la muestra permite alcanzar el nivel del 80% del poder estadístico. Habría que ver a Cohen (1988) en la sección de Poder estadístico, para analizar el tamaño de la muestra. También, cómo jugar con los números en el *software* de *Gpower* (para más detalles de este *software*, véase a Ponce [2019]).

Suele pasar que en los estudios de investigación educativa, se reporte simplemente el coeficiente que se obtuvo sin considerar el margen de error. Dicho reporte estaría incompleto si solo se establece un punto de estimación (*i.e.*, un coeficiente). Lo apropiado es, según la Asociación Americana de Psicología y muchos otros autores, que se reporte un intervalo de confianza donde se incluya este margen de error. Un intervalo de confianza con un nivel de confianza del 99%, tendrá más posibilidades de capturar el valor verdadero de una población que uno de menor nivel de confianza.

Una última sugerencia del r , es que no se puede utilizar en todas las ocasiones. Hay que considerar las escalas en las que se midieron las variables, la linealidad entre las variables, la distribución normal, la homocedasticidad, observaciones atípicas, etcé-

tera. Existen múltiples maneras de correlacionar variables que no reúnen los requisitos antes mencionados. Por ello, recomiendo consultar a Kotz (2006) y Salkind (2007), quienes mostraron en sus respectivas obras varias maneras de correlacionar variables en forma paramétrica y no-paramétrica.

CAPÍTULO 4

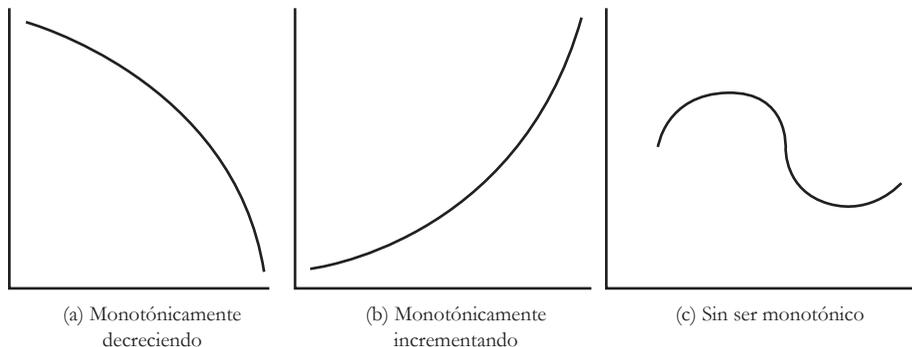
Coeficiente de Correlación ρ de Spearman

Una pregunta de investigación puede plantear la posibilidad de una relación entre variables, *e.g.*, ¿existe una relación entre el ausentismo y el aprendizaje? Antes de pensar en algún tipo de correlación, como la r de Pearson, habría que ver la forma operacional de las variables. La forma operacional del ausentismo podría ser una escala dicotómica también conocida como nominal, *i.e.*, la o el estudiante tiene o no muchas faltas. Asimismo, el aprendizaje podría ser medido en una escala dicotómica: aprobó o no aprobó una clase. Dada esta situación, el r de Pearson no sería la estadística adecuada, porque son variables sin escalas continuas o, por lo menos, ordinales, como se usa en la práctica el r . De hecho, cuando se tienen dos variables medidas en forma dicotómica, pero que implican una escala continua (*i.e.*, el ausentismo se pudo haber medido en número de días ausentes y el aprendizaje, en calificaciones), se puede emplear la correlación tetracórica (para más detalles, véase: Jöreskog, & Sörbom, 1996). Además de la escala para correlacionar, hay que cerciorarse de otros requisitos: normalidad de los datos, linealidad, uso en la literatura de ese análisis, etcétera. Cuando *no* se cumplen algunos de los supuestos del r de Pearson, habría que considerar la correlación de Spearman como una alternativa de sustitución, entre otras posibilidades.

EXISTEN VARIAS ALTERNATIVAS AL R DE PEARSON EN LA investigación educativa (*i.e.*, coeficientes de correlación como *Gamma de Goodman* y *Kruskal, biseriales* y *Tau de Kendall*, entre otros; consultar a: Kotz [2006] y Salkind [2007], para más detalles). Una de estas alternativas es el Coeficiente de Correlación de Orden Jerárquico de *Spearman*. También, conocido simplemente por correlación *ro* de Spearman (*i.e.*, el *r* de Spearman o el *r_s*). A diferencia de la *r* de Pearson, el coeficiente *r_s* es una medida no-paramétrica. Al respecto, Kotz (2006, p. 1792) explica que las “...técnicas estadísticas no-paramétricas implican una estimación o inferencia que no está directamente relacionada con parámetros”. Los parámetros se encuentran en las poblaciones (*e.g.*, promedio, desviación estándar y el *r*, entre otros). Según Kotz (2006), como una técnica no-paramétrica, el *r* de Spearman tendría las siguientes ventajas:

1. Facilidad para entenderlo y aplicarlo;
2. Uso con pequeñas muestras;
3. Medición de los datos en escalas ordinal o de *ranking*;
4. Aplicación en datos incompletos o imprecisos;
5. Uso en una gran cantidad de disciplinas.

Al igual que el *r*, el ρ de Spearman indica la fuerza y la dirección de asociación entre dos variables (*i.e.*, va desde -1 hasta +1). La dirección de la asociación puede ser negativa o positiva, pero los datos deben seguir un patrón *monotónico*, *i.e.*, una variable se incrementa y la otra también, o una aumenta y la otra no (Laerd Statistics, 2017). En la Figura 4.1, se muestran las dos opciones para calcular un ρ y la tercera, es que posea un set de datos sin ser monotónico, con los cuales no se podría calcular el coeficiente.

Figura 4.1 Correlación r de Spearman

El ρ de Spearman se emplea con escalas que están en forma de *ranking* (*i.e.*, la distancia entre los puntos no es necesariamente la misma). Por ejemplo, en una Olimpiada del Conocimiento, un participante puede quedar en primer lugar al obtener el 100% de los aciertos; otro participante, en segundo lugar con el 90%; y otro participante, en tercer lugar con el 60% de los aciertos. La distancia entre el 100% y el 90% no es la misma que entre el 90% y el 60%. Sin embargo, lo importante para el uso del ρ de Spearman es el *ranking*: 1.º, 2.º y 3.º, que serían empleados para calcular este coeficiente al utilizar una segunda variable convertida en *ranking* también. Asimismo, se puede usar el ρ de Spearman con frecuencias, escalas ordinales (*i.e.*, tipo Likert) o datos continuos cuando no se cumple con los supuestos necesarios para realizar la r de Pearson. El primer paso es convertir estas escalas a *rankings*, para luego aplicar la fórmula para calcular el ρ de Spearman (Ecuación 4.1):

$$\rho = 1 - \left(\frac{6\sum d^2}{n(n^2 - 1)} \right), \quad (\text{Ecuación 4.1})$$

donde:

- d^2 = Diferencia entre los *rankings* elevada al cuadrado;
- n = Número de *rankings* en pares.

Ejemplo heurístico con la correlación de Spearman

SE TIENEN DOS MAESTROS A QUIENES SE LES PIDE QUE ORDENEN una serie de doce mejores prácticas en la educación, *i.e.*, se le asigna el número uno a la práctica que es la más importante (*ranking* 1); el número dos a la segunda más importante (*ranking* 2); ... y el número doce a la última más importante (*ranking* 12). Estas mejores prácticas fueron recopiladas por Brophy (2001) y se tradujeron al español para este capítulo (Tabla 4.1).

Tabla 4.1 Mejores prácticas de Brophy (2001)

- | |
|---|
| <p>A. Aprendizaje de cooperación: los estudiantes muchas veces se benefician de trabajar en parejas o en pequeños grupos, para construir o ayudar los unos a los otros a dominar destrezas y habilidades.</p> <p>B. Establecimiento de metas para el aprendizaje: los docentes pueden preparar a los alumnos para el aprendizaje, al proveer una estructura para clarificar los productos buscados y describir estrategias de aprendizaje deseadas.</p> <p>C. Construir un andamiaje para el involucramiento en completar asignaturas: los maestros proveen la asistencia necesaria para involucrar a los estudiantes productivamente en actividades de aprendizaje.</p> <p>D. Evaluaciones con un objetivo: los docentes usan una variedad de métodos de evaluación formales e informales, para monitorear el progreso hacia los objetivos del aprendizaje.</p> <p>E. Discurso reflexivo: las preguntas son planeadas para involucrar a los alumnos en debates alrededor de ideas centrales sobre temas.</p> <p>F. Clima de apoyo en el aula: los estudiantes aprenden mejor dentro de un medio de cohesión y solidaridad.</p> <p>G. Estrategias para ser autodidacta: los maestros modelan e instruyen a los alumnos con estrategias de cómo aprender por su cuenta.</p> <p>H. Expectativas de progreso: los docentes establecen y se cercioran de que las expectativas de los productos del aprendizaje sean apropiadas.</p> <p>I. Oportunidades para aprender: los estudiantes aprenden más cuando el tiempo de las clases es dedicado a actividades relacionadas con la currícula, y el sistema de disciplina del aula promueve la dedicación de los pupilos a las tareas encomendadas.</p> <p>J. Práctica y aplicación de actividades: los alumnos necesitan suficientes oportunidades para practicar y aplicar lo que están aprendiendo, así como recibir retroalimentación al respecto.</p> <p>K. Contenido coherente: para facilitar un aprendizaje con significado y retención para los estudiantes, el contenido es explicado claramente y desarrollado con énfasis en su estructura y conexiones con conocimientos previos.</p> <p>L. Alineación de la currícula: todos los componentes de la currícula están alineados para crear un programa compacto para alcanzar los propósitos y metas institucionales.</p> |
|---|

En la Tabla 4.2, se muestran los resultados de los dos maestros con sus evaluaciones de las doce mejores prácticas: los *rankings*. Por ejemplo, el docente 1 estableció que la mejor práctica A es la más importante: 1. En cambio, el maestro 2 dijo que la mejor práctica A es la segunda más importante: 2.

Tabla 4.2 Resultados de las evaluaciones de mejores prácticas de Brophy (2001)

Mejor práctica	Maestro 1	Maestro 2	d Diferencia entre el ranking de los maestros 1 y 2	d^2
A	1	2	$(1 - 2) = -1$	1
B	2	1	$(2 - 1) = 1$	1
C	3	5	$(3 - 5) = -2$	4
D	4	3	$(4 - 3) = 1$	1
E	5	4	$(5 - 4) = 1$	1
F	6	6	$(6 - 6) = 0$	0
G	7	8	$(7 - 8) = -1$	1
H	8	7	$(8 - 7) = 1$	1
I	9	9	$(9 - 9) = 0$	0
J	10	12	$(10 - 12) = -2$	4
K	11	10	$(11 - 10) = 1$	1
L	12	11	$(12 - 11) = 1$	1
$\sum d^2$				16
Constante				6
Multiplicación $(6 \sum d^2)$				$16 \times 6 = 96$
n				12
n^2				144
$n(n^2 - 1)$				$12(144 - 1) = 1,716$
$(6 \sum d^2) / n(n^2 - 1)$				$96 / 1716 = .056$
Constante				1
r				$1 - .056 = .94$

Nota: Usando la fórmula de la correlación de Spearman (Ecuación 3.15).

Para darle significado al ρ que se encontró, se toman en cuenta algunos de los tamaños que se manejan para el ρ de Spearman, los cuales son los siguientes:

0 a $|.19|$: muy débil;

- |.20| a |.39| : débil;
- |.40| a |.59| : moderado;
- |.60| a |.79| : fuerte;
- |.80| a |1| : muy fuerte.

Entonces, se obtuvo un $\rho = .94$ al correlacionar los *rankings* de las evaluaciones de los dos maestros. Por lo tanto, se concluye que hay una correlación *muy fuerte* entre los *rankings* de ambos docentes. Se podría especular que comporten ciertos valores, experiencias, etcétera. Sin embargo, no se podría decir que dichos valores o experiencias sean las causas de que hayan evaluado a las mejores prácticas, porque las correlaciones no explican en sí mismas las relaciones de causa y efecto. El *software* SPSS ofrece la posibilidad de hacer un Test de Significancia Estadística para el ρ de Spearman, pero está más allá de los propósitos de la presente obra ahondar en este tema. Para más al respecto, se recomienda ver a: Kotz (2006) y Salkind (2007).

Preguntas para reflexionar

1. ¿Qué otros coeficientes de correlaciones existen además de los listados en este capítulo?
2. ¿Qué desventajas tiene el r con respecto al ρ de Spearman?
3. ¿En qué ocasiones es más apropiado usar el ρ que el r ?

Opinión del autor

EL ρ DE SPEARMAN ES UNA ALTERNATIVA AL R CUANDO NO SE CUMPLEN algunos requisitos de este último. Solo habría que cerciorarse de que los datos están en forma monotónica en un plano cartesiano para usarlo. Una vez que esto se identifica, se puede convertir la escala original a la escala de *rankings* y aplicar la fórmula. También, el *software* SPSS se utiliza para calcular este coeficiente.

Referencias

- American Psychological Association (APA) (2015). *APA Dictionary of Psychology* (2nd Ed.). American Psychological Association.
- Berenson, M. L., Levine, D. M., & Krehbiel, T. C. (2012). *Basic Business Statistics: Concepts and Applications* (12nd Ed.). Pearson.
- Brophy, J. (2001). Introduction. In J. Brophy (Ed.), *Advances in Research on Teaching* (Vol. 8, pp. 1-23). JAI Elsevier.
- Byrne, B. M. (2009). *Structural Equation Modeling with Lisrel, PRELIS, and SIMPLIS: Basic Concepts, Applications and Programing*. Psychology Press.
- Cain, M. K., Zhang, Z., & Yuan, K. H. (2017). Univariate and Multivariate Skewness and Kurtosis for Measuring Nonnormality: Prevalence, Influence and Estimation. *Behavior Research Methods*, 49(5), 1716-1735.
- Chandra, R. (2014). *Combinations and Permutations* (1st Ed.). Notion Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd Ed.). Psychology Press.
- Cumming, G. (2013). *Understanding the New Statistics: Effect Sizes, Confidence Intervals, and Meta-analysis*. Routledge.
- Ercikan, K., & Roth, W. M. (2006). What Good is Polarizing Research into Qualitative and Quantitative? *Educational Researcher*, 35(5), 14-23.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2014). *Multivariate Data Analysis* (7th Ed.). Pearson New International Edition.

- Helen, W. M. (1940). Degree Freedom. *Journal of Educational Psychology*, 31(14), 253-269.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for Behavioral Sciences* (5th Ed.). Houghton Mifflin.
- Holgado-Tello, F. P., Chacón-Moscoso, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric Versus Pearson Correlations in Exploratory and Confirmatory Factor Analysis of Ordinal Variables. *Quality and Quantity*, 44.
- Hurley, P. J. (2014). *A Concise Introduction to Logic* (12th Ed.). Cengage Learning.
- Johnson, R., & Bhattacharyya, G. (1987). *Statistics Principles and Methods: Revised Printing*. John Wiley & Sons.
- Jöreskog, K. G., & Sörbom, D. (1996). *Lisrel 8: User's Reference Guide*. Scientific Software International.
- Kotz, S. (Ed.) (2006). *Encyclopedia of Statistical Sciences* (16 Vols., 2nd Ed.). Wiley-Interscience.
- Laerd Statistics (2017). <https://statistics.laerd.com/spss-tutorials/spearmans-rank-order-correlation-using-spss-statistics.php>
- Lane, D. M., Scott, D., Hebl, M., Guerra, R., Osherson, D., & Zimmer, H. (2014). *Introduction to Statistics*. Rice University, 474-476.
- Ponce, H. F. (2019). *Conceptos básicos de estadísticas inferenciales aplicadas a la investigación educativa*. Universidad Autónoma de Ciudad Juárez.
- Ponce, H. F., Domínguez, C. T., & Arriaga, M. (2016). La importancia de la investigación en la educación especial. *Nóesis*, 25(50), 216-242.
- Salkind, N. J. (2007). *Encyclopedia of Measurement and Statistics* (Vols. 1, 2, & 3). Sage.
- (2011). *Statistics for People Who Think They Hate Statistics* (4th Ed.). Sage.
- Schneider, B., Carnoy, M., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating Causal Effects using Experimental and Observational Designs*. American Educational Research Association.
- Schumacker, R., & Tomek, S. (2013). *Understanding Statistics: Using R*. Springer.
- Spiegel, M. R., Schiller, J., & Srinivasan, R. A. (2000). *Probability and Statistics* (2nd Ed.). McGraw-Hill.
- Tabachnick, B. G., & Fidell, L. S. (2012). *Using Multivariate Statistics* (6th Ed.). Pearson.

Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. American Psychological Association.

Weisstein, E. W. (2020). Bivariate Normal Distribution. *MathWorld* – A Wolfram Web Resource. <http://mathworld.wolfram.com/BivariateNormalDistribution.html>

Recursos de internet

Calculadora de ©Richard Lowry 2001-2019. All Rights Reserved: <http://vassarstats.net/rho.html>

CalculatorSoup®: <http://www.calculatorsoup.com/calculators/discretemathematics/combinations.php>

Coursera: <https://es.coursera.org/>

Gpower (Allgemeine Psychologie und Arbeitspsychologie): <http://www.gpower.hhu.de/>

Lisrel Scientific Software International (SSI): <http://www.ssicentral.com/lisrel/>

Mathcracker, Solvers Statistics (Critical z Values): https://mathcracker.com/z_critical_values.php

Minitab: <http://www.minitab.com/es-mx/>

Psychometrica: <https://www.psychometrica.de/correlation.html>

R (The R Project for Statistical Computing): <https://www.r-project.org/>

SAS (Analytics Platform): https://www.sas.com/en_us/home.html

Social Sciences Statistics (valores p de los puntajes t): <https://www.socscistatistics.com/pvalues/tdistribution.aspx>

----- (valores p de los puntajes z): <http://www.socscistatistics.com/pvalues/normal-distribution.aspx>

SPSS (IBM Software): <https://www.ibm.com/analytics/mx/es/technology/spss/index.html>

Wolfram MathWorld: <http://mathworld.wolfram.com/BivariateNormalDistribution.html>

UACJ

