

Exp

6^a
ed

Manual de
**Estadística,
Método diferencial
y Psicometría**

APIR

Dirección editorial López Frutos, Alicia
Campos Pavón, Jaime
Suárez Barrientos, Aida
Tajima Pozo, Kazuhiro

**MANUAL APIR
ESTADÍSTICA, MÉTODO EXPERIMENTAL Y PSICOMETRÍA
6.ª edición - Marzo 2024**

ISBN
978-84-19592-02-6

DEPÓSITO LEGAL
M-16636-2016

ACADEMIA DE PREPARACIÓN PIR, S.L.
www.academiapir.com
admin@academiapir.com

APIR es una marca registrada de la **Academia de Preparación PIR S.L.**

MAQUETACIÓN
Equipo de Diseño y Maquetación AMIR

DISEÑO E ILUSTRACIONES
Equipo de Diseño y Maquetación AMIR

La protección de los derechos de autor se extiende tanto al contenido redaccional de la publicación como al diseño, ilustraciones y fotografías de la misma, por lo que queda prohibida su reproducción total o parcial sin el permiso del propietario de los derechos de autor.



Este manual ha sido impreso con papel ecológico, sostenible y libre de cloro, y ha sido certificado según los estándares del FSC (*Forest Stewardship Council*) y del PEFC (*Programme for the Endorsement of Forest Certification*).

Estadística

Variable

Representación numérica de una característica con al menos dos modalidades (\neq constante: 1 modalidad). Las VD deben ser válidas, fiables y sensibles.

VARIABLES		
CUANTITATIVAS	SEMI-CUANTITATIVAS	CUALITATIVAS
Son de Intervalo o razón	Ordinales	Nominales
Subtipos: Discretas Continuas		Dicotómica si sólo puede tomar dos valores; p. ej., sí/no

Conceptos estadísticos habituales

Estadística

La estadística es una ciencia que recoge, ordena y analiza datos que provienen de muestras, y que realiza inferencias acerca de las poblaciones de las que estas muestras proceden.

- **Estadística descriptiva:**
Se ocupa de la ordenación y del análisis de datos. Describe las características de una muestra utilizando estadísticos muestrales (PIR).
- **Estadística inferencial:**
Se ocupa de la realización de inferencias a partir de datos recogidos en muestras sobre poblaciones. Es decir, describe a una población mediante los parámetros inferidos a partir de los datos de una muestra.

Población

Una población estadística reúne todos los elementos que cumplen una o varias características, por ejemplo, las mujeres de nacionalidad española, los hombres, etc. (PIR).

Muestra

Una muestra es un subconjunto de elementos de una población (p. ej., una parte de las mujeres de nacionalidad española).

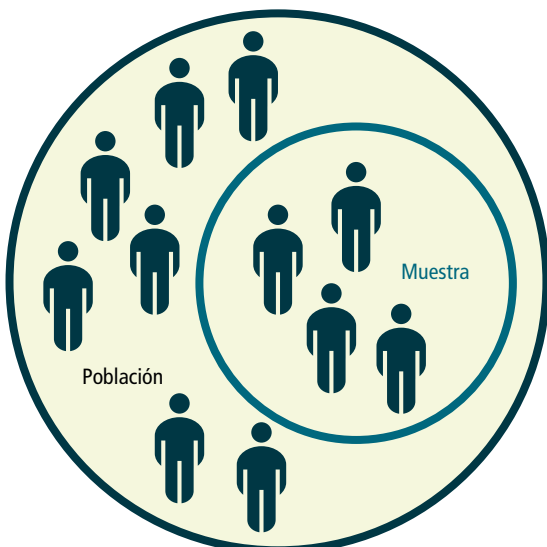


Figura 1. Diferencia entre población y muestra.

Parámetro

Un parámetro es una propiedad que describe una población. Salvo que trabajemos con todos los datos de una población los parámetros deberemos inferirlos de los datos obtenidos de la muestra (PIR).

Estadístico

Un estadístico es una propiedad que describe una muestra. Asimismo, la media de una escala de medición del estrés es un parámetro si se calcula a partir de una población (p. ej., todas las mujeres españolas), y es un estadístico si se calcula sobre una muestra (p. ej., parte de las mujeres españolas).

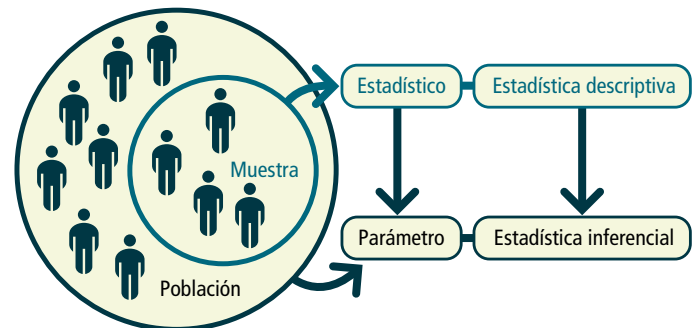


Figura 2. Representación visual de correspondencia entre estadística descriptiva e inferencial.

CONJUNTO O SUBCONJUNTO DE ELEMENTOS	PROPIEDAD DESCRIPTIVA	TIPO DE ESTADÍSTICA
Población	Parámetro	Inferencial
Muestra	Estadístico	Descriptiva

Tabla 2. Correspondencia entre estadística descriptiva e inferencial.

Las variables

Una variable es una representación numérica de una característica. Se pueden clasificar de diferentes maneras.

Medidas de variación

OTRAS MEDIDAS DE VARIACIÓN			
AMPLITUD (recorrido, rango)	AMPLITUD SEMI-INTERCUARTIL	DESVIACIÓN MEDIA	COEFICIENTE DE VARIACIÓN
Diferencia entre puntuación máxima y mínima.	Diferencia entre Q1 y Q3 entre 2.	Media de las diferencias, en valor absoluto, de las puntuaciones respecto a su media.	Dato que se obtiene de dividir la DT de cada grupo por su correspondiente media, y multiplicado por 100. Indica cómo varían las distancias en un grupo al comparlas con las distancias de otro grupo diferente.
Muy sensible a valores extremos y no toma en cuenta valores intermedios. Solo para variables cuantitativas (intervalo o razón).	Menos influenciable por valores extremos (adecuada para distribuciones asimétricas). Mnemotécnica: Amplitud Semi-Intercuartil (ASI) → distribuciones ASimétricas. Nos permite conocer en cuántas unidades de los valores que toma la variable se concentra en 50% de los casos.	Útil cuando el sumatorio de las diferencias es nulo. A menor desviación media, mayor homogeneidad de las puntuaciones (más próximos a la media).	Se usa cuando se quiere: <ul style="list-style-type: none"> • Comparar variabilidades muestrales de características de distinta naturaleza (p. ej., peso y altura). • Medir la misma característica en dos muestras de diferente naturaleza o de la misma naturaleza en diferentes grupos. Permite comparar la variabilidad de distintas distribuciones.

	DISTRIBUCIÓN	
	SIMÉTRICA	ASIMÉTRICA
TENDENCIA CENTRAL	Media (sensible a variaciones en las puntuaciones).	Mediana (menos sensible a variaciones de las puntuaciones).
VARIACIÓN	Varianza y desviación típica.	ASI.

$$CV = \frac{S}{\bar{X}} 100$$

Puntuaciones típicas y escalas derivadas (PIR 20, 100)

La puntuación típica

La **puntuación típica** (z_x) indica el número de desviaciones típicas que separa una observación de la media de su grupo (PIR). Es la calificación estándar más elemental y útil (PIR), ya que permite comparar unidades de distintos grupos, variables medidas de distintas formas o variables diferentes (PIR). Su fórmula es:

$$z_x = \frac{X - \bar{X}}{S_x}$$

Tomemos un ejemplo. Se sabe que la media de peso en Italia es de 70 kg, con una desviación típica de 15 kg. Ahora imaginemos un país hipotético, el Glorbistán, en el que la media de

peso es de 2456 glorbs, con una desviación típica de 348 glorbs. Durante sus vacaciones, Marco, de nacionalidad italiana, se va de viaje a este país, y quiere comparar su peso con el de un amigo Glorbistaní que se encuentra ahí. Marco pesa 84 kg, y su amigo 2600 glorbs (PIR).

En esta situación, en la que ni el uno ni el otro conocen la correspondencia entre las diferentes unidades de peso, la puntuación típica es de particular ayuda. Aplicando la fórmula, se encuentra que la puntuación típica de Marco es de 0.93 y la de su amigo Glorbistaní de 0.41. **Estos datos significan que Marco se encuentra 0.93 desviaciones típicas por encima de la media de peso italiana, y que su amigo se encuentra 0.41 desviación típica por encima de la media Glorbistaní** (o sea, que este último se encuentra más cerca de la media de peso de su país que Marco de la media de peso de Italia).

Una de las propiedades más destacables de las puntuaciones típicas es que su media es igual a 0, mientras que su varianza y desviación típica son iguales a 1 (PIR).

Los **eneatipos** o **estatinos** son un tipo de puntuación típica normalizada (PIR), de media 5 y desviación típica 2 (PIR 18, 56). Se llega a ella transformando las puntuaciones típicas normali-

Distribución de los datos

DISTRIBUCIÓN DE PROBABILIDAD PARA VARIABLES

DISCRETAS

Binomial:

la variable analizada tiene que ser dicotómica o dicotomizada (cara-cruz, acierto-error). Cuando en un ensayo sólo puede ocurrir uno de los dos resultados mutuamente excluyentes como éxito-fracaso. Una secuencia de ensayos bernoulli repetidos n veces.

Multinomial:

se utiliza al tener N ensayos independientes con más de dos resultados posibles.

Poisson:

con variables dicotómicas. Probabilidad de ocurrencia de sucesos con probabilidades muy pequeñas o sucesos "raros".

Geométrica:

se repite un experimento aleatorio en el que sólo dos resultados son posibles, la distribución geométrica indicaría el número de ensayos necesarios para obtener el suceso X .

CONTINUAS

Rectangular:

si todos los valores con probabilidad no nula tienen la misma función de densidad de probabilidad.

Normal:

se divide alrededor de la esperanza matemática en dos áreas simétricas de igual superficie. Puede ser normal no estandarizada (a partir de las variables observadas) o normal estandarizada (a partir de puntuaciones tipificadas). **Propiedades:**

- Simétrica y asintótica.
- Dos puntos de inflexión.
- El área representada supone una probabilidad de 1.
- Media, Q2 y P50 coinciden.

Chi cuadrado:

se construye a partir de la curva normal. El parámetro que caracteriza a esta distribución son sus grados de libertad ($n-1$), originando una distribución para cada grado de libertad. Asimétrica. Siempre valores positivos.

T de Student:

a partir de dos variables aleatorias X e Y . X se distribuye según una curva normal estandarizada e Y según la curva chi cuadrado. Es simétrica con respecto a valor 0. Cuando los grados de libertad tienden a infinito, la curva t de Student tiende a la curva normal.

Exponencial:

análoga a la distribución geométrica para variables aleatorias discretas. Para medir variables relacionadas con el tiempo y su percepción (p. ej., "tiempo que tarda un enfermo en manifestar los primeros síntomas").

Fisher:

se calcula a partir de la chi cuadrado y comparte muchas de sus características: asimétrica, valores siempre positivo.

la curva t de Student tiende a la curva normal (PIR 15, 05; PIR 14, 25).

La distribución exponencial

Esta distribución es análoga a la distribución geométrica para variables aleatorias discretas. Se utiliza en general para medir variables relacionadas con el tiempo y su percepción (p. ej., serviría para representar la variable "tiempo que tarda un enfermo en manifestar los primeros síntomas tras una infección" (PIR).

La distribución de Fischer (o distribución de F de Snedecor)

Esta distribución se calcula a partir de la distribución χ^2 , y comparte por lo tanto muchas de sus características (p. ej., su curva es asimétrica, sus valores siempre positivos). Al igual que la distribución χ^2 , dicha asimetría va disminuyendo según aumentan los grados de libertad, y tiende a la normal cuando los grados de libertad tienden a infinito.

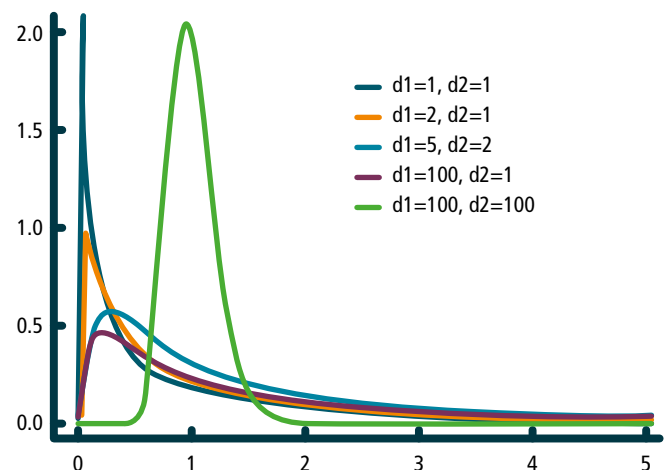


Figura 20. Distribución de Fischer.

La distribución muestral y el estadístico de contraste

Antes de describir los resultados obtenidos, los estudios suelen presentar el método de extracción de datos, o sea el tipo de muestreo utilizado, así como describir la distribución de la muestra (si la muestra es grande, se supone que su distribución es normal).

Los estadísticos de contraste, como lo hemos visto anteriormente, tienen distribuciones específicas (p. ej., la prueba χ^2 según una curva que le es específica, el estadístico "z" según la curva normal, etc.) cuyos valores se encuentran generalmente en los manuales de estadística. A partir de estos valores, se puede calcular el valor del α y definir el **nivel crítico**.

La región crítica

La región crítica define la probabilidad de rechazar la hipótesis nula (o sea, el valor de α) (PIR 15, 07; PIR). De manera complementaria, define igualmente la probabilidad de aceptación de la hipótesis nula ($1-\alpha$) (PIR).

Gráficamente, si se representan los posibles valores del estadístico de contraste en un segmento, se delimitan dos regiones:

- La región crítica o región de rechazo, delimitada por el valor de α .
- La región no crítica que define la probabilidad de mantener la hipótesis nula.

Existen tres tipos de contrastes estadísticos, que definen simultáneamente tres tipos de regiones críticas (PIR):

• **Bilateral.**

Este caso se presenta cuando la hipótesis plantea una igualdad y la hipótesis alternativa una desigualdad (PIR). Por ejemplo:

- $H_0: \mu=0,7$
- $H_1: \mu \neq 0,7$
- En este caso, es necesario dividir el valor de α entre 2, porque existe una probabilidad que el resultado encontrado sea superior o inferior, en el caso de nuestro ejemplo, a 0.7.

(Ver figura 21)

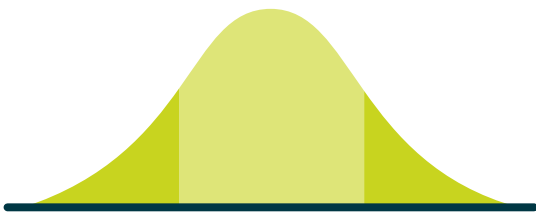


Figura 21. Contraste bilateral.

• **Unilateral izquierdo.**

Se dice que un contraste es unilateral izquierdo cuando se rechaza la hipótesis nula al obtener un valor del estadístico de contraste menor al punto crítico inferior (PIR). Los contrastes unilaterales izquierdos se formulan matemáticamente de la siguiente manera:

- $H_0: \mu \geq 0,7$
- $H_1: \mu < 0,7$

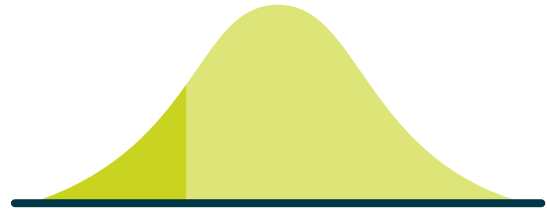


Figura 22. Contraste unilateral izquierdo.

• **Unilateral derecho.**

Se dice que un contraste es unilateral derecho cuando se rechaza la hipótesis nula al obtener un valor del estadístico de contraste superior al punto crítico superior. Los contrastes unilaterales derechos se formulan matemáticamente de la siguiente manera:

- $H_0: \mu \leq 0,7$
- $H_1: \mu > 0,7$

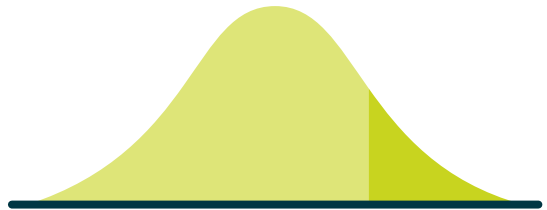


Figura 23. Contraste unilateral derecho.

Una hipótesis nula que ha sido rechazada en un contraste unilateral, a un determinado nivel de confianza, puede no ser rechazada a ese mismo nivel de confianza si el contraste es bilateral (PIR).

Mantenimiento o rechazo de la hipótesis nula

La decisión de mantener o rechazar una hipótesis nula es el fruto de una comparación entre el valor fijado de α y el nivel crítico (p). **Se define éste como el nivel de significación más pequeño al que una hipótesis nula puede ser rechazada con el estadístico de contraste obtenido (PIR)**. También se define como la probabilidad de obtener un resultado para un estadístico de contraste tan o más extremo que el encontrado en nuestra muestra, siendo H_0 verdadera (PIR 16, 43).

Si $p < \alpha$, entonces la probabilidad de cometer un error de tipo I real es inferior al riesgo que hemos fijado a priori (p. ej., decidiendo de fijar el nivel de significación a $\alpha=0.05$), y se rechaza por lo tanto la hipótesis nula.

Si $p > \alpha$, entonces la probabilidad de cometer un error de tipo I real es superior al riesgo que hemos fijado a priori, y se mantiene por lo tanto la hipótesis nula.

Relaciones existentes entre los números para medir

PISO



TIPOS DE RELACIONES:

PROPORCIONALES

Se da en características de los objetos en las que además de igualdad, orden, suma y resta podemos establecer operaciones de división y multiplicación.

IGUALDAD/DESIGUALDAD

Se da en características de los objetos en las que sólo podemos decir que son iguales o diferentes. $A=B$ ó $A \neq B$.

SUMA Y RESTA

Se da en características de los objetos en las que además de igualdad y orden podemos realizar operaciones de suma y resta. $A+B=C$, $C-B=A$.

ORDEN

Se da en características de los objetos en las que además de igualdad podemos establecer si unos son mayores que otros, es decir, ordenarlos. $A > B > C$ ó $C < B < A$.

COEFICIENTE DE SPEARMAN

Coefficiente que ÚNICAMENTE se utiliza para variables ordinales

hay que guardar el ORDEN en la cola para el concierto de BRITNEY SPEARS



Escala de medida

La definición de una escala de medida, nivel de medida o modelo de medida es una plantilla o un patrón para realizar la medición en el que a cada una de las modalidades de las características de los fenómenos se le asigna un número según una relación biunívoca.

tienes RAZÓN cuando INTENTas ORDENar tu Nevera

TIPOS DE ESCALAS

RAZÓN

INTERVALO

ORDINAL

NOMINAL



OTROS COEFICIENTES DE CORRELACIÓN

FILOMENA, te pones la VISERA y PUNTO.
Te pones la VISERA y te dejas de TRETAS.

EN ORDEN:

PHI



2 V Dicotómicas

BISERIAL PUNTUAL



1 V Dicotómica +
1 V Continua

BISERIAL



1 V Continua +
1 V Dicotomizada

TRETACÓRICO



2 V Dicotomizadas

Se calcula un índice de fiabilidad: Índice de KAPPA

Gorra en inglés: CAP

