# Introducción a Modelos Psicométricos

## Clase 2: Herramientas estadísticas para la psicometría

Iwin Leenen y Ramsés Vázquez-Lira

Facultad de Psicología, UNAM

Programa de Licenciatura y Posgrado en Psicología
Semestre 2019–1

# Índice

# Índice

# Índice

# Objetos, variables y observaciones

En la estadística aplicada, se dispone de los datos de varios objetos o unidades experimentales.

Por ejemplo:

- Hemos registrado la interacción social de 50 niños en el patio de la escuela.
- Tenemos las respuestas en un cuestionario de 60 participantes en un taller de *mindfulness*.
- Tenemos los resultados de 1,000 alumnos en el examen de entrada de la UNAM.

Para hacer referencia a las unidades experimentales, se suele asignarles un número. Por ejemplo, se habla del resultado del "alumno *i*" en el examen.

# Objetos, variables y observaciones

En la estadística aplicada, se dispone de los datos de varios objetos o unidades experimentales.

Por ejemplo:

- Hemos registrado la interacción social de 50 niños en el patio de la escuela.
- Tenemos las respuestas en un cuestionario de 60 participantes en un taller de *mindfulness*.
- Tenemos los resultados de 1,000 alumnos en el examen de entrada de la UNAM.

Para hacer referencia a las unidades experimentales, se suele asignarles un número. Por ejemplo, se habla del resultado del “alumno $i$” en el examen.

# Objetos, variables y observaciones

En la estadística aplicada, se dispone de los datos de varios objetos o unidades experimentales.

Por ejemplo:

- Hemos registrado la interacción social de 50 niños en el patio de la escuela.
- Tenemos las respuestas en un cuestionario de 60 participantes en un taller de *mindfulness*.
- Tenemos los resultados de 1,000 alumnos en el examen de entrada de la UNAM.

Para hacer referencia a las unidades experimentales, se suele asignarles un número. Por ejemplo, se habla del resultado del “alumno $i$” en el examen.

# Objetos, variables y observaciones

Para poder sistematizar los datos, se definen una o más variables.
Es común representar las variables por letras mayúsculas $X$, $Y$, etc.

Por ejemplo:

- Para el registro de la interacción social, se define la variable $X$: "Número de otros niños con los que habla".
- Para evaluar el efecto del taller, se definen las variables:
  - $X_{\text{pre}}$: "Puntuación en la aplicación *pre*"
  - $X_{\text{post}}$: "Puntuación en la aplicación *post*"
  - $Y$: "Diferencia entre las puntuación *pre* y *post*".
- Para el examen de opción múltiple, se definen las variables $X_1$ a $X_{120}$: "Puntuación en la pregunta $j$ del examen".

# Objetos, variables y observaciones

Para poder sistematizar los datos, se definen una o más variables.
Es común representar las variables por letras mayúsculas $X$, $Y$, etc.

Por ejemplo:

- Para el registro de la interacción social, se define la variable $X$: "Número de otros niños con los que habla".
- Para evaluar el efecto del taller, se definen las variables:
  - $X_{\text{pre}}$: "Puntuación en la aplicación *pre*"
  - $X_{\text{post}}$: "Puntuación en la aplicación *post*"
  - $Y$: "Diferencia entre las puntuación *pre* y *post*".
- Para el examen de opción múltiple, se definen las variables $X_1$ a $X_{120}$: "Puntuación en la pregunta $j$ del examen".

# Objetos, variables y observaciones

Se observan para los objetos valores en las variables.

Por ejemplo:

- El niño $i$ tuvo el valor de 8 en la variable $X$ "Número de de otros niños con los que habla":
$$x_i = 8.$$
- El alumno $i$ contestó correctamente en la pregunta $j$ del examen de opción múltiple.
Es decir, tiene el valor 1 en la variable $X_j$:
$$x_{ij} = 1.$$

# Objetos, variables y observaciones

Se observan para los objetos valores en las variables.

Por ejemplo:

- El niño *i* tuvo el valor de 8 en la variable *X* "Número de de otros niños con los que habla":
$$x_i = 8.$$
- El alumno *i* contestó correctamente en la pregunta *j* del examen de opción múltiple.
  Es decir, tiene el valor 1 en la variable $X_j$:
$$x_{ij} = 1.$$

# Índice

# El signo sumatorio Σ (Sigma)

¿Qué significa la siguiente expresión?

$$\sum_{k=3}^{6} \frac{k^2}{2} = ?$$

# El signo sumatorio Σ (Sigma)

¿Qué significa la siguiente expresión?

$$\sum_{k=3}^{6} \frac{k^2}{2} = \underbrace{\frac{3^2}{2}}_{k=3} + \underbrace{\frac{4^2}{2}}_{k=4} + \underbrace{\frac{5^2}{2}}_{k=5} + \underbrace{\frac{6^2}{2}}_{k=6} = 43$$

- índice sumatorio: $k$
- término genérico: $\frac{k^2}{2}$
- límites: 3 (límite inferior) y 6 (límite superior)

# El signo sumatorio Σ (Sigma)

Otro ejemplo:

$$\sum_{i=1}^{n} x_i = ?$$

# El signo sumatorio Σ (Sigma)

Otro ejemplo:

$$\sum_{i=1}^{n} x_i = \underbrace{x_1}_{i=1} + \underbrace{x_2}_{i=2} + \underbrace{x_3}_{i=3} + \ldots + \underbrace{x_n}_{i=n}$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

"Sumas dobles":

$$\sum_{i=2}^{4}\sum_{j=3}^{5} x_{ij} = ?$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

“Sumas dobles”:

$$\sum_{i=2}^{4}\sum_{j=3}^{5} x_{ij} = \sum_{i=2}^{4}\left(\sum_{j=3}^{5} x_{ij}\right)$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

"Sumas dobles":

$$\sum_{i=2}^{4}\sum_{j=3}^{5} x_{ij} = \sum_{i=2}^{4}\left(\sum_{j=3}^{5} x_{ij}\right)$$

$$= \underbrace{\sum_{j=3}^{5} x_{2j}}_{i=2} + \underbrace{\sum_{j=3}^{5} x_{3j}}_{i=3} + \underbrace{\sum_{j=3}^{5} x_{4j}}_{i=4}$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

"Sumas dobles":

$$
\begin{aligned}
\sum_{i=2}^{4}\sum_{j=3}^{5} x_{ij} &= \sum_{i=2}^{4}\left(\sum_{j=3}^{5} x_{ij}\right) \\
&= \underbrace{\sum_{j=3}^{5} x_{2j}}_{i=2} + \underbrace{\sum_{j=3}^{5} x_{3j}}_{i=3} + \underbrace{\sum_{j=3}^{5} x_{4j}}_{i=4} \\
&= (x_{23} + x_{24} + x_{25}) + (x_{33} + x_{34} + x_{35}) + (x_{43} + x_{44} + x_{45})
\end{aligned}
$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

En una suma doble, el límite inferior o superior (o ambos) del segundo signo sumatorio puede contener el índice del primer signo sumatorio

$$\sum_{i=1}^{4} \sum_{j=i+1}^{4} (x_i - x_j)^2 = ?$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

En una suma doble, el límite inferior o superior (o ambos) del segundo signo sumatorio puede contener el índice del primer signo sumatorio

$$\sum_{i=1}^{4}\sum_{j=i+1}^{4}(x_i - x_j)^2 = \sum_{i=1}^{4}\left[\sum_{j=i+1}^{4}(x_i - x_j)^2\right]$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

En una suma doble, el límite inferior o superior (o ambos) del segundo signo sumatorio puede contener el índice del primer signo sumatorio

$$\sum_{i=1}^{4}\sum_{j=i+1}^{4}(x_i - x_j)^2 = \sum_{i=1}^{4}\left[\sum_{j=i+1}^{4}(x_i - x_j)^2\right]$$

$$= \underbrace{\sum_{j=1+1}^{4}(x_1 - x_j)^2}_{i=1} + \underbrace{\sum_{j=2+1}^{4}(x_2 - x_j)^2}_{i=2} + \underbrace{\sum_{j=3+1}^{4}(x_3 - x_j)^2}_{i=3} + \underbrace{\sum_{j=4+1}^{4}(x_4 - x_j)^2}_{i=4}$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

En una suma doble, el límite inferior o superior (o ambos) del segundo signo sumatorio puede contener el índice del primer signo sumatorio

$$\sum_{i=1}^{4}\sum_{j=i+1}^{4}(x_i-x_j)^2=\sum_{i=1}^{4}\left[\sum_{j=i+1}^{4}(x_i-x_j)^2\right]$$

$$=\underbrace{\sum_{j=1+1}^{4}(x_1-x_j)^2}_{i=1}+\underbrace{\sum_{j=2+1}^{4}(x_2-x_j)^2}_{i=2}+\underbrace{\sum_{j=3+1}^{4}(x_3-x_j)^2}_{i=3}+\underbrace{\sum_{j=4+1}^{4}(x_4-x_j)^2}_{i=4}$$

$$=\sum_{j=2}^{4}(x_1-x_j)^2+\sum_{j=3}^{4}(x_2-x_j)^2+\sum_{j=4}^{4}(x_3-x_j)^2+\sum_{j=5}^{4}(x_4-x_j)^2$$

# El signo sumatorio Σ (Sigma)

## Sumas dobles

En una suma doble, el límite inferior o superior (o ambos) del segundo signo sumatorio puede contener el índice del primer signo sumatorio

$$\sum_{i=1}^{4}\sum_{j=i+1}^{4}(x_i - x_j)^2 = \sum_{i=1}^{4}\left[\sum_{j=i+1}^{4}(x_i - x_j)^2\right]$$

$$= \underbrace{\sum_{j=1+1}^{4}(x_1 - x_j)^2}_{i=1} + \underbrace{\sum_{j=2+1}^{4}(x_2 - x_j)^2}_{i=2} + \underbrace{\sum_{j=3+1}^{4}(x_3 - x_j)^2}_{i=3} + \underbrace{\sum_{j=4+1}^{4}(x_4 - x_j)^2}_{i=4}$$

$$= \sum_{j=2}^{4}(x_1 - x_j)^2 + \sum_{j=3}^{4}(x_2 - x_j)^2 + \sum_{j=4}^{4}(x_3 - x_j)^2 + \sum_{j=5}^{4}(x_4 - x_j)^2$$

$$= \left[(x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_1 - x_4)^2\right] + \left[(x_2 - x_3)^2 + (x_2 - x_4)^2\right] + \left[(x_3 - x_4)^2\right]$$

# Índice

# Índice

# Distribuciones univariadas empíricas

Datos observados de una muestra

Supongamos que hemos registrado de 50 pacientes el índice de masa corporal (IMC):

| No. | IMC | No. | IMC | No. | IMC | No. | IMC | No. | IMC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 25 | 11 | 25 | 21 | 42 | 31 | 26 | 41 | 29 |
| 2 | 25 | 12 | 21 | 22 | 27 | 32 | 27 | 42 | 28 |
| 3 | 24 | 13 | 22 | 23 | 22 | 33 | 24 | 43 | 23 |
| 4 | 20 | 14 | 30 | 24 | 25 | 34 | 33 | 44 | 27 |
| 5 | 22 | 15 | 26 | 25 | 34 | 35 | 24 | 45 | 27 |
| 6 | 23 | 16 | 28 | 26 | 22 | 36 | 22 | 46 | 19 |
| 7 | 24 | 17 | 23 | 27 | 22 | 37 | 23 | 47 | 24 |
| 8 | 26 | 18 | 20 | 28 | 24 | 38 | 29 | 48 | 23 |
| 9 | 20 | 19 | 19 | 29 | 26 | 39 | 24 | 49 | 26 |
| 10 | 25 | 20 | 28 | 30 | 22 | 40 | 23 | 50 | 27 |

# Distribuciones univariadas empíricas

## Datos observados de una muestra

Supongamos que hemos registrado de 50 pacientes el índice de masa corporal (IMC):

| No. | IMC | No. | IMC | No. | IMC | No. | IMC | No. | IMC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 25 | 11 | 25 | 21 | 42 | 31 | 26 | 41 | 29 |
| 2 | 25 | 12 | 21 | 22 | 27 | 32 | 27 | 42 | 28 |
| 3 | 24 | 13 | 22 | 23 | 22 | 33 | 24 | 43 | 23 |
| 4 | 20 | 14 | 30 | 24 | 25 | 34 | 33 | 44 | 27 |
| 5 | 22 | 15 | 26 | 25 | 34 | 35 | 24 | 45 | 27 |
| 6 | 23 | 16 | 28 | 26 | 22 | 36 | 22 | 46 | 19 |
| 7 | 24 | 17 | 23 | 27 | 22 | 37 | 23 | 47 | 24 |
| 8 | 26 | 18 | 20 | 28 | 24 | 38 | 29 | 48 | 23 |
| 9 | 20 | 19 | 19 | 29 | 26 | 39 | 24 | 49 | 26 |
| 10 | 25 | 20 | 28 | 30 | 22 | 40 | 23 | 50 | 27 |

$frec(25) = 5$

# Distribuciones univariadas empíricas

## Datos observados de una muestra

Supongamos que hemos registrado de 50 pacientes el índice de masa corporal (IMC):

| No. | IMC | No. | IMC | No. | IMC | No. | IMC | No. | IMC |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 25 | 11 | 25 | 21 | 42 | 31 | 26 | 41 | 29 |
| 2 | 25 | 12 | 21 | 22 | 27 | 32 | 27 | 42 | 28 |
| 3 | 24 | 13 | 22 | 23 | 22 | 33 | 24 | 43 | 23 |
| 4 | 20 | 14 | 30 | 24 | 25 | 34 | 33 | 44 | 27 |
| 5 | 22 | 15 | 26 | 25 | 34 | 35 | 24 | 45 | 27 |
| 6 | 23 | 16 | 28 | 26 | 22 | 36 | 22 | 46 | 19 |
| 7 | 24 | 17 | 23 | 27 | 22 | 37 | 23 | 47 | 24 |
| 8 | 26 | 18 | 20 | 28 | 24 | 38 | 29 | 48 | 23 |
| 9 | 20 | 19 | 19 | 29 | 26 | 39 | 24 | 49 | 26 |
| 10 | 25 | 20 | 28 | 30 | 22 | 40 | 23 | 50 | 27 |

$frec(25) = 5$

$p(29) = \frac{2}{50} = .04$

# Distribuciones univariadas empíricas

## Función de frecuencia y proporción

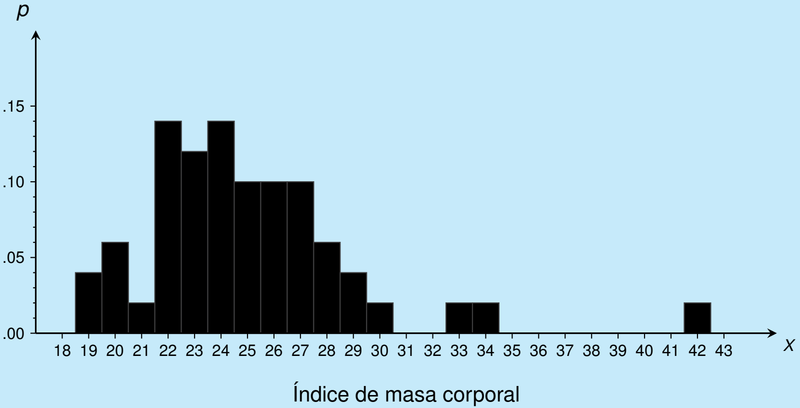
Se puede resumir la distribución de la variable $X$ (IMC) en una tabla, que muestra la función de frecuencia y la función de proporción:

| $j$ | $x_j$ | $frec(x_j)$ | $p(x_j)$ |
|---|---|---|---|
| 1 | 19 | 2 | .04 |
| 2 | 20 | 3 | .06 |
| 3 | 21 | 1 | .02 |
| 4 | 22 | 7 | .14 |
| 5 | 23 | 6 | .12 |
| 6 | 24 | 7 | .14 |
| 7 | 25 | 5 | .10 |
| 8 | 26 | 5 | .10 |
| 9 | 27 | 5 | .10 |
| 10 | 28 | 3 | .06 |
| 11 | 29 | 2 | .04 |
| 12 | 30 | 1 | .02 |
| 13 | 33 | 1 | .02 |
| 14 | 34 | 1 | .02 |
| 15 | 42 | 1 | .02 |
| Total | | 50 | 1.00 |

# Distribuciones univariadas empíricas

## Función de frecuencia y proporción

Se puede representar gráficamente la función de frecuencia o proporción:



# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.
- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.
- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \ldots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.
- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.

- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.

- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \dots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.

- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.
- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.
- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \ldots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.
- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.

- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.

- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \ldots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.

- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.
- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.
- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \ldots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.
- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.
- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.
- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \ldots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.
- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Modelo estadístico

Supongamos que:

- Se administra un examen de física cuántica que consiste en 10 preguntas de opción múltiple con tres opciones de respuesta.
  La puntuación $X$ en el examen es el número de respuestas correctas.
- Una persona que no sabe nada de física cuántica lo responde.
  Elige en cada pregunta aleatoriamente una respuesta.
- ¿Cuál es la probabilidad de que esta persona obtenga un puntaje de $x$ ($x = 0, 1, 2, \ldots, 10$) en el examen?

Para contestar esta pregunta, se especifica un modelo estadístico, donde, por ejemplo, se supone que:

- La probabilidad de elegir la respuesta correcta en cada pregunta es $\frac{1}{3}$.
- Las respuestas son independientes.

Lo anterior, estadísticamente, se traduce diciendo que $X$ tiene una distribución binomial:

$$X \sim \text{Binomial}\left(10, \frac{1}{3}\right).$$

# Distribuciones univariadas teóricas para variables discretas

## Función de probabilidad

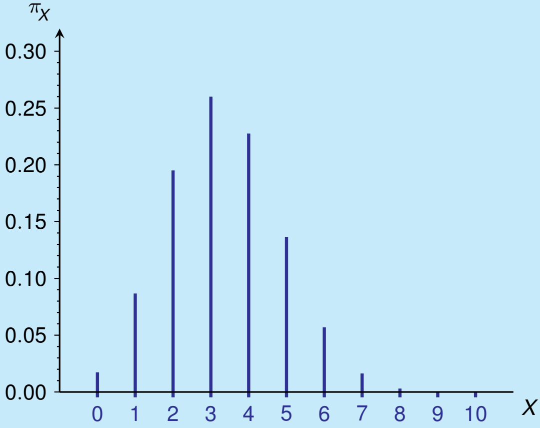
Esta distribución teórica de la variable $X$ (puntaje en el examen) se puede resumir en una tabla, mostrando la función de probabilidad de $X$:

| $j$ | $x_j$ | $\pi_X(x_j)$ |
|---|---|---|
| 1 | 0 | .017 |
| 2 | 1 | .087 |
| 3 | 2 | .195 |
| 4 | 3 | .260 |
| 5 | 4 | .228 |
| 6 | 5 | .137 |
| 7 | 6 | .057 |
| 8 | 7 | .016 |
| 9 | 8 | .003 |
| 10 | 9 | .000 |
| 11 | 10 | .000 |
| Total | | 1.000 |

# Distribuciones univariadas teóricas para variables discretas

## Función de probabilidad

Y se puede representar gráficamente:



# Distribuciones univariadas teóricas para variables continuas

## Modelo estadístico y función de densidad

Supongamos que:

- Se aplica un examen en la computadora y, para una pregunta concreta, se registra el tiempo $X$ (en segundos) entre el momento en que se muestra la pregunta en la pantalla y el momento de la selección de la respuesta;
- La variable $X$ tiene la siguiente distribución:

# Distribuciones univariadas teóricas para variables continuas

## Modelo estadístico y función de densidad

Supongamos que:

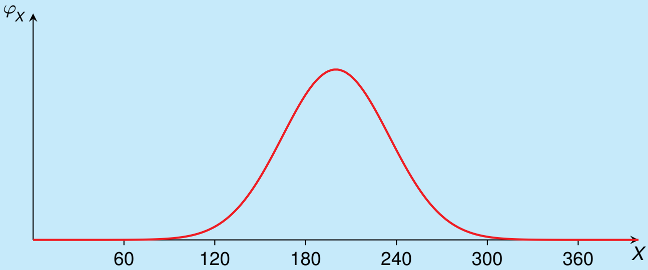
- Se aplica un examen en la computadora y, para una pregunta concreta, se registra el tiempo $X$ (en segundos) entre el momento en que se muestra la pregunta en la pantalla y el momento de la selección de la respuesta;
- La variable $X$ tiene la siguiente distribución:

# Distribuciones univariadas teóricas para variables continuas

## Modelo estadístico y función de densidad

Supongamos que:

- Se aplica un examen en la computadora y, para una pregunta concreta, se registra el tiempo $X$ (en segundos) entre el momento en que se muestra la pregunta en la pantalla y el momento de la selección de la respuesta;
- La variable $X$ tiene la siguiente distribución:

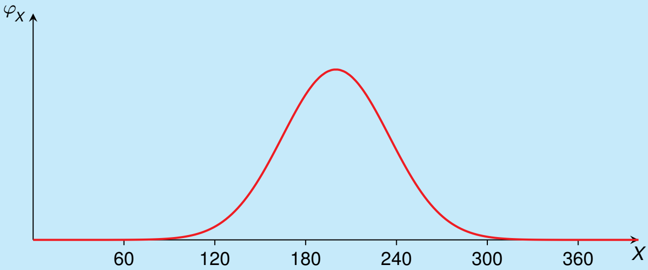


# Distribuciones univariadas teóricas para variables continuas

## Modelo estadístico y función de densidad

Supongamos que:

- Se aplica un examen en la computadora y, para una pregunta concreta, se registra el tiempo $X$ (en segundos) entre el momento en que se muestra la pregunta en la pantalla y el momento de la selección de la respuesta;
- La variable $X$ tiene la siguiente distribución:




Nota: $\varphi_X$ se llama la función de densidad de la variable $X$.

# Índice

# Estadísticos y parámetros de tendencia central

## La media aritmética y el valor esperado

- Para distribuciones empíricas (la variable $X$ observada en una muestra), se define la media aritmética:

$$\overline{x} = \sum_{j=1}^{m} p(x_j) x_j,$$

donde la suma es entre todos los *distintos* valores que se han observado en la muestra.

- Para distribuciones teóricas, se define la media poblacional o el valor esperado.
  - Para variables discretas:

$$\mathscr{E}(X) = \sum_{j=1}^{m} \pi_X(x_j)\, x_j,$$

donde la suma es entre todos los posibles valores que puede asumir la variable $X$.

  - Para variables continuas:

$$\mathscr{E}(X) = \int_{-\infty}^{+\infty} \varphi_X(x)\, x\, dx.$$

# Estadísticos y parámetros de tendencia central

## La media aritmética y el valor esperado

- Para distribuciones empíricas (la variable $X$ observada en una muestra), se define la media aritmética:

$$\overline{x} = \sum_{j=1}^{m} p(x_j) x_j ,$$

donde la suma es entre todos los *distintos* valores que se han observado en la muestra.

- Para distribuciones teóricas, se define la media poblacional o el valor esperado.
  - Para variables discretas:

$$\mathscr{E}(X) \;=\; \sum_{j=1}^{m} \pi_X(x_j)\, x_j ,$$

donde la suma es entre todos los posibles valores que puede asumir la variable $X$.

  - Para variables continuas:

$$\mathscr{E}(X) \;=\; \int_{-\infty}^{+\infty} \varphi_X(x)\, x\, dx .$$

# Estadísticos y parámetros de tendencia central

## La media aritmética y el valor esperado

- Para distribuciones empíricas (la variable $X$ observada en una muestra), se define la media aritmética:

  $$\overline{x} = \sum_{j=1}^{m} p(x_j)x_j,$$

  donde la suma es entre todos los *distintos* valores que se han observado en la muestra.

- Para distribuciones teóricas, se define la media poblacional o el valor esperado.
  - Para variables discretas:

    $$\mathscr{E}(X) \;=\; \sum_{j=1}^{m} \pi_X(x_j)\, x_j,$$

    donde la suma es entre todos los posibles valores que puede asumir la variable $X$.

  - Para variables continuas:

    $$\mathscr{E}(X) \;=\; \int_{-\infty}^{+\infty} \varphi_X(x)\, x\, dx.$$

# Estadísticos y parámetros de tendencia central

## Efecto de una transformación lineal en la media

- Definimos la nueva variable $Y$ a partir de la variable $X$:

$$Y = aX + b,$$

donde $a$ y $b$ son dos constantes cualesquiera.

- Entonces, la media aritmética de la variable $Y$ se obtiene por:

$$\overline{y} = a\overline{x} + b.$$

- Entonces, la media poblacional de la variable $Y$ se obtiene por:

$$\mathscr{E}(Y) = a\mathscr{E}(X) + b.$$

# Estadísticos y parámetros de tendencia central

## Efecto de una transformación lineal en la media

- Definimos la nueva variable $Y$ a partir de la variable $X$:

$$Y = aX + b,$$

donde $a$ y $b$ son dos constantes cualesquiera.

- Entonces, la media aritmética de la variable $Y$ se obtiene por:

$$\bar{y} = a\bar{x} + b.$$

- Entonces, la media poblacional de la variable $Y$ se obtiene por:

$$\mathscr{E}(Y) = a\mathscr{E}(X) + b.$$

# Estadísticos y parámetros de tendencia central

## Efecto de una transformación lineal en la media

- Definimos la nueva variable $Y$ a partir de la variable $X$:

$$Y = aX + b,$$

donde $a$ y $b$ son dos constantes cualesquiera.

- Entonces, la media aritmética de la variable $Y$ se obtiene por:

$$\overline{y} = a\overline{x} + b.$$

- Entonces, la media poblacional de la variable $Y$ se obtiene por:

$$\mathscr{E}(Y) = a\mathscr{E}(X) + b.$$

# Estadísticos y parámetros de tendencia central

## La media de una suma de variables

- Definimos la nueva variable $Y$ a partir de las variable $X_1$ y $X_2$ como:

$$Y = X_1 + X_2.$$

- Entonces, la media aritmética de la variable $Y$ se obtiene por:

$$\overline{y} = \overline{x}_1 + \overline{x}_2.$$

- Entonces, la media poblacional de la variable $Y$ se obtiene por:

$$\mathscr{E}(Y) = \mathscr{E}(X_1) + \mathscr{E}(X_2).$$

# Estadísticos y parámetros de tendencia central

## La media de una suma de variables

- Definimos la nueva variable $Y$ a partir de las variable $X_1$ y $X_2$ como:

$$Y = X_1 + X_2.$$

- Entonces, la media aritmética de la variable $Y$ se obtiene por:

$$\overline{y} = \overline{x}_1 + \overline{x}_2.$$

- Entonces, la media poblacional de la variable $Y$ se obtiene por:

$$\mathscr{E}(Y) = \mathscr{E}(X_1) + \mathscr{E}(X_2).$$

# Estadísticos y parámetros de tendencia central

## La media de una suma de variables

- Definimos la nueva variable $Y$ a partir de las variable $X_1$ y $X_2$ como:

$$Y = X_1 + X_2.$$

- Entonces, la media aritmética de la variable $Y$ se obtiene por:

$$\overline{y} = \overline{x}_1 + \overline{x}_2.$$

- Entonces, la media poblacional de la variable $Y$ se obtiene por:

$$\mathscr{E}(Y) = \mathscr{E}(X_1) + \mathscr{E}(X_2).$$

# Índice

# Estadísticos y parámetros de variabilidad o dispersión

La varianza

- Para distribuciones empíricas, se define la varianza muestral:

$$s_X^2 = \sum_{j=1}^{m} p(x_j)(x_j - \overline{x})^2,$$

donde la suma es entre todos los *distintos* valores que se han observado en la muestra.

- Para distribuciones teóricas, se define la varianza poblacional como:

$$\sigma_X^2 = \mathscr{E}\left[(X - \mu_X)^2\right],$$

donde $\mu_X$ es $\mathscr{E}(X)$.

  - Para variables discretas:

$$\sigma_X^2 = \sum_{j=1}^{m} \pi_X(x_j)\,(x_j - \mu_X)^2$$

  - para variables continuas:

$$\sigma_X^2 = \int_{-\infty}^{+\infty} \varphi_X(x)\,(x - \mu_X)^2\,dx$$

# Estadísticos y parámetros de variabilidad o dispersión

La varianza

- Para distribuciones empíricas, se define la varianza muestral:

$$s_X^2 = \sum_{j=1}^{m} p(x_j)(x_j - \overline{x})^2,$$

donde la suma es entre todos los *distintos* valores que se han observado en la muestra.

- Para distribuciones teóricas, se define la varianza poblacional como:

$$\sigma_X^2 = \mathscr{E}\left[(X - \mu_X)^2\right],$$

donde $\mu_X$ es $\mathscr{E}(X)$.

- Para variables discretas:

$$\sigma_X^2 = \sum_{j=1}^{m} \pi_X(x_j)\,(x_j - \mu_X)^2$$

- para variables continuas:

$$\sigma_X^2 = \int_{-\infty}^{+\infty} \varphi_X(x)\,(x - \mu_X)^2\,dx$$

# Estadísticos y parámetros de variabilidad o dispersión

## La varianza

- Para distribuciones empíricas, se define la varianza muestral:

$$s_X^2 = \sum_{j=1}^{m} p(x_j)(x_j - \overline{x})^2,$$

donde la suma es entre todos los *distintos* valores que se han observado en la muestra.

- Para distribuciones teóricas, se define la varianza poblacional como:

$$\sigma_X^2 = \mathscr{E}\left[(X - \mu_X)^2\right],$$

donde $\mu_X$ es $\mathscr{E}(X)$.

  - Para variables discretas:

$$\sigma_X^2 \ = \ \sum_{j=1}^{m} \pi_X(x_j)\,(x_j - \mu_X)^2$$

  - para variables continuas:

$$\sigma_X^2 \ = \ \int_{-\infty}^{+\infty} \varphi_X(x)\,(x - \mu_X)^2\,dx$$

# Estadísticos y parámetros de variabilidad o dispersión

## La desviación estándar

La desviación estándar es la raiz cuadrada de la varianza:

- Para distribuciones empíricas:

$$s_X = \sqrt{s_X^2}$$

- Para distribuciones teóricas:

$$\sigma_X = \sqrt{\sigma_X^2}$$

# Estadísticos y parámetros de variabilidad

## Efecto de una transformación lineal en la varianza y desviación estándar

Si la variable $Y$ se obtiene por una transformación lineal de la variable $X$, es decir $Y = aX + b$, entonces:

- la varianza y desviación estándar muestral de la variable $Y$ se obtiene por:

$$s_Y^2 = a^2\, s_X^2$$
$$s_Y = |a|\, s_X$$

- la varianza y desviación estándar poblacional de la variable $Y$ se obtiene por:

$$\sigma_Y^2 = a^2\, \sigma_X^2$$
$$\sigma_Y = |a|\, \sigma_X$$

# Estadísticos y parámetros de variabilidad

## Efecto de una transformación lineal en la varianza y desviación estándar

Si la variable $Y$ se obtiene por una transformación lineal de la variable $X$, es decir $Y = aX + b$, entonces:

- la varianza y desviación estándar muestral de la variable $Y$ se obtiene por:

$$s_Y^2 = a^2\, s_X^2$$
$$s_Y = |a|\, s_X$$

- la varianza y desviación estándar poblacional de la variable $Y$ se obtiene por:

$$\sigma_Y^2 = a^2\, \sigma_X^2$$
$$\sigma_Y = |a|\, \sigma_X$$

# Índice

# Función de frecuencia y proporción acumulada

Definición

## Definición

Para una variable $X$, que asume $m$ distintos valores $x_1, x_2, x_3, \ldots, x_m$ donde $x_1 < x_2 < x_3 < \cdots < x_m$, se define:

La función de frecuencia acumulada:

$$\boxed{afrec(x_j) = \sum_{k=1}^{j} frec(x_k)}$$

La función de proporción acumulada:

$$\boxed{F(x_j) = \sum_{k=1}^{j} p(x_k)}$$

# Función de frecuencia y proporción acumulada

Definición

## Definición

Para una variable $X$, que asume $m$ distintos valores $x_1, x_2, x_3, \ldots, x_m$ donde $x_1 < x_2 < x_3 < \cdots < x_m$, se define:

La función de frecuencia acumulada:

$$\boxed{afrec(x_j) = \sum_{k=1}^{j} frec(x_k)}$$

La función de proporción acumulada:

$$\boxed{F(x_j) = \sum_{k=1}^{j} p(x_k)}$$

# Función de frecuencia y proporción acumulada

Definición

## Definición

Para una variable $X$, que asume $m$ distintos valores $x_1, x_2, x_3, \ldots, x_m$ donde $x_1 < x_2 < x_3 < \cdots < x_m$, se define:

La función de frecuencia acumulada:

$$\boxed{afrec(x_j) = \sum_{k=1}^{j} frec(x_k)}$$

La función de proporción acumulada:

$$\boxed{F(x_j) = \sum_{k=1}^{j} p(x_k)}$$

# Función de frecuencia y proporción acumulada

## Función de frecuencia y proporción acumulada

| $j$ | $x_j$ | $frec(x_j)$ | $p(x_j)$ |
|---|---|---|---|
| 1 | 19 | 2 | .04 |
| 2 | 20 | 3 | .06 |
| 3 | 21 | 1 | .02 |
| 4 | 22 | 7 | .14 |
| 5 | 23 | 6 | .12 |
| 6 | 24 | 7 | .14 |
| 7 | 25 | 5 | .10 |
| 8 | 26 | 5 | .10 |
| 9 | 27 | 5 | .10 |
| 10 | 28 | 3 | .06 |
| 11 | 29 | 2 | .04 |
| 12 | 30 | 1 | .02 |
| 13 | 33 | 1 | .02 |
| 14 | 34 | 1 | .02 |
| 15 | 42 | 1 | .02 |
| Total | | 50 | 1.00 |

# Función de frecuencia y proporción acumulada

## Función de frecuencia y proporción acumulada

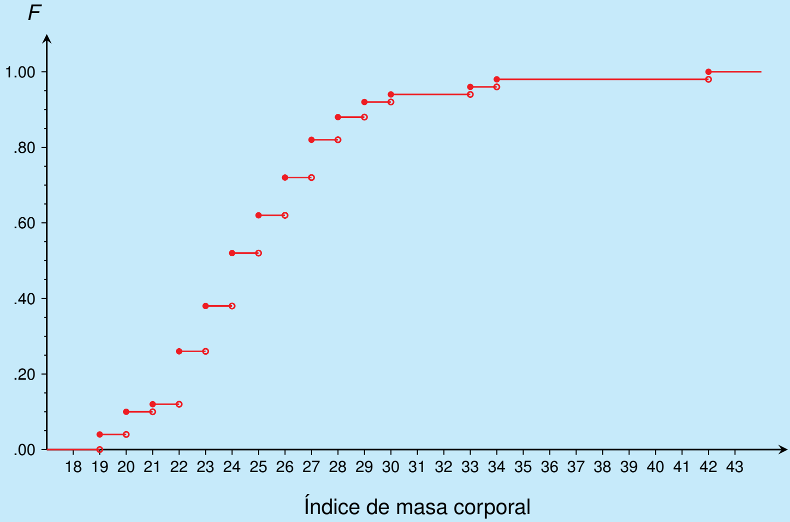
| $j$ | $x_j$ | $frec(x_j)$ | $p(x_j)$ | $afrec(x_j)$ |
|---|---|---|---|---|
| 1 | 19 | 2 | .04 | 2 |
| 2 | 20 | 3 | .06 | 5 |
| 3 | 21 | 1 | .02 | 6 |
| 4 | 22 | 7 | .14 | 13 |
| 5 | 23 | 6 | .12 | 19 |
| 6 | 24 | 7 | .14 | 26 |
| 7 | 25 | 5 | .10 | 31 |
| 8 | 26 | 5 | .10 | 36 |
| 9 | 27 | 5 | .10 | 41 |
| 10 | 28 | 3 | .06 | 44 |
| 11 | 29 | 2 | .04 | 46 |
| 12 | 30 | 1 | .02 | 47 |
| 13 | 33 | 1 | .02 | 48 |
| 14 | 34 | 1 | .02 | 49 |
| 15 | 42 | 1 | .02 | 50 |
| Total | | 50 | 1.00 | |

# Función de frecuencia y proporción acumulada

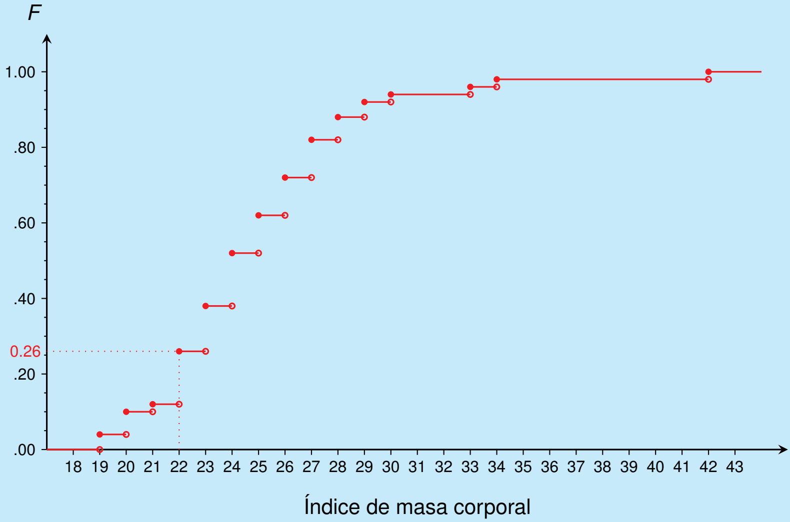
## Función de frecuencia y proporción acumulada

| $j$ | $x_j$ | $frec(x_j)$ | $p(x_j)$ | $afrec(x_j)$ | $F(x_j)$ |
|---|---|---|---|---|---|
| 1 | 19 | 2 | .04 | 2 | .04 |
| 2 | 20 | 3 | .06 | 5 | .10 |
| 3 | 21 | 1 | .02 | 6 | .12 |
| 4 | 22 | 7 | .14 | 13 | .26 |
| 5 | 23 | 6 | .12 | 19 | .38 |
| 6 | 24 | 7 | .14 | 26 | .52 |
| 7 | 25 | 5 | .10 | 31 | .62 |
| 8 | 26 | 5 | .10 | 36 | .72 |
| 9 | 27 | 5 | .10 | 41 | .82 |
| 10 | 28 | 3 | .06 | 44 | .88 |
| 11 | 29 | 2 | .04 | 46 | .92 |
| 12 | 30 | 1 | .02 | 47 | .94 |
| 13 | 33 | 1 | .02 | 48 | .96 |
| 14 | 34 | 1 | .02 | 49 | .98 |
| 15 | 42 | 1 | .02 | 50 | 1.00 |
| Total | | 50 | 1.00 | | |

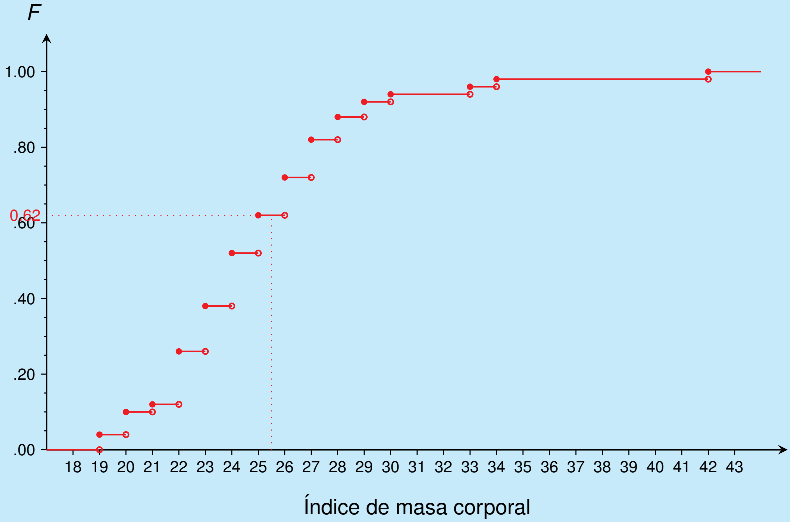
# Representación gráfica de la función de proporción acumulada



# Representación gráfica de la función de proporción acumulada



# Representación gráfica de la función de proporción acumulada



# Función de distribución

Definición

## Definición

Para una variable $X$ discreta, que asume $m$ distintos valores $x_1, x_2, x_3, \ldots, x_m$ donde $x_1 < x_2 < x_3 < \cdots < x_m$, se define la función de distribución como:

$$\boxed{\Phi_X(x_j) = \sum_{k=1}^{j} \pi_X(x_k)}$$

## Definición

Para una variable $X$ continua, se define la función de distribución como:

$$\Phi_X(x) = \int_{-\infty}^{x} \varphi_X(u)du,$$

lo cual corresponde con el área debajo de $\varphi_X$ a la izquierda de $x$.

# Función de distribución

Definición

## Definición

Para una variable $X$ discreta, que asume $m$ distintos valores $x_1, x_2, x_3, \ldots, x_m$ donde $x_1 < x_2 < x_3 < \cdots < x_m$, se define la función de distribución como:

$$\boxed{\Phi_X(x_j) = \sum_{k=1}^{j} \pi_X(x_k)}$$

## Definición

Para una variable $X$ continua, se define la función de distribución como:

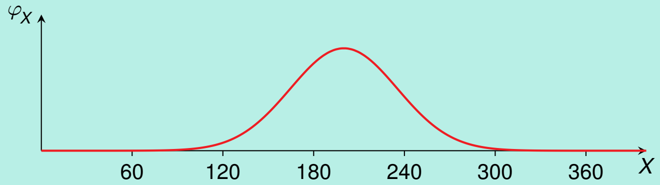
$$\Phi_X(x) = \int_{-\infty}^{x} \varphi_X(u)du,$$

lo cual corresponde con el área debajo de $\varphi_X$ a la izquierda de $x$.

# Función de distribución

Definición

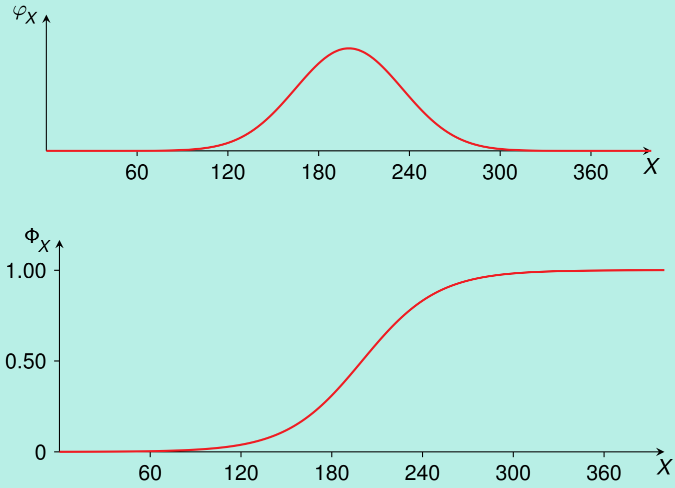
## La distribución normal acumulada



# Función de distribución

Definición

## La distribución normal acumulada



# Índice

# Índice

# Índice

# Índice

# Índice

# Índice

# Índice

# Índice