Inferencias con cifras

Métodos de investigación cuantitativa

UAH-EMAPE

Enfoques Metodológicos para el Análisis de Políticas Educativas

Septiembre 28

2024

Profesor invitado

Carrasco, D., PhD,
Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile

Centro UC Medición - MIDE

Interpretación de cifras

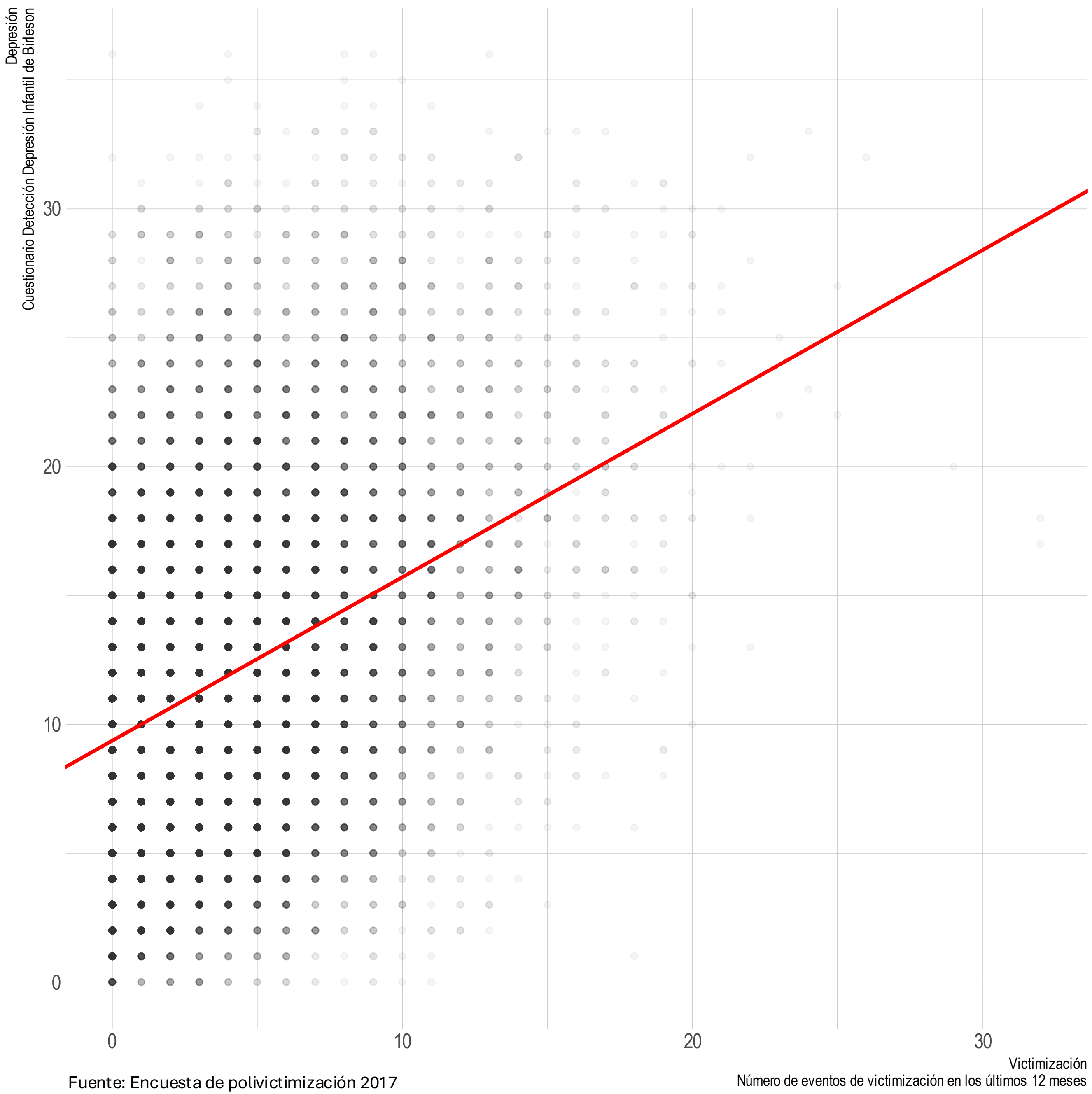
En la investigación cuantitativa se emplean modelos estadísticos para producir análisis de datos.

Estos análisis nos entregan resultados que han de ser interpretados para realizar afirmaciones acerca de:

a) Cuál es la proporción (i.e., cantidad) de observaciones que presenta un atributo.

b) Cuál es relación de entre dos variables
   a) Si hay o no relación.
   b) Que dirección toma esta relación
   c) Que tamaño tiene esta relación.

De todos los modelos estadísticos, el modelo más fundamental es el modelo de regresión. Esta herramienta nos permite obtener estimados de cantidad, así como de relaciones entre variables.



Fuente: Encuesta de polivictimización 2017

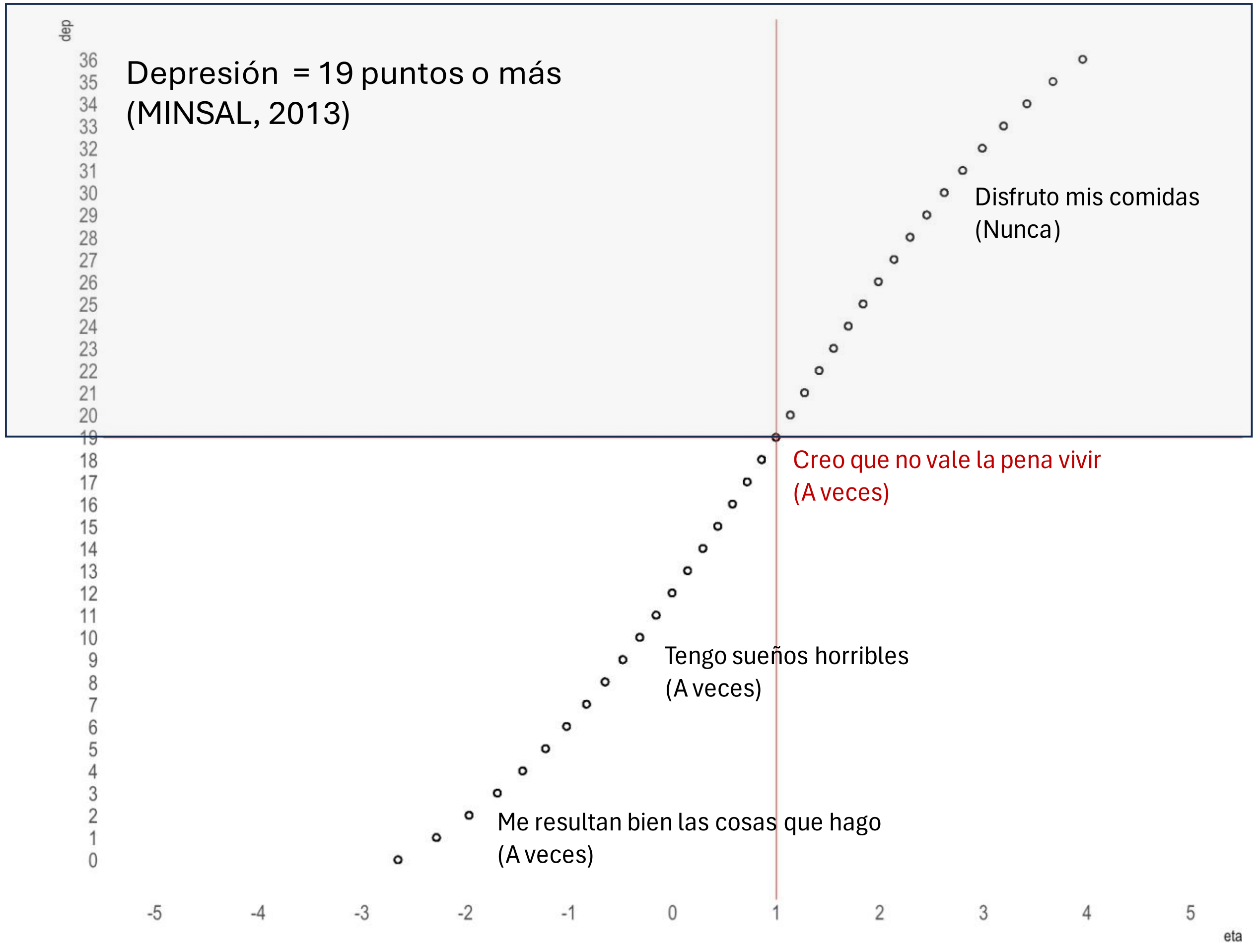
Diseño e interpretación de cifras

En la clase anterior nos centramos en como diferentes aspectos de diseño cambian la interpretación de resultados obtenidos. En particular, que condicional a como obtuvimos las observaciones con las que contamos, que tipo de afirmaciones podemos realizar:

a) Si nuestras afirmaciones realizadas con cifras se extienden a la población de origen de las observaciones.

b) Y en el otro extremo, si nuestras afirmaciones realizadas con cifras se extienden solo a las observaciones con las que contamos.

En la clase de hoy veremos ejemplos de interpretación de cifras, condicionales a diseño. Además, incluiremos elementos de **inferencia** para entender por qué se puede realizar las afirmaciones que vamos a realizar.



Fuente: Encuesta de polivictimización 2017

Inferencia

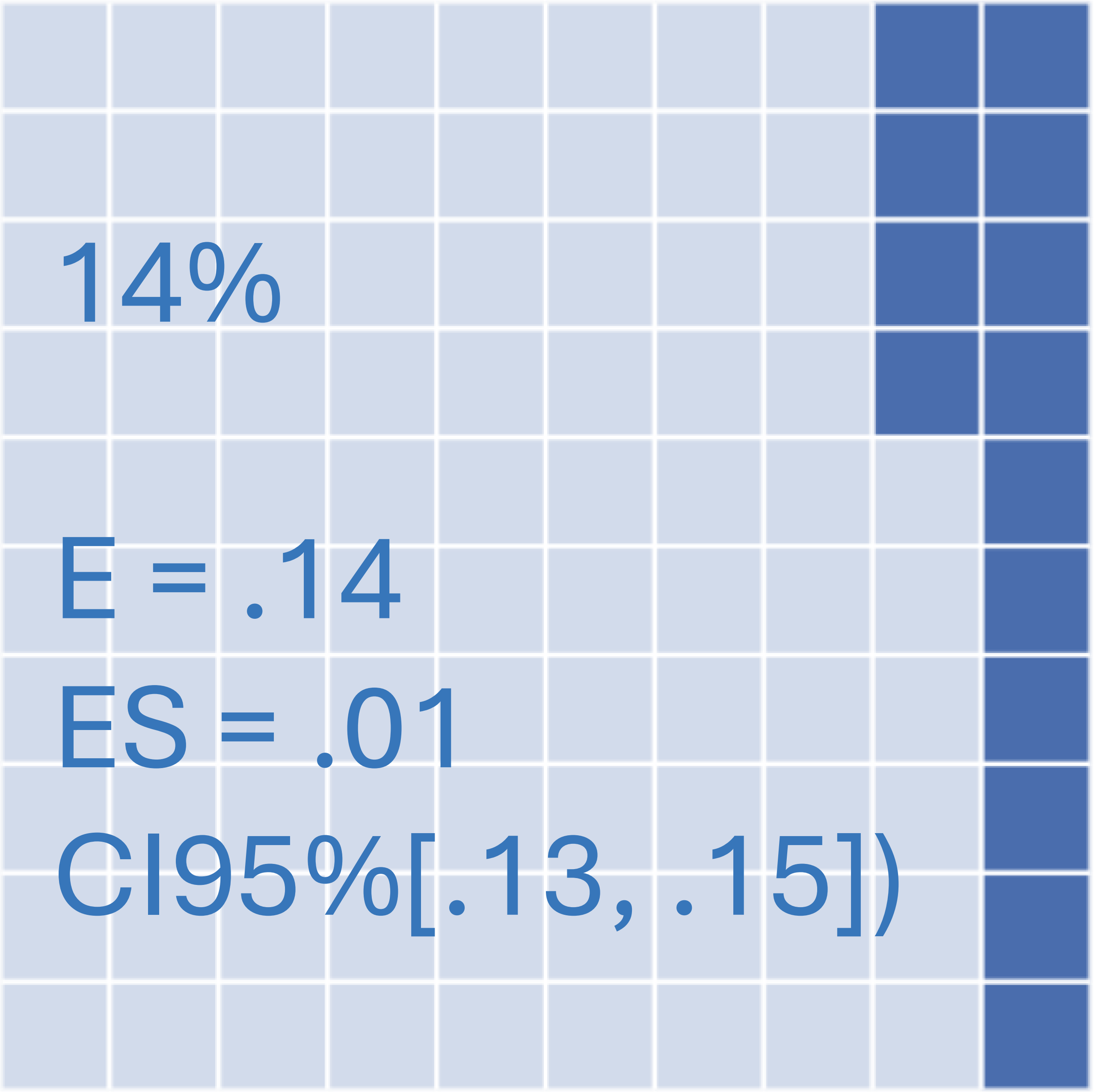
Inferencia basada en diseño

¿Cuántos estudiantes se encuentran en riesgo de depresión?

Encuesta de Polivictimización en niños, niñas, y adolescentes (2017)

Estimado poblacional

Empleando datos de la Encuesta Polivictimización en niños, niñas, y adolescentes (2017) de la prevalencia de riesgo de depresión en escolares en Chile de séptimo básico a tercero medio es de 14% (E = .14, ES = .01, CI95%[.13, .15]). Los resultados anteriores implican que 14 de 100 estudiantes auto reporta síntomas depresivos, los cuales incluyen a la ideación suicida.



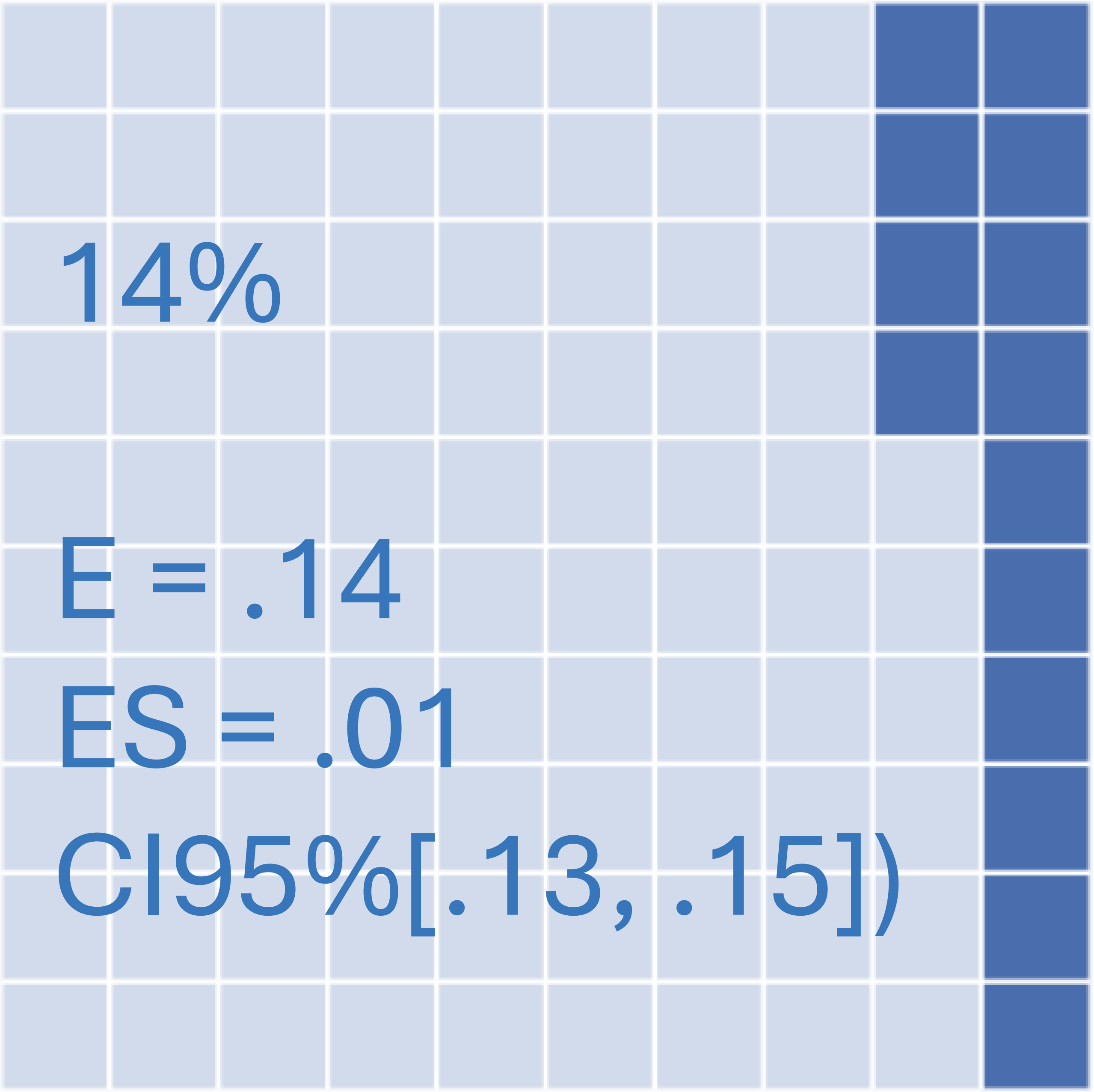
Cifras reportadas

Proporción estimada
a la población finita.

Punto estimado

Error de estimación

Intervalo de confianza al 95%



¿Por qué creerle a esta cifra?

Hay diferentes elementos, que permiten que esta cifra sea creíble.

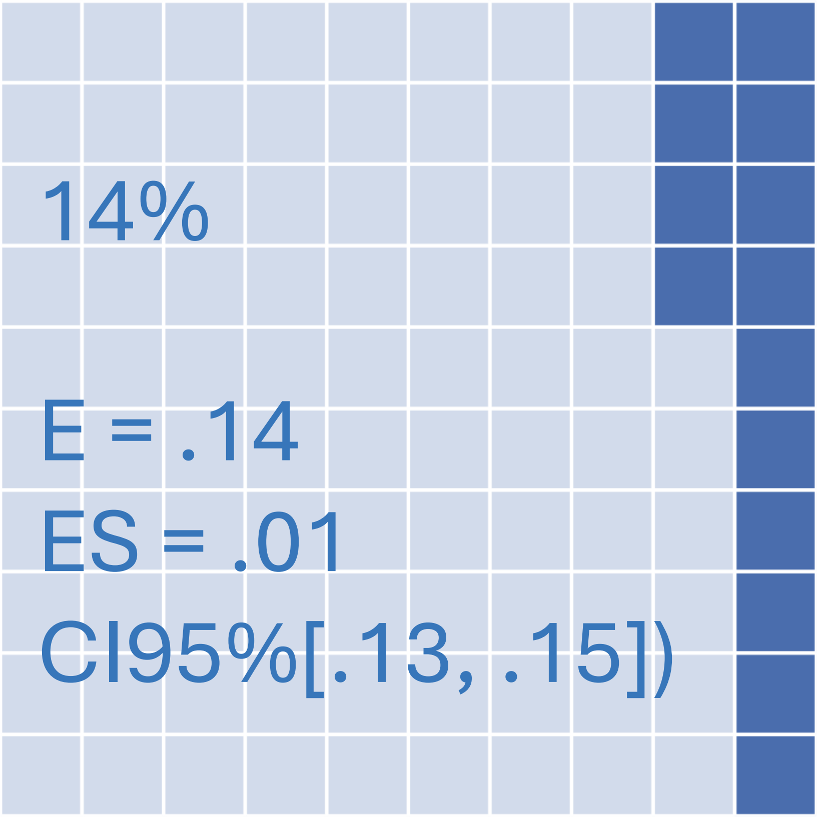
Instrumento adecuado

Puntaje Interpretable

Diseño Muestral

Distribución Muestral de estadígrafos

- Ley de los grandes Números
- Teorema del límite central

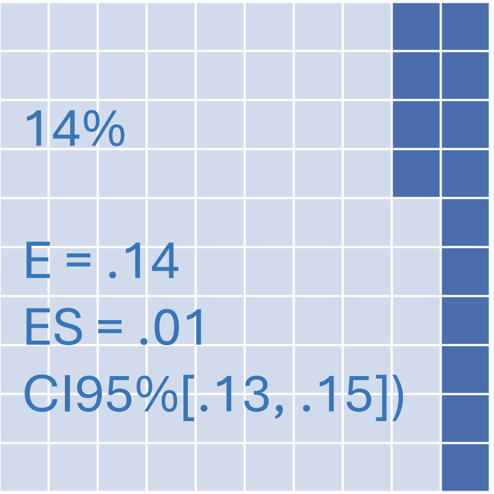


Inferencia

Inferencia basada en diseño

Elementos de soporte a la inferencia realizada

¿Por qué creerle a esta cifra?



Hay diferentes elementos, que permiten que esta cifra sea creíble.

El primer elemento es que estamos empleando un **instrumento pertinente** al objetivo.

Empleamos los puntajes de la escala de Birleson. Este instrumento presenta una alta sensibilidad de 87.5% (proporción de casos con depresión clínica correctamente clasificados); y alta especificidad de 93% (proporción de casos correctamente clasificados como sin depresión) (MINSAL, 2013).

ANEXO 5. Cuestionario Detección Depresión Infantil de Birleson

Instrucciones: Por favor contesta tan honestamente como puedas; las frases se refieren a como tú te has sentido la semana pasada.

No hay respuestas correctas, es importante que cuentes cómo te has sentido.

Nº		Siempre	A veces	Nunca
1	Pienso que los días que vienen serán agradables	0	1	2
2	Siento ganas de llorar	2	1	0
3	Tengo ganas de arrancar o escapar	2	1	0
4	Tengo dolores de estómago	2	1	0
5	Quiero salir a jugar fuera de casa	0	1	2
6	Duermo muy bien	0	1	2
7	Tengo mucha energía	0	1	2
8	Disfruto mis comidas	0	1	2
9	Puedo arreglármelas solo, valerme por mi mismo	0	1	2
10	Creo que no vale la pena vivir	2	1	0
11	Me resultan bien las cosas que hago	0	1	2
12	Disfruto lo que hago igual que antes	0	1	2
13	Me gusta hablar con mi familia	0	1	2
14	Tengo sueños horribles	2	1	0
15	Me siento muy solo	2	1	0
16	Me animo fácilmente, me entusiasmo con mucha facilidad	0	1	2
17	Me siento tan triste que difícilmente lo soporto	2	1	0
18	Me siento muy aburrido, lateado, choreado	2	1	0

Asignación de puntaje:

Nº	orden	Nº	orden	Nº	orden	Nº	orden	Nº	orden	Nº	orden
1	0-1-2	**4**	2-1-0	**7**	0-1-2	**10**	2-1-0	**13**	0-1-2	**16**	0-1-2
2	2-1-0	**5**	0-1-2	**8**	0-1-2	**11**	0-1-2	**14**	2-1-0	**17**	2-1-0
3	2-1-0	**6**	0-1-2	**9**	0-1-2	**12**	0-1-2	**15**	2-1-0	**18**	2-1-0

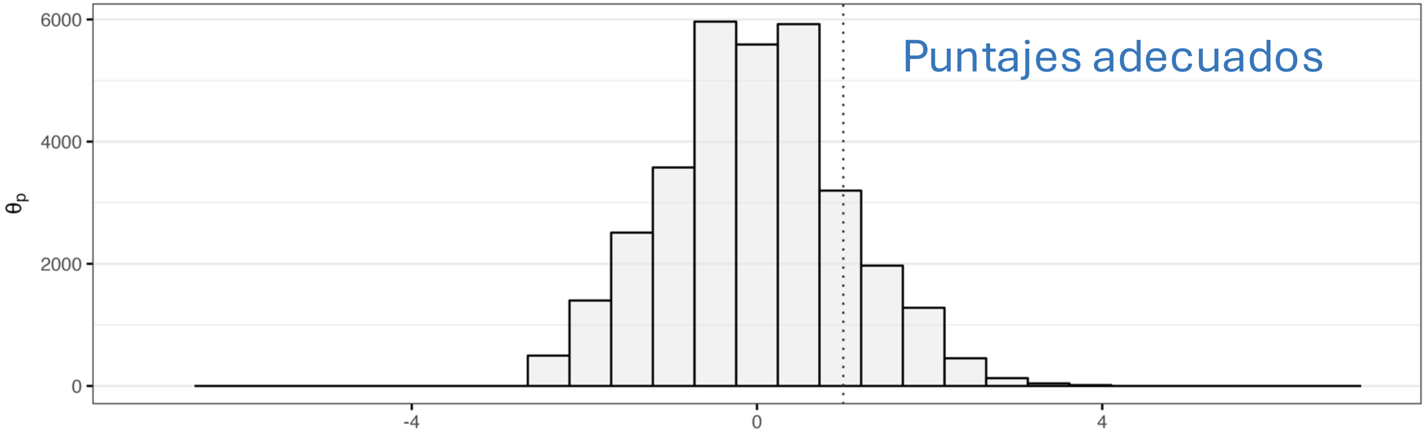
Puntaje de corte: Mayor o igual a 19 puntos. Sensibilidad 87.5% y especificidad cercanas a 93%. Valor predictivo 50%.

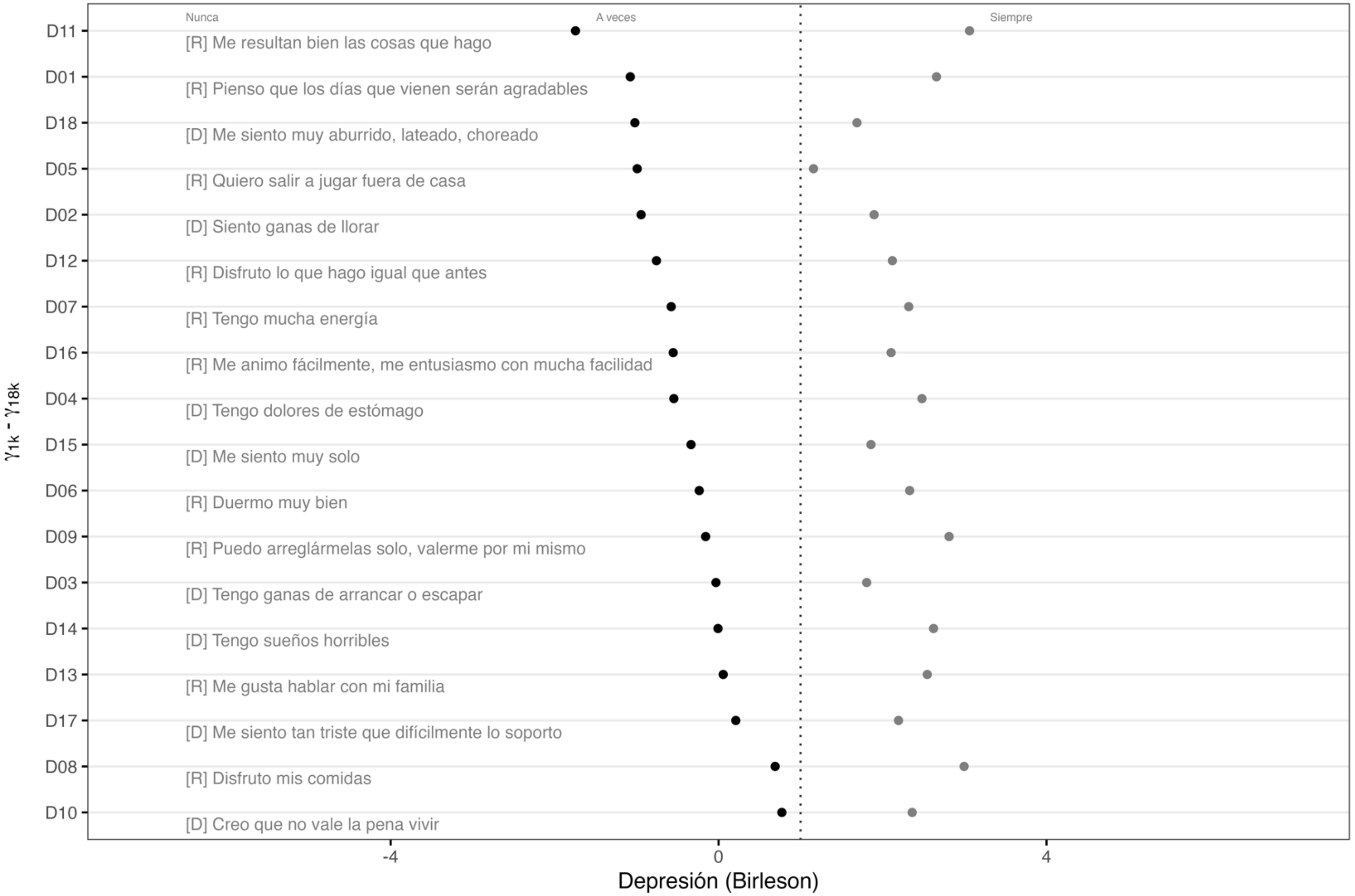
¿Por qué creerle a esta cifra?

Hay diferentes elementos, que permiten que esta cifra sea creíble.

Con las respuestas recogidas, **podemos construir un puntaje total** (suma de las respuestas) con las cuales **podemos clasificar a los estudiantes** como en riesgo de depresión, o bajo riesgo de depresión.

Estudiantes que presentan 19 o más puntos, son más proclives a reportar ideación suicida.



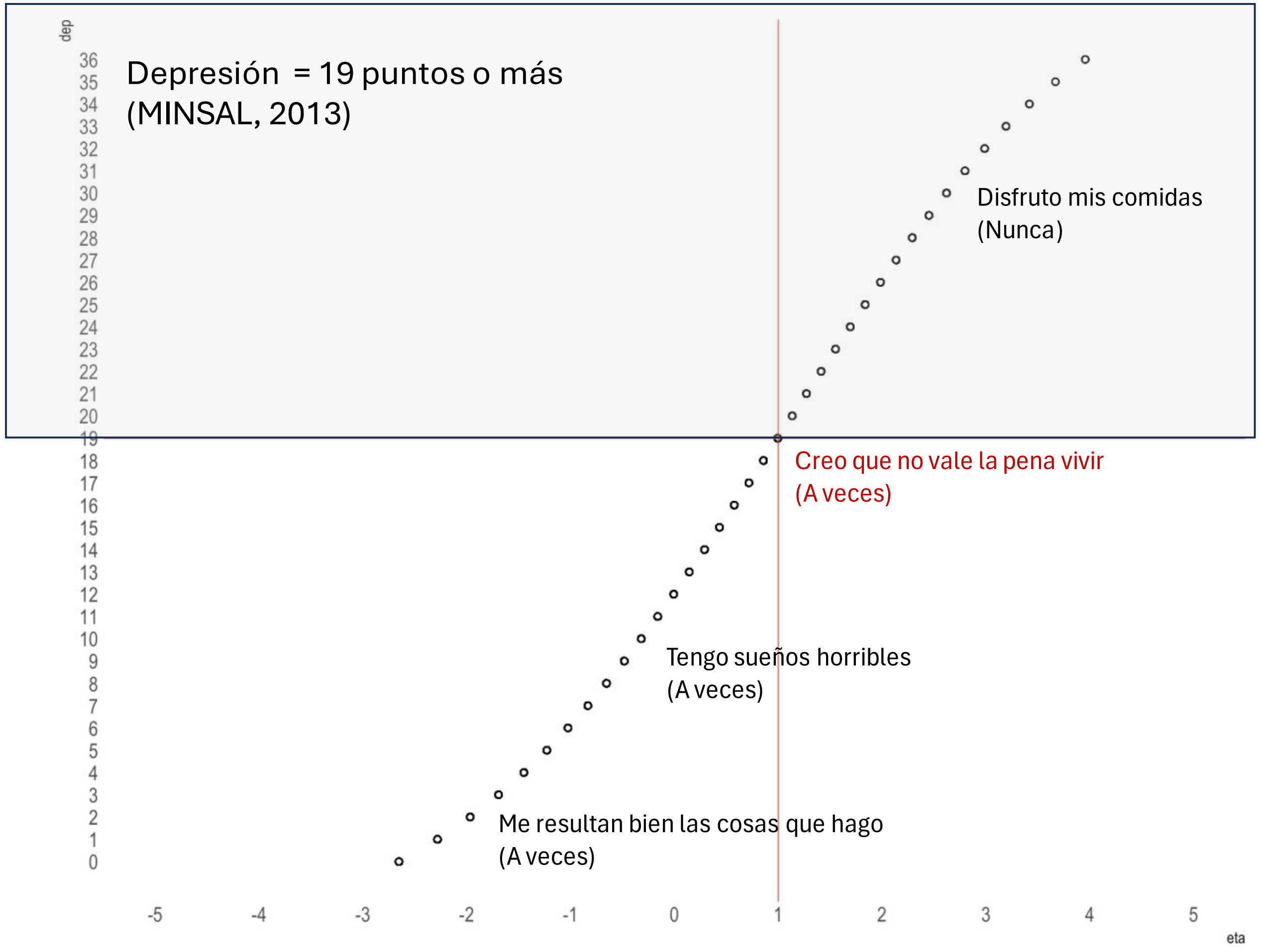


¿Por qué creerle a esta cifra?

Hay diferentes elementos, que permiten que esta cifra sea creíble.

Con las respuestas recogidas, podemos construir un puntaje total (suma de las respuestas) con las cuales podemos clasificar a los estudiantes como en riesgo de depresión, o bajo riesgo de depresión.

Estudiantes que **presentan 19 o más puntos**, son más proclives a reportar ideación **suicida**.

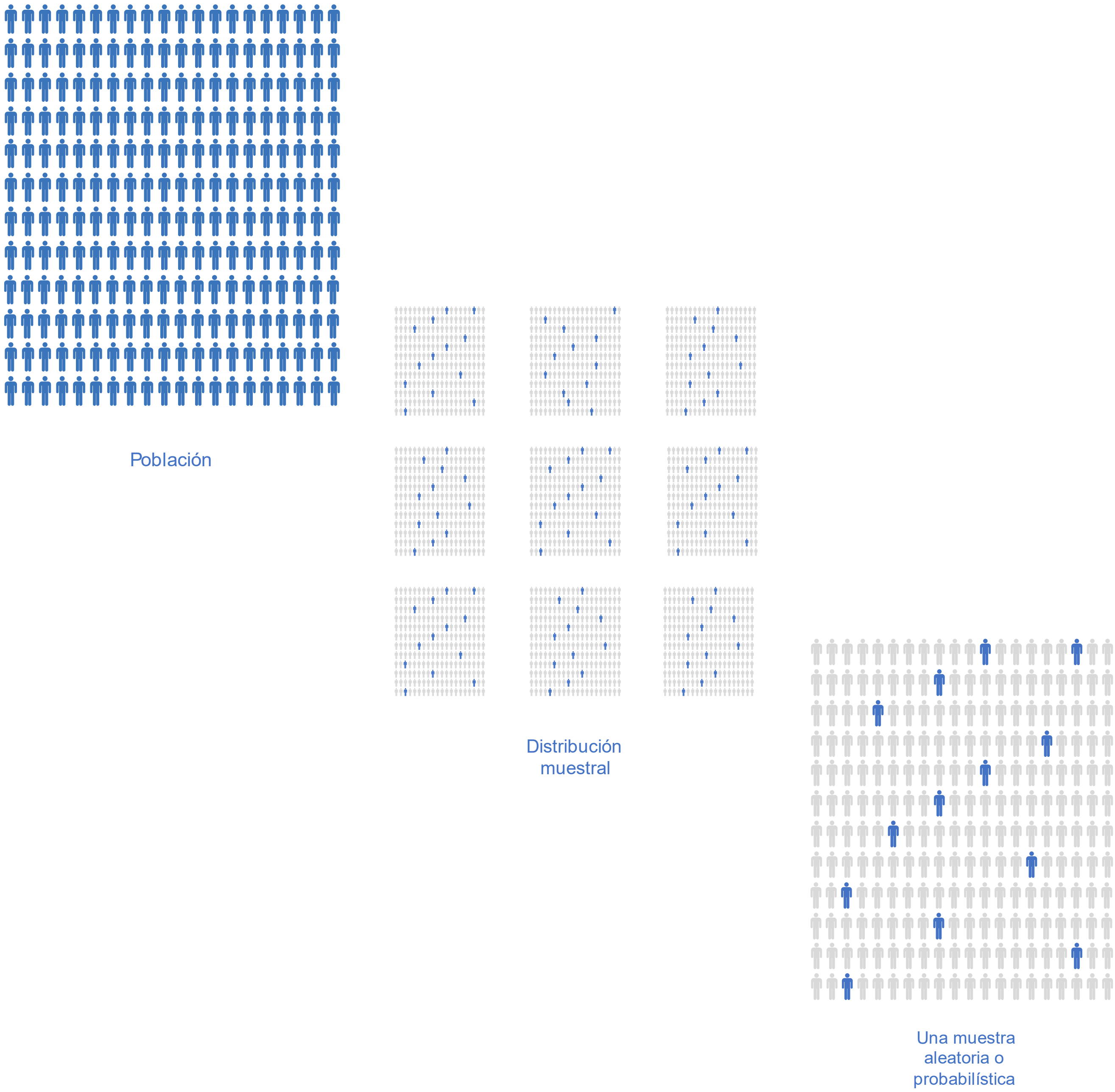


¿Por qué creerle a esta cifra?

Hay diferentes elementos, que permiten que esta cifra sea creíble.

La Encuesta Polivictimización en niños, niñas, y adolescentes (2017) es un estudio que emplea una La **muestra es probabilística** de tres etapas, donde se seleccionan escuelas, luego cursos, y finalmente estudiantes.

Incluye a un total de 19.864 estudiantes, de séptimo básico a tercero medio de diferentes regiones del país, de 699 escuelas, tanto particulares, subvencionadas, y municipales.



Diseño Muestral adecuado

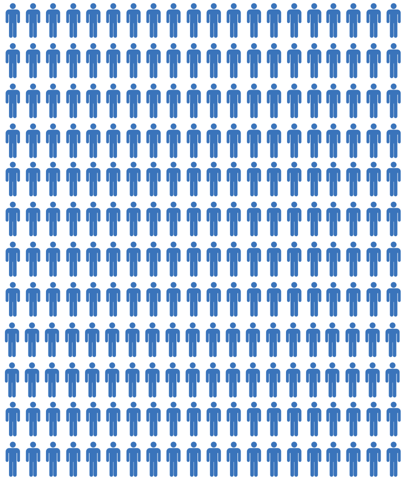
¿Por qué creerle a esta cifra?

Hay diferentes elementos, que permiten que esta cifra sea creíble.

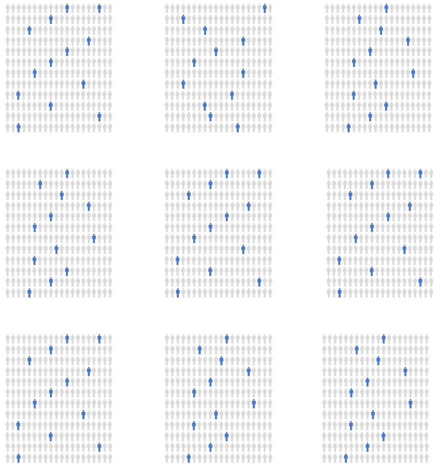
¿Por qué con menos casos que toda la población de observaciones podemos hacer generalizaciones a la población de estudiantes?

¿Qué es lo que tienen los diseños muestrales aleatorios y probabilísticos, que nos permiten realizar inferencias a la población?

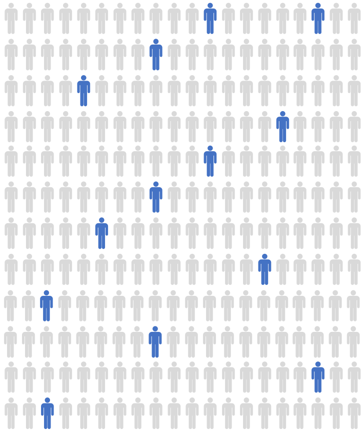
Para que me crean que podemos hacer estas afirmaciones ("*14 de 100 estudiantes está en riesgo de depresión"*), necesitamos revisar que son las distribuciones muestrales de los estadígrafos, y porque son útiles.



Población



Distribución muestral



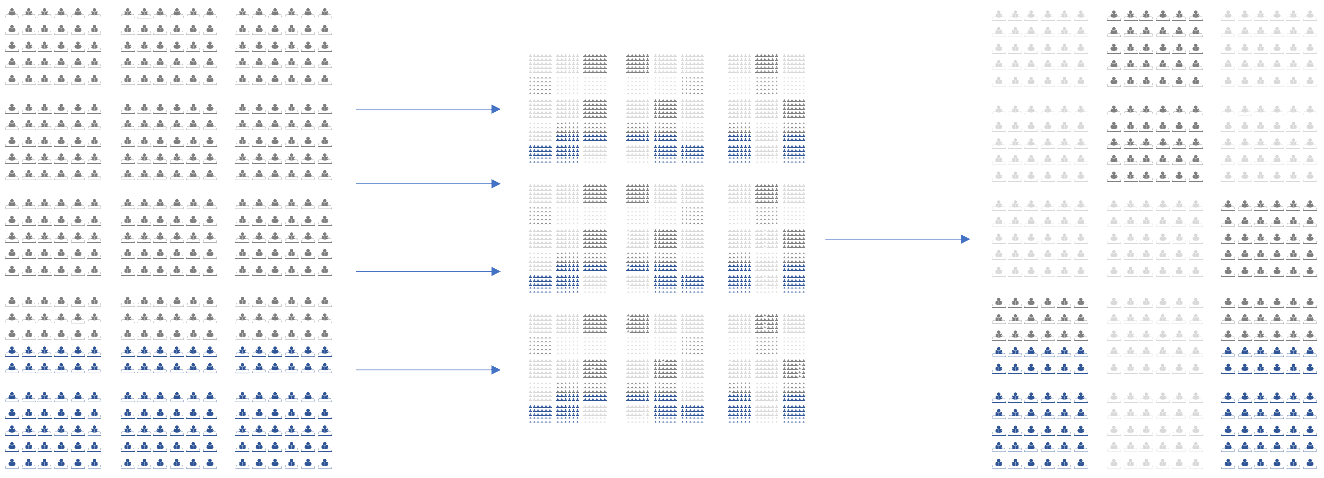
Una muestra aleatoria o probabilística

Inferencia

¿En qué consiste la inferencia?

Cómo sacamos conclusiones a partir de cifras

Lógica de la inferencia



población

muestras posibles

muestra compleja

Una intuición sencilla, respecto a la acción de realizar inferencias, es comenzar con una población finita (i.e., un listado exhaustivo de observaciones).

De este listado, podemos extraer una muestra, la cual es un caso especial de todas las muestras posibles (i.e., la distribución muestral).

Debido a que la distribución muestral posee ciertas características (i.e., conforma a una distribución normal), nos atrevemos a realizar ciertas afirmaciones.

la prevalencia de riesgo de depresión en escolares en Chile de séptimo básico a tercero medio es de 14% (E = .14, ES = .01, CI95%[.13, .15]), lo que implican que 14 de 100 estudiantes auto reporta síntomas depresivos, los cuales incluyen a la ideación suicida.

O podemos decir que un 57% de los mismos estudiantes (CI95%[54%, 59%]), **indica correctamente cual es el rol de una constitución**, empleando datos de ICCS 2009.

¿**Por qué** uno podría creer que estas cifras nos permiten hacer **generalizaciones** sobre la población de estudiantes de un país?

¿**Qué significan las cifras** anteriores (e.g., E, SE, CI95%)?

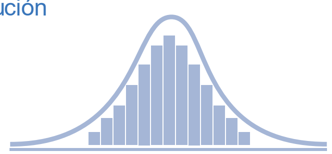
Tipos de inferencia

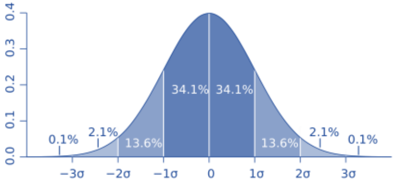
$$\mu$$

La población finita

$$\hat{x}$$

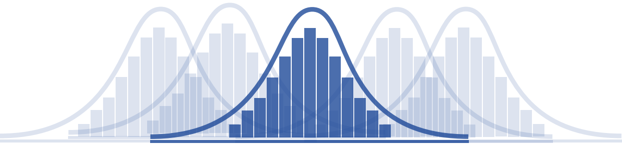
La distribución muestral





$$\bar{x}_1$$

Una muestra aleatoria

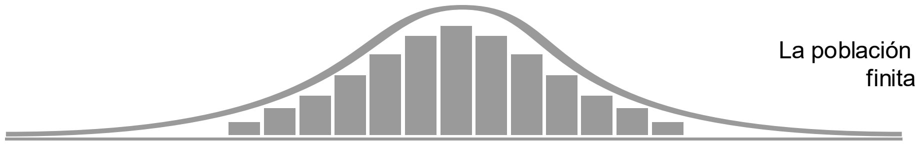


La población infinita, o de valores posibles

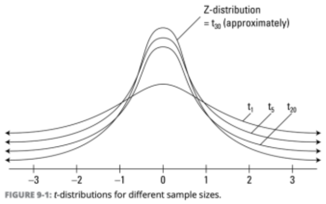
$$y_i = \beta_0 + \beta_1 x_i + e_i$$

Mecanismo generador de datos

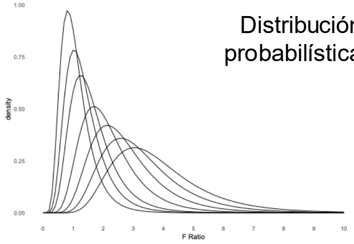
La población finita



Distribución probabilística

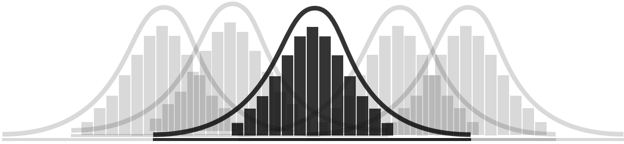


$$t, F$$



$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$

Una muestra (e.g., aleatoria, condicionada, o experimental)



Sterba, 2009; Binder & Roberts, 2016; Cobb, 2017

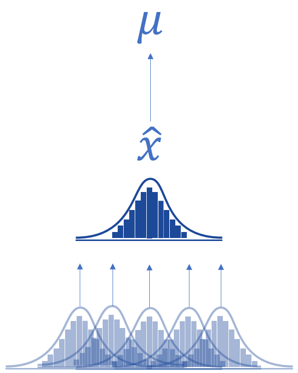
Inferencias con cifras

Para que tenga sentido que es posible producir cifras que nos permiten hacer **inferencias** es necesario considerar diferentes elementos, sobre los que se posa la idea de hacer inferencias.

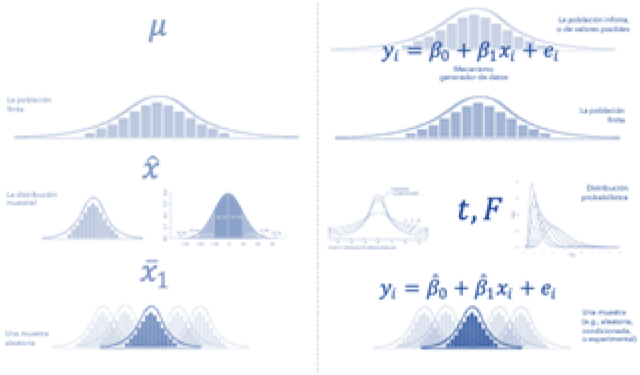
Una de ellas, es la idea de producir inferencias basados en **distribuciones de probabilidad**. Tanto, la inferencia basada en modelo, como la inferencia basada en diseño, emplean a este elemento.

Adicionalmente, es importante considerar al **diseño con el cual se produjeron las observaciones**. Ya que, que la credibilidad de una afirmación planteada con resultados (i.e., *tenable*), depende de las características del diseño que producen las observaciones.

Diseños e inferencias



Distribución probabilística de estadigrafos



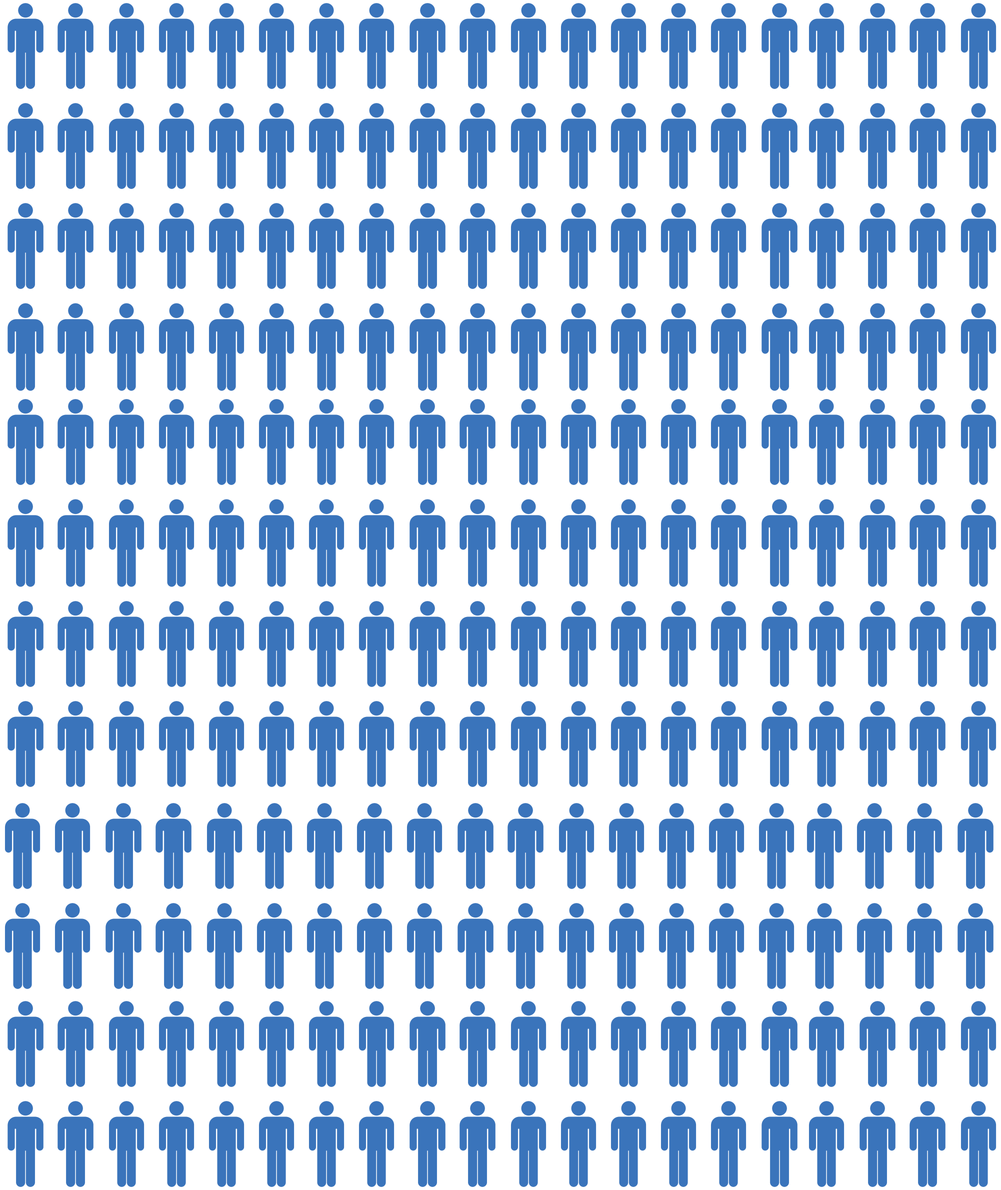
Marcos de inferencia (basada en diseño, y basada en modelo)

Inferencia

Distribución Muestral

Tamaño muestral y precisión de los estimados

Inferencia a la población



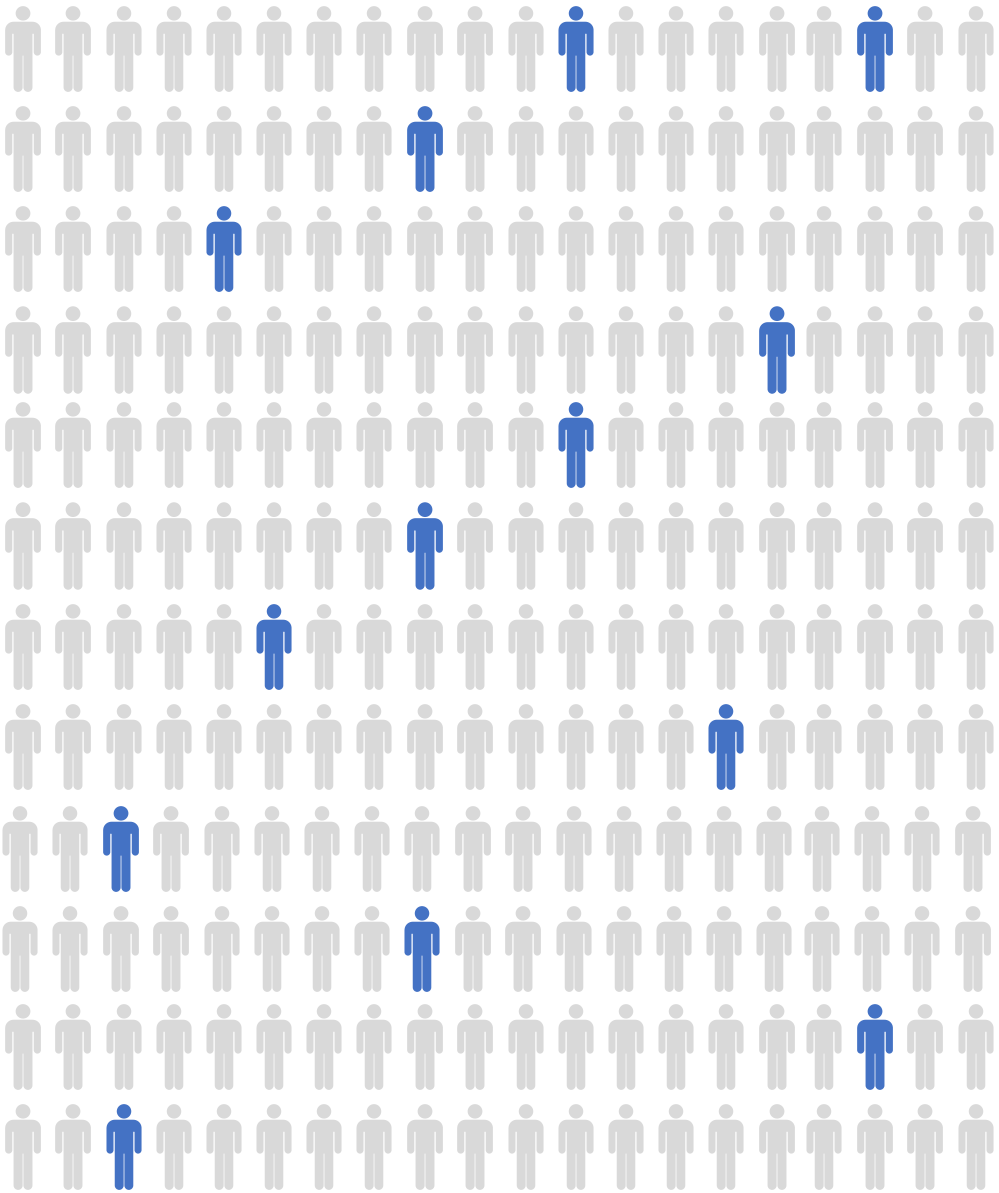
μ

Parámetro poblacional

Podemos describir las poblaciones, en base a parámetros. Los mismos que vimos respecto a las **distribuciones**, como las proporciones, media, mediana, desviación estándar, varianza y percentiles.

Proporciones
Media
Mediana
Deviación Estándar
Varianza

Inferencia a la población



μ

Parámetro poblacional

Proporciones
Media
Mediana
Deviación Estándar
Varianza

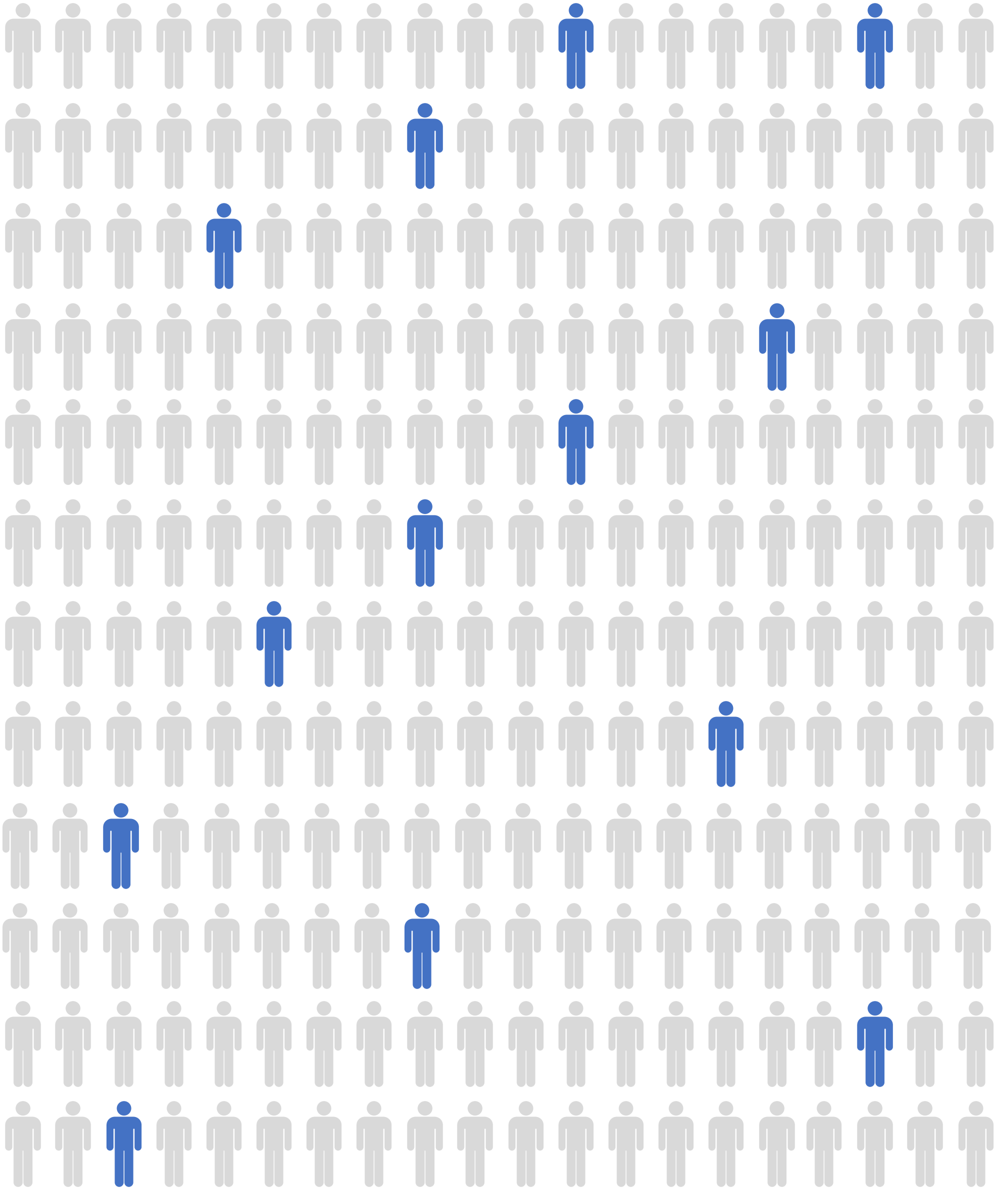
La idea general de la inferencia a la población finita es qué con muestras aleatorias, podemos acercarnos a los parámetros poblacionales.

A estas cifras obtenidas con las muestras, les llamamos estadígrafos.

$\bar{x}$

estadígrafo

Inferencia a la población



μ

Parámetro poblacional

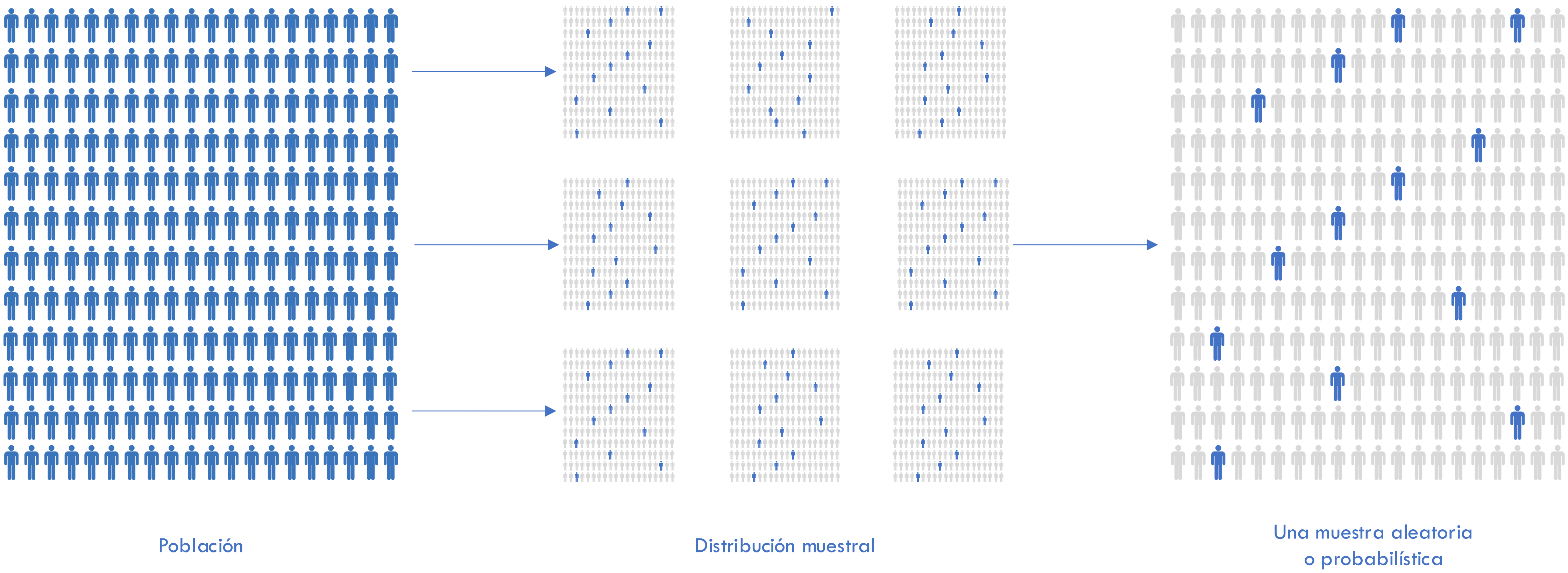
Si la forma en que se obtuvieron las observaciones es **aleatoria**, o en un sentido más amplio es **probabilístico**, podemos inferir los parámetros poblacionales con los estadígrafos calculados.

Proporciones
Media
Mediana
Deviación Estándar
Varianza

$\bar{x}$

estadígrafo

Inferencia a la población



Inferencia a la población

Gracias a que podemos asegurar que nuestra muestra es aleatoria (i.e., **que conocemos su mecanismo de selección**), y que la distribución muestral toma una distribución conocida (normal), entonces podemos tener una expectativa razonable respecto a la dispersión de los estadígrafos.

Y como tal, tener una expectativa razonable de que tan cerca estamos del parámetro población, con algún grado de incertidumbre.

Los elementos críticos del proceso de inferencia, es que la muestra proviene de una distribución de muestras posibles, y por tanto tiene ciertas propiedades (i.e., **conocemos su distribución**).

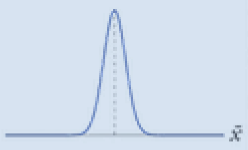
Antes de desarrollar esta idea, enfoquémonos primero en la idea de población.

población

μ

Parámetro poblacional

Distribución muestral

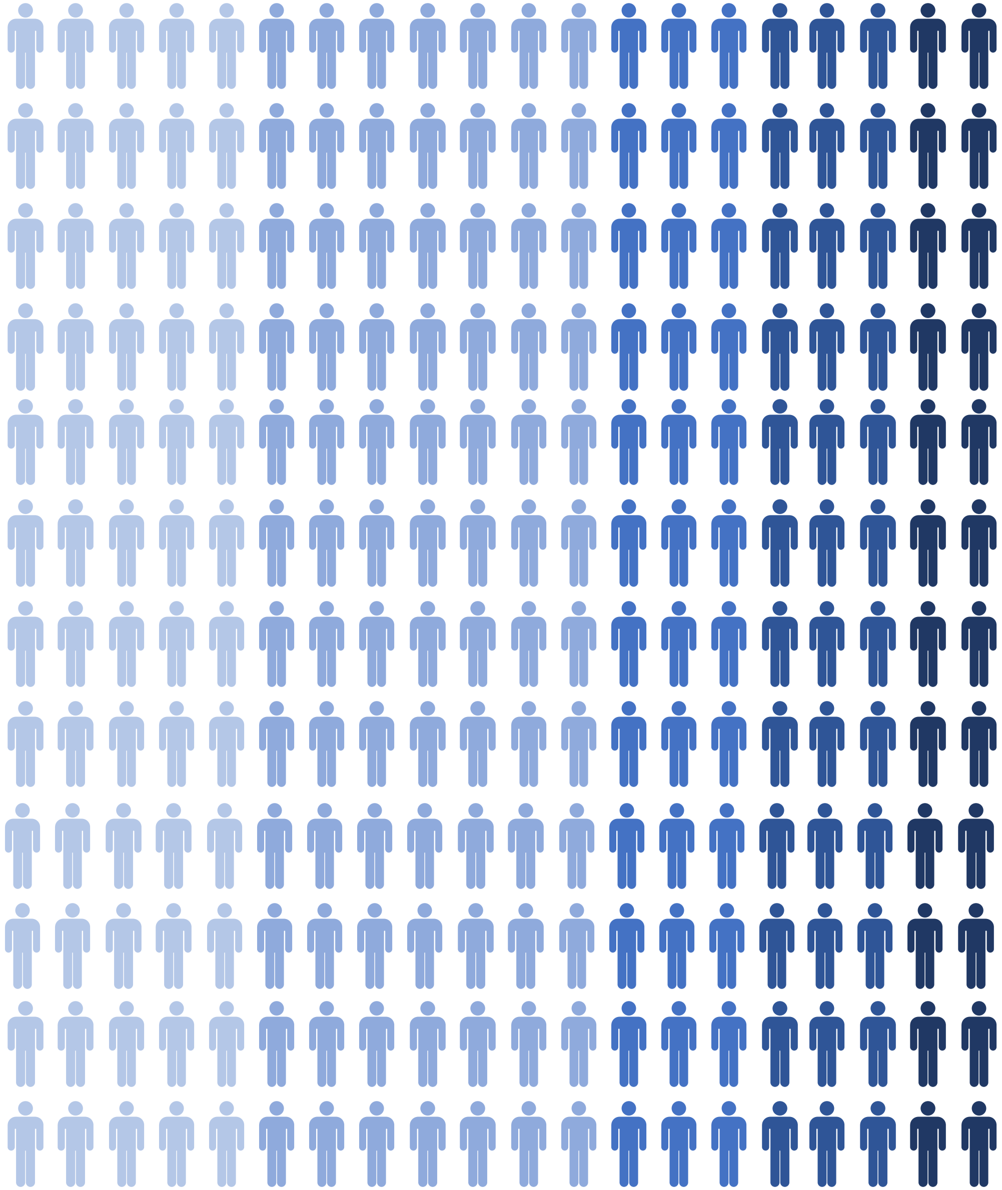


Una muestra aleatoria

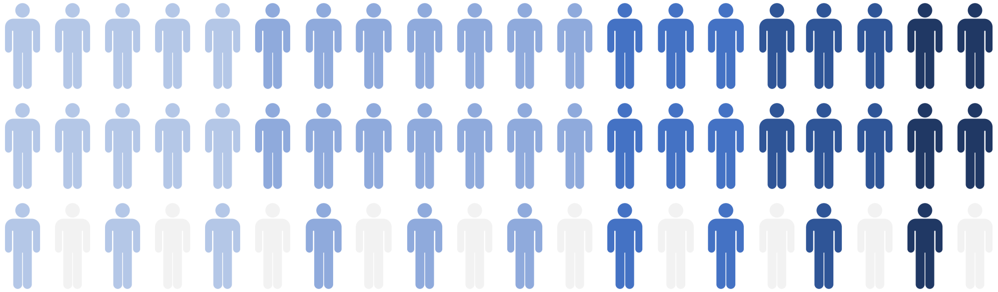
$\bar{x}$

estadígrafo

Inferencia a la población



Muestra de 400 casos



Mientras nuestra selección de casos sea aleatoria, una muestra de cerca de 400 casos tiene un intervalo de confianza de 95% con un error de cerca de 5% (Ross, 2005).

¿Por qué?

¿Por qué podemos ocupar una muestra, para inferir acerca de una población?

Ley de los grandes números

A medida que aumenta el tamaño de una colección de observaciones independientes e idénticamente distribuidas [iid] (i.e., igualmente informativas), la media de esta colección se acerca a la media de la población.

Weak law of great Numbers
Bernoulli (1693) (Cobb, 2017)

Teorema del límite central

Si tomamos una serie de muestras independientes de una población, la distribución de medias de las muestras toma una forma normal (n<30).

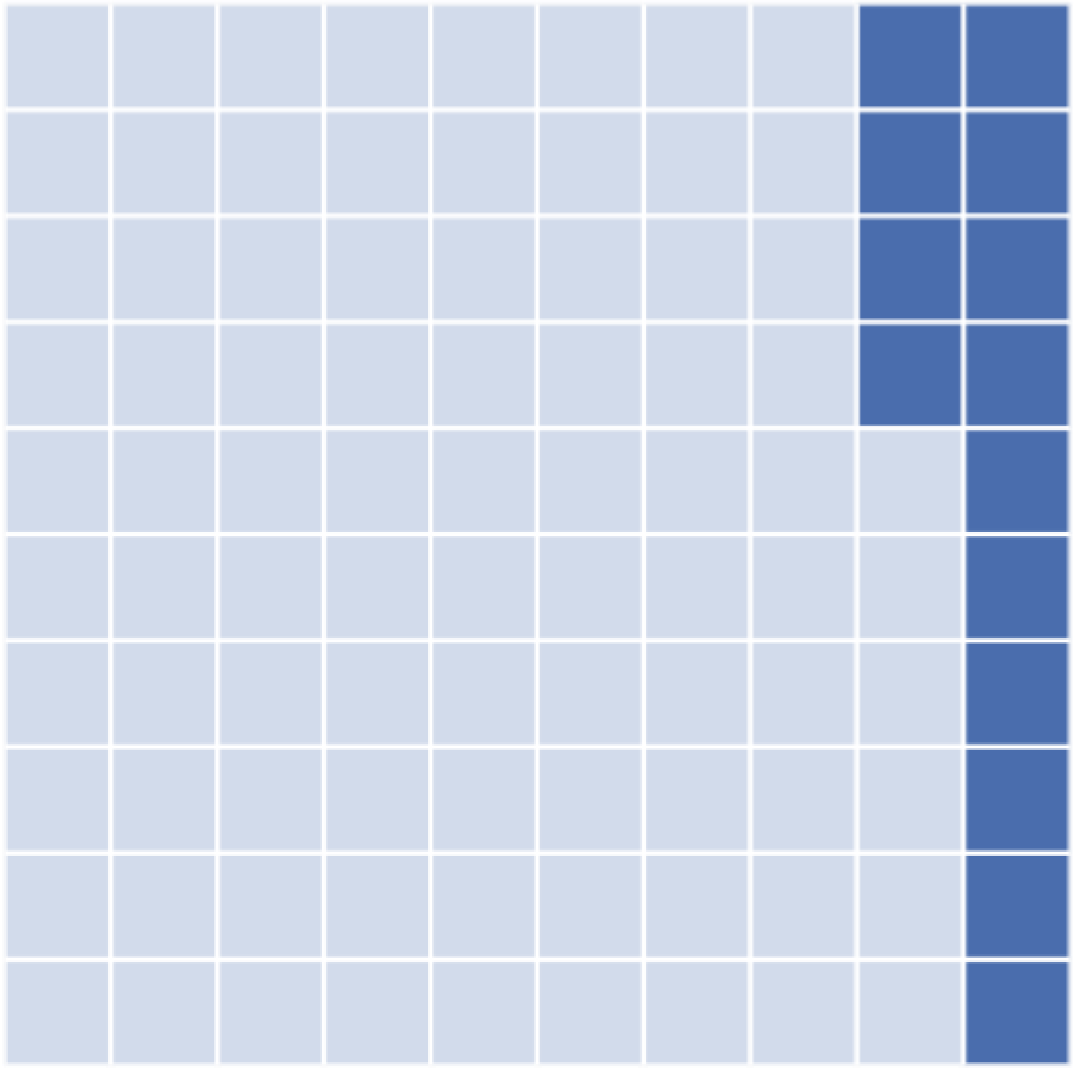
Central limit theorem
De Moivre (1720) (Cobb, 2017)

Inferencia

Ley de los grandes números

Weak law of great numbers

Ley de los grandes números



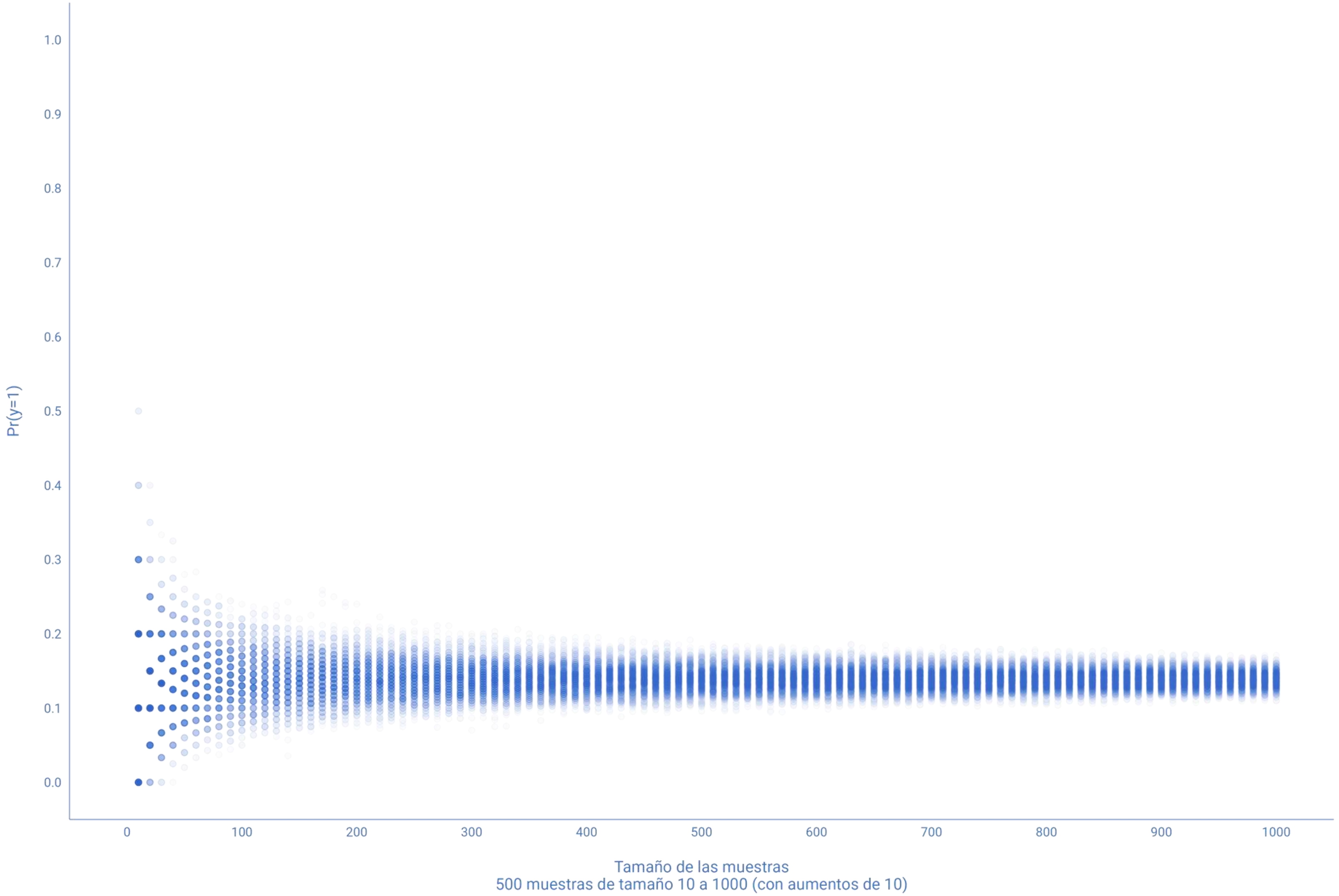
Supongamos que tenemos una población de finita de valores.

Esta población finita de valores corresponde una lista de valores 0 a 1, que corresponde a 100 mil observaciones, donde la proporción de valores 1 es de 14%.

Es decir, que tenemos una lista exhaustiva de valores, que posee la misma proporción de 1, que obtuvimos en los análisis anteriores.

Vamos a generar colecciones de muestras de diferente tamaño, para ilustrar la ley de los grandes números.

Ley de los grandes números

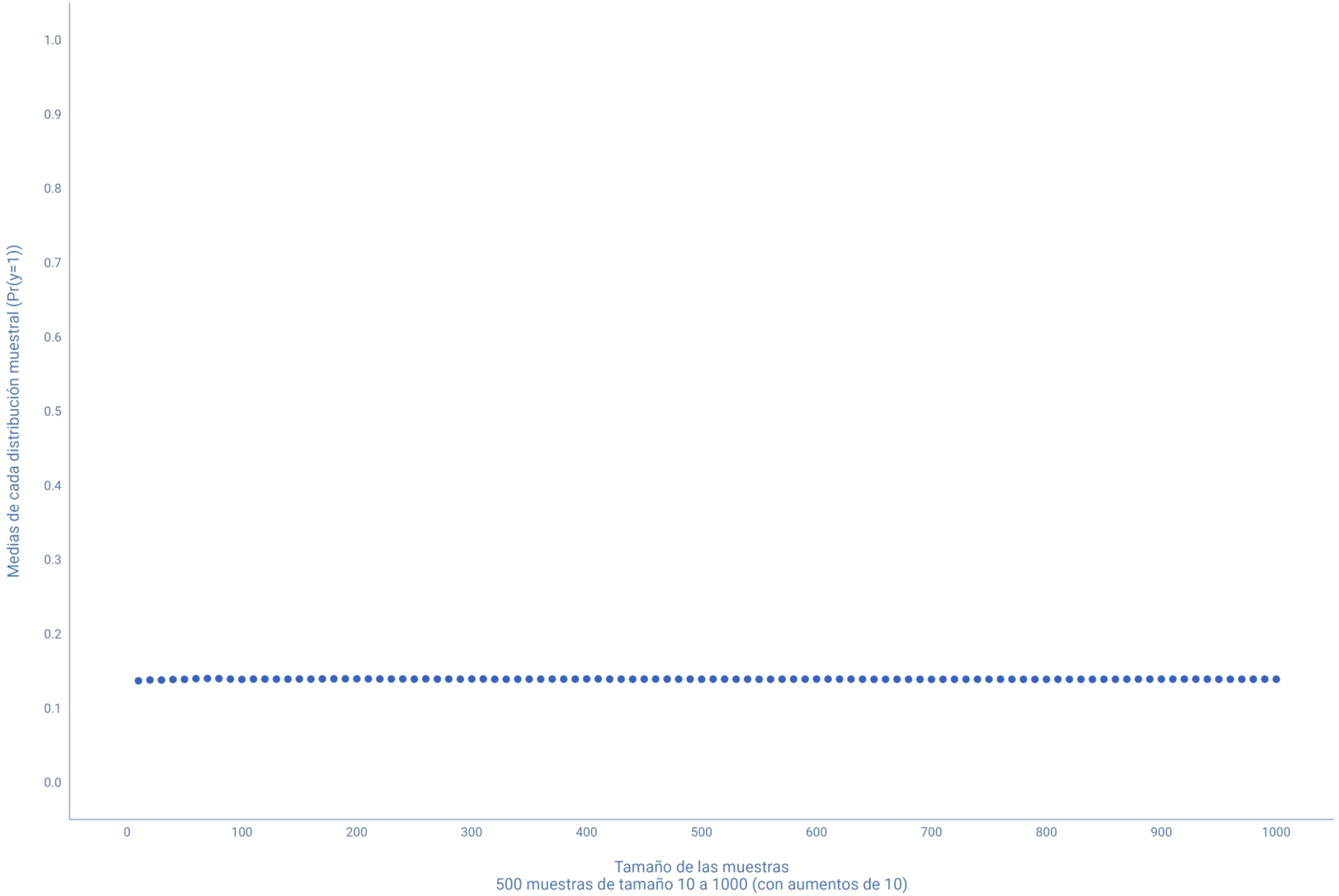


Mientras más grande sea una muestra aleatoria simple, menos se aleja del promedio de la población (i.e., la proporción de la población). Mientras más se acerque al infinito el tamaño muestral, la expectativa es que la media de cada muestra de mayor tamaño se aproxima al promedio de la población.

La versión más simple de la ley de los grandes números es que a medida que una muestra es más grande, su promedio se acerca al promedio de la población.

En la presente figura tenemos la media de cada muestra generada, condicionada por el tamaño muestral.

Ley de los grandes números y teorema del límite central

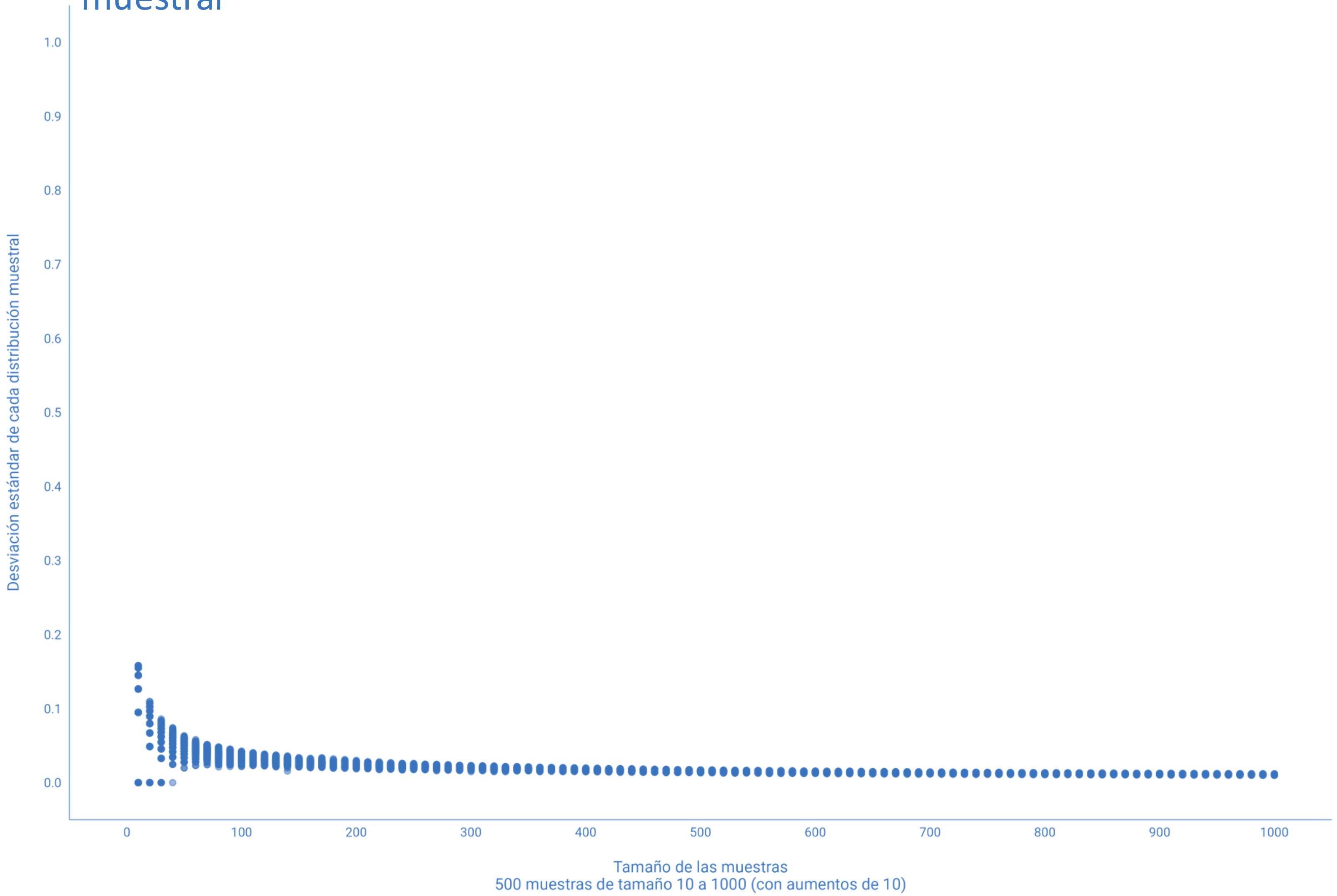


En la presente figura tenemos la media de cada distribución muestral, condicionada por el tamaño muestral.

El promedio de todos los promedios de cada muestra, es decir, el promedio de cada distribución muestral converge al promedio de la población.

Mientras la ley de los grandes números versa sobre el promedio de una muestra (condicional al tamaño); el teorema del límite central versa sobre la distribución muestral.

Dispersión de las medias de la distribución muestral, condicional al tamaño muestral



Entre muestras de 10 a 100, mientras mayor sea el tamaño muestral, la dispersión de las proporciones de cada distribución muestral decae rápidamente.

Entre 100 y 400 observaciones, sigue habiendo una menor dispersión, a medida que aumenta el tamaño de las muestras.

De 400 observaciones en adelante, el efecto del tamaño muestra sobre la precisión (i.e., dispersión de la distribución muestral marginal).

Ley de los grandes números

$$\lim_{n \to \infty} P(|\bar{X}_n - \mu| < \epsilon) = 1$$

Formalmente la formula indica:

Que tenemos una expectativa de que la probabilidad de que la media de una muestra tenga una distancia cero, con la media de la población, se acerca a 1, mientras más cerca del infinito este el tamaño de la muestra.

Genial. Ahora tenemos una primera intuición, que tan sesgada puede estar una muestra, depende su tamaño.

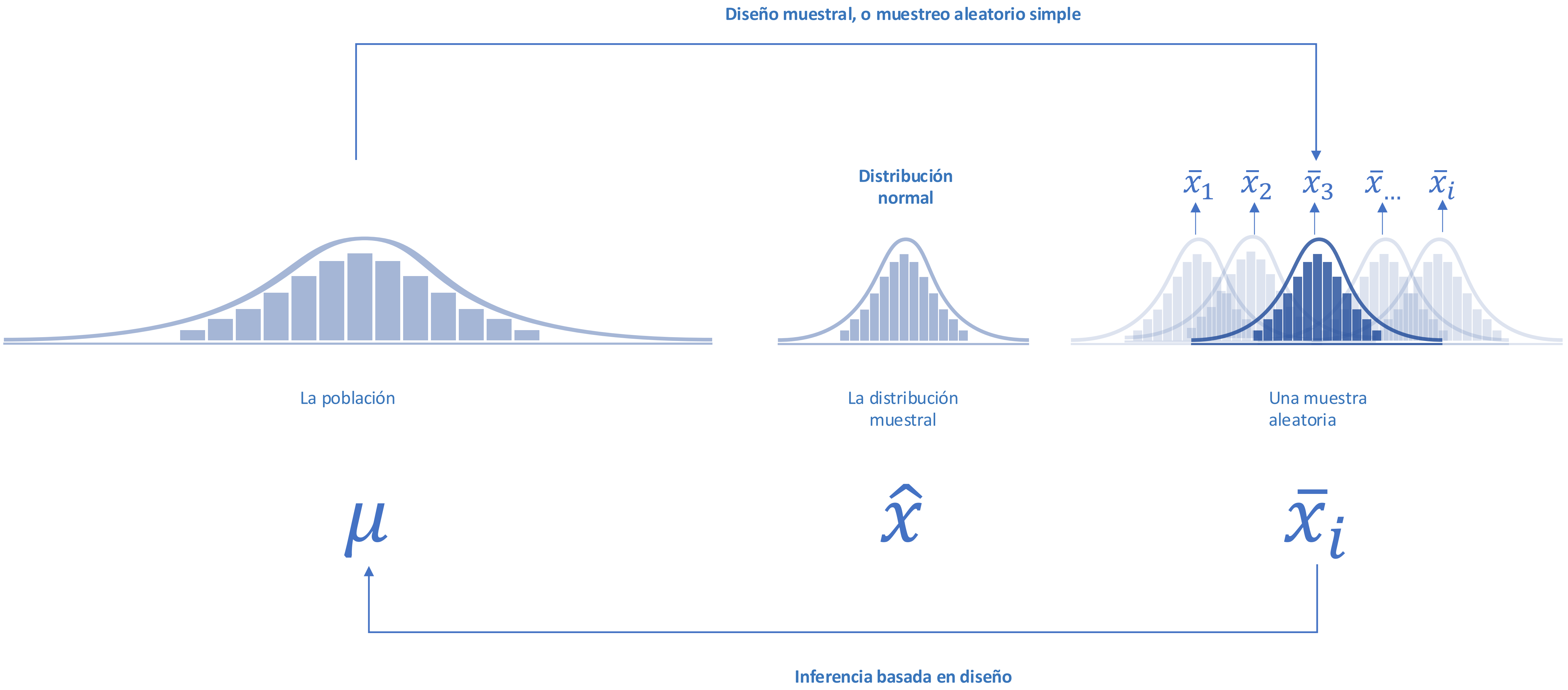
¿Cuál es el tamaño deseable?

Inferencia

Teorema del límite central

La distribución muestral de la media es normal

Muestra, Distribución Muestral, y Población



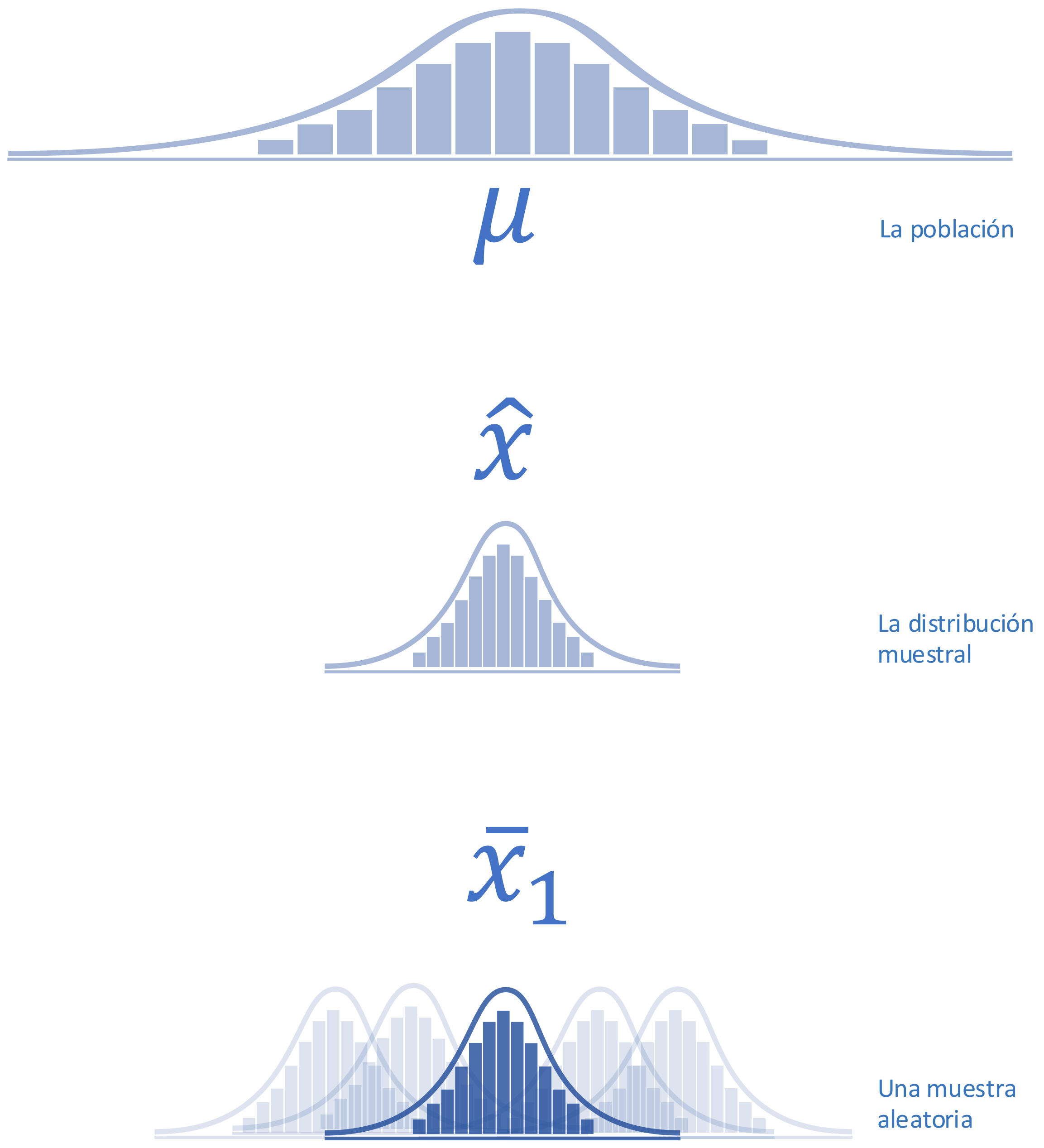
De estadígrafos a parámetros

Nuestro foco de inferencia son los **parámetros de una población**. Esta puede ser un conjunto de observaciones seleccionables (e.g., todos los estudiantes de sexto grado de Chile) (población finita).

Los parámetros de interés de una población pueden ser la media, una proporción, un total, un ratio, u otra cifra, como la varianza, y la desviación estándar. En la población, **el parámetro se simboliza con μ** (la letra griega “mu”). En la clase de hoy, vamos a partir con la idea de una media (una proporción).

Para poder realizar inferencias basadas en diseño, **descansamos en la idea de la distribución muestral del estadígrafo $\hat{x}$**. Este es el promedio de todas las medias de las muestras aleatorias posibles.

La distribución de medias, de todas las muestras posibles toma una forma particular, es **una distribución normal**.



Distribución Muestral de un estadígrafo

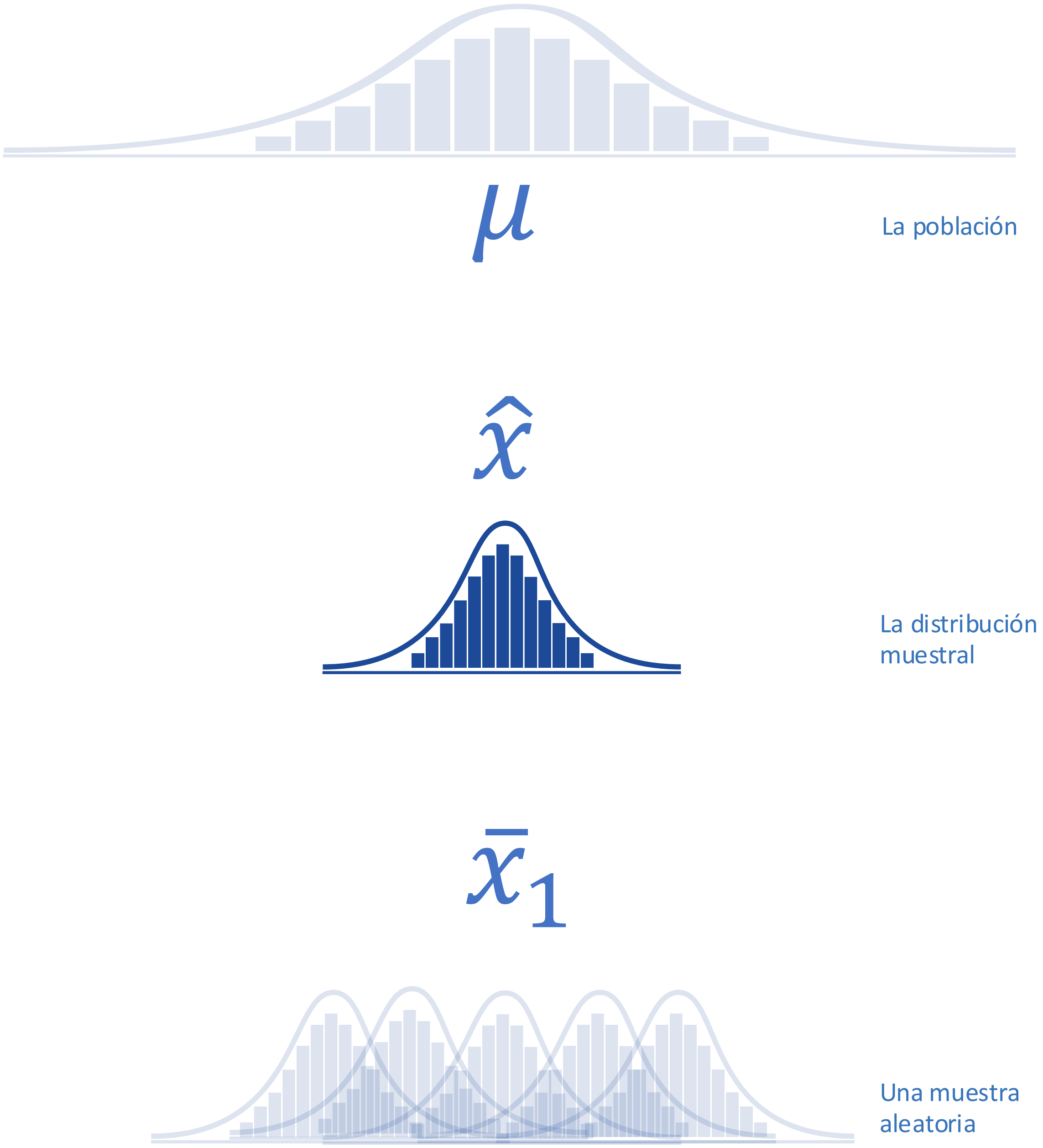
La distribución muestral de la media es el **promedio de todas las medias, de cada muestra**.

Podemos aproximarnos a esta distribución, por medio de generar una serie de muestras aleatorias simples (con reemplazo). Es decir, un conjunto de muestras que poseen la misma probabilidad de ser seleccionadas. Calcular sus medias, y luego calcular la media de todas las muestras.

Esta distribución, **siempre toma una forma normal**. Esta afirmación, a veces es formalizada con la siguiente expresión (Devlin et al., 2018):

$$\hat{x} \xrightarrow{iid} N(\mu, \frac{\sigma^2}{n})$$

Nota: Esta notación indica que el promedio de las medias de muchas muestras (aleatorias simples), toma como media al parámetro de la población, y posee una varianza conocida, y una forma normal.

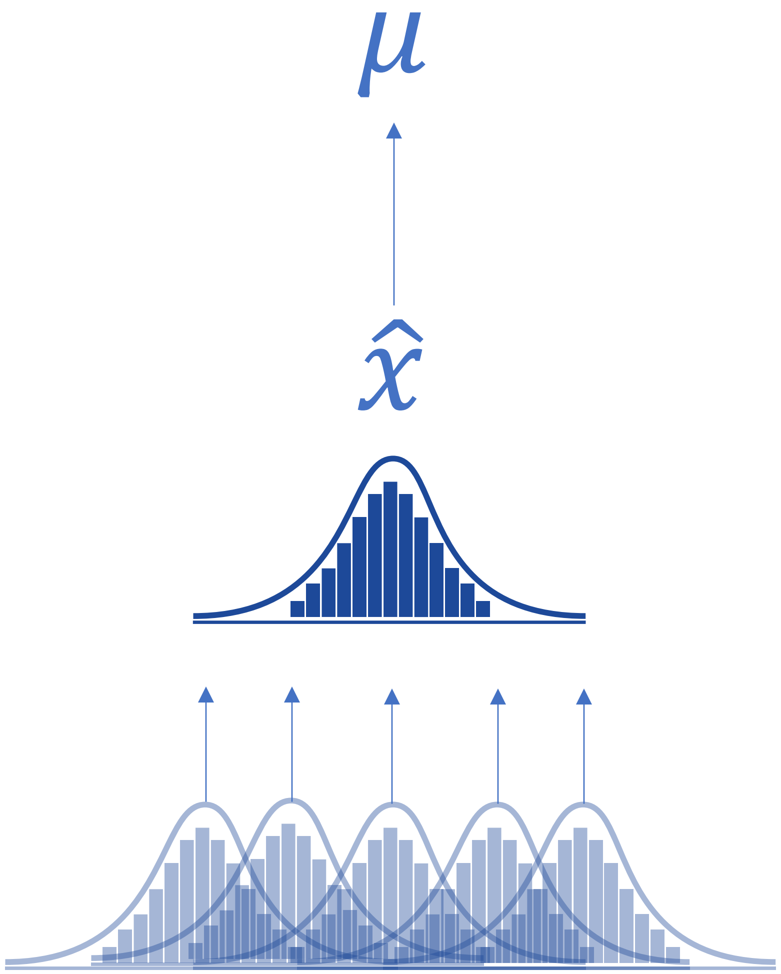


Teorema del límite central

En términos verbales, la idea de que la distribución muestral de la media tome una forma normal, es conocida como el **teorema del límite central**.

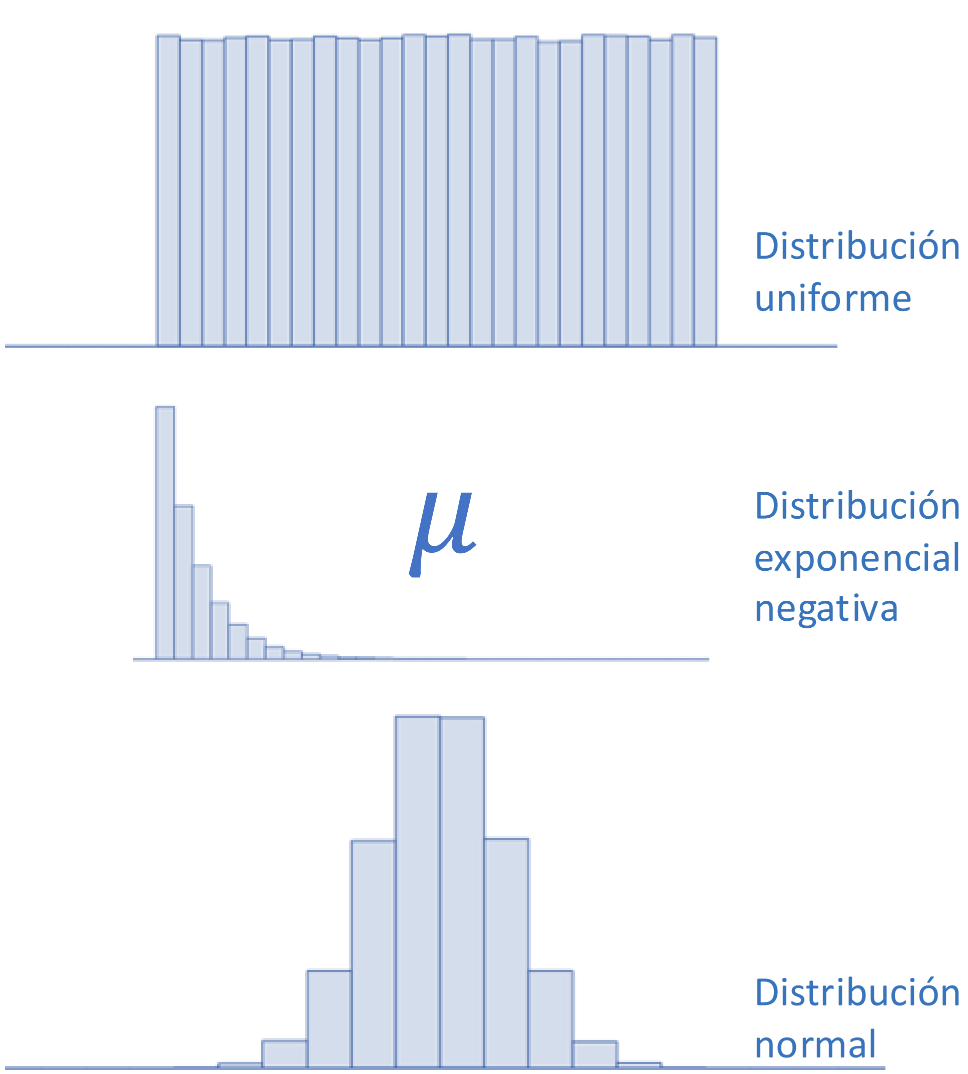
A medida que el **tamaño muestral crece, la distribución de medias**, de todas las muestras posibles, **toma una forma normal** (Huck, 2009), sin importar la forma original de la distribución de la población (Devlin et al., 2018).

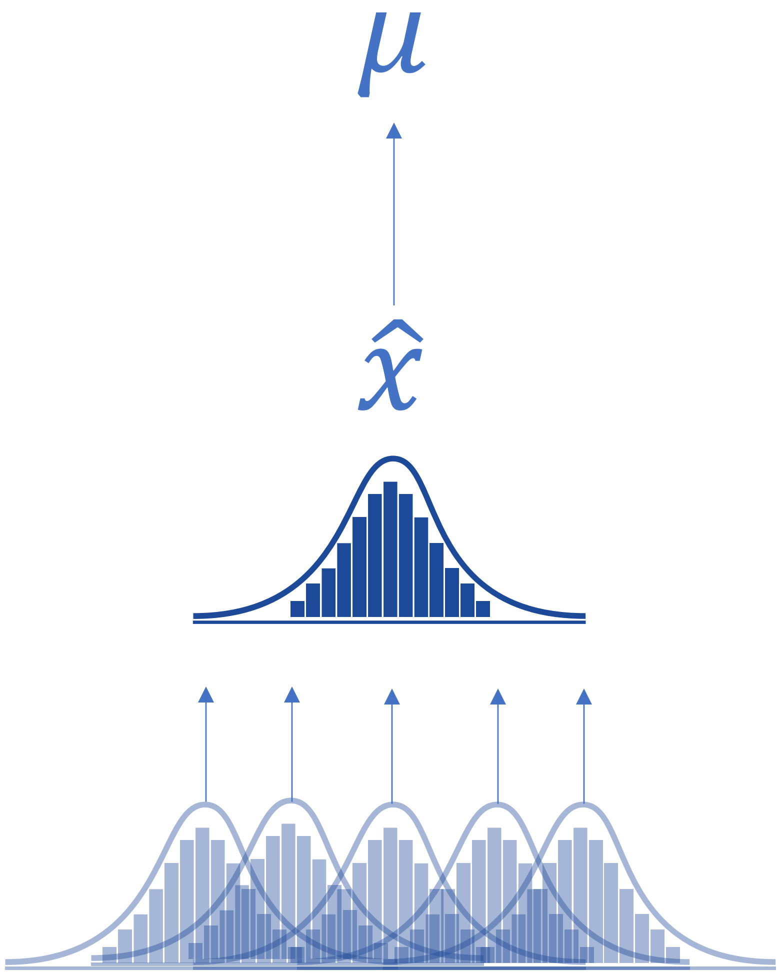
No obstante, un aspecto a considerar es que la **varianza** de la distribución muestral de los estadígrafos **disminuye** en la medida de que el **tamaño muestral es mayor**. Para el ejercicio de realizar inferencias con una sola muestra aleatoria, esto es importante. Porque implica que, muestras de tamaños más grandes (e.g., 100, 200, 300, 400 casos), provienen de distribuciones muestrales con menos dispersión, en contraste a muestras de menor tamaño (e.g., 10, 20, 30).



Ya sabemos que, la media de todas las muestras posibles (distribución muestral), converge a la media de la población. Lo vimos en la simulación de los grandes números.

Teorema del límite central





Otro aspecto relevante: **No importa la forma de la distribución de la población**, la distribución muestral **nos permite recuperar parámetros de la población**.

Dispersión de la distribución muestral

La varianza de la distribución muestral tiene como expresión a sigma cuadrado, partido por el tamaño muestral.

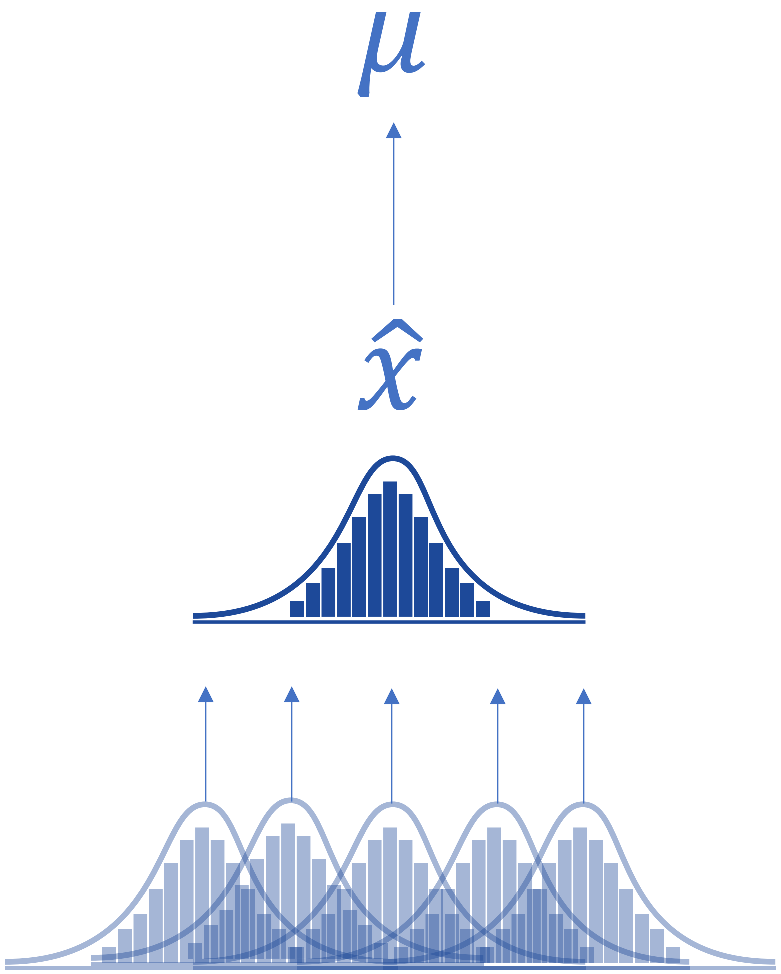
$$\hat{x} \xrightarrow{iid} N(\mu, \frac{\sigma^2}{n})$$

La desviación estándar de la distribución muestral de la media tiene por expresión a

$$\frac{\sigma}{\sqrt{n}}$$

La desviación estándar de la distribución muestral de las proporciones tiene por expresión a

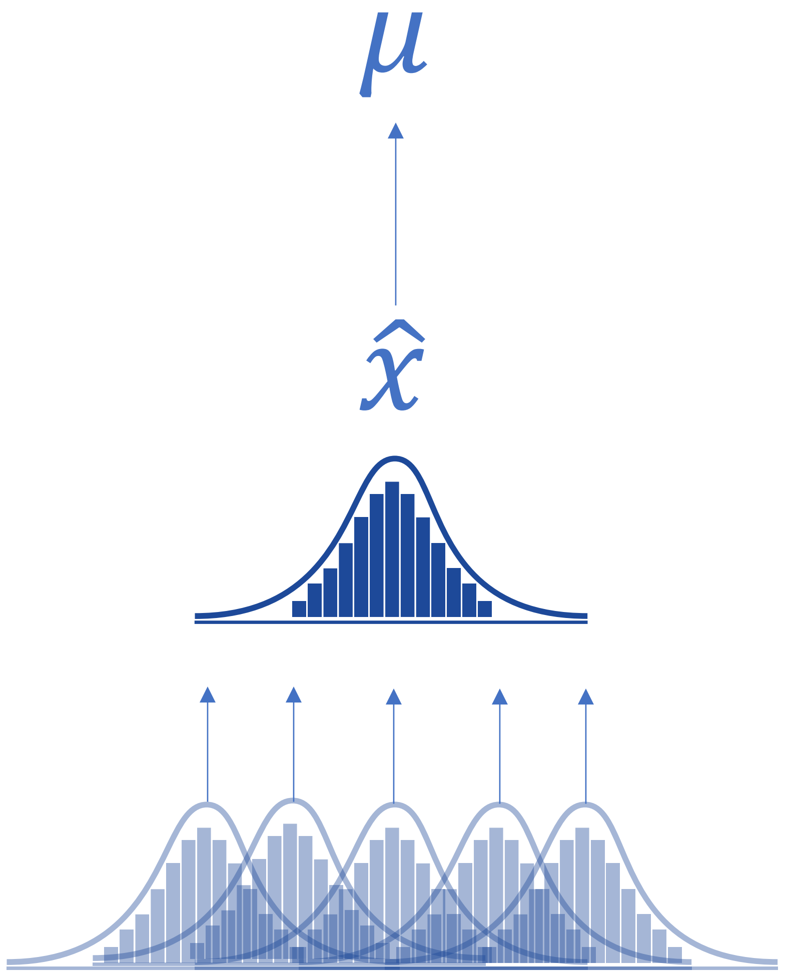
$$\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$



No importa la forma de la distribución de la población, la distribución muestral de la media toma una forma conocida.

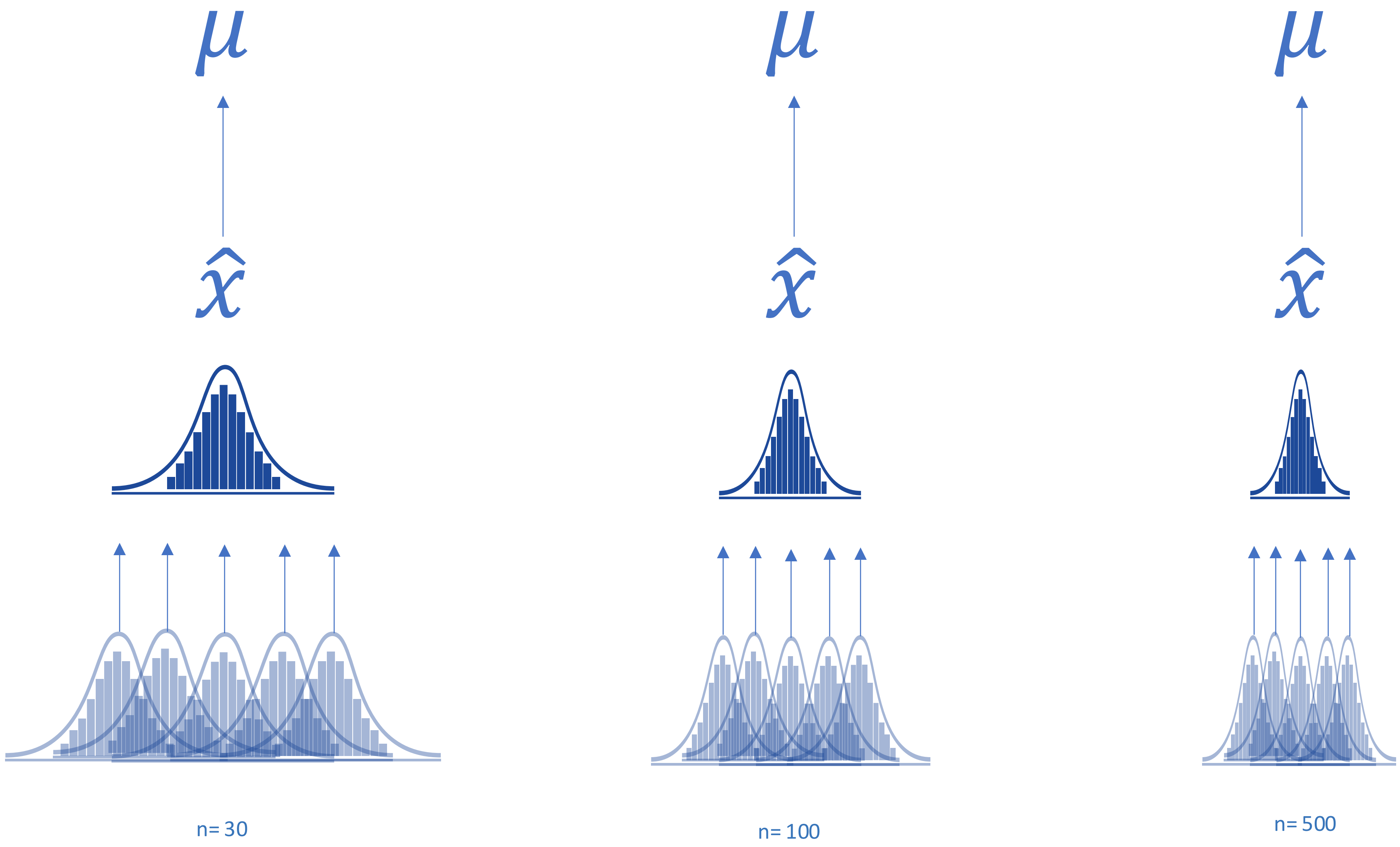
En consecuencia, la precisión de una muestra aleatoria no se incrementa de forma lineal.

Dicho de otro modo, si se quiere duplicar la precisión de una muestra aleatoria, **el tamaño muestral podría requerir ser cuadruplicado** (Stigler, 2016).



No importa la forma de la distribución de la población, la distribución muestral de la media toma una forma conocida.

Dispersión de la distribución muestral



A medida que el tamaño muestral aumenta, la dispersión de la distribución muestral de los estadígrafos se hace más estrecha.

En otras palabras, la distribución del estadígrafo (i.e., medias, proporciones), de la distribución muestral, es menos dispersa.

En consecuencia, se plantea que, muestras aleatorias de mayor tamaño podrían ser más precisas para realizar inferencias a la población.

Inferencia

Forma de la distribución muestral

La distribución muestral de la media es una distribución normal

Teorema del límite central y distribución muestral

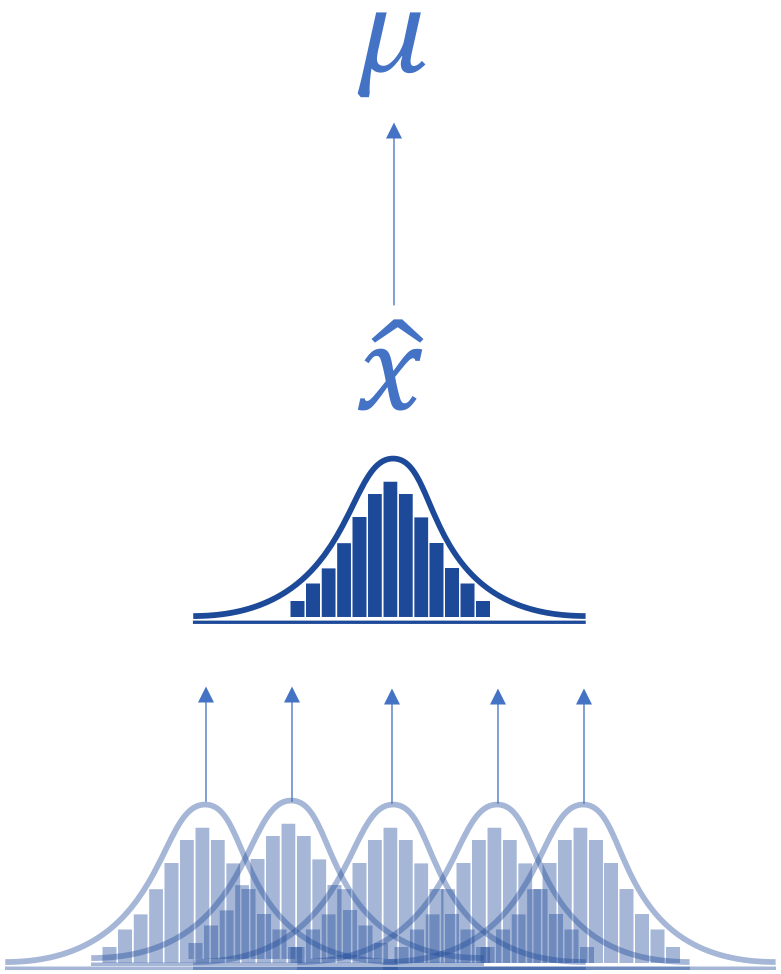
Tenemos cuatro ideas centrales:

Mientras **más grande una muestra**, más chances tiene de acercarse al promedio de la población.

El **promedio de la distribución muestral**, de seguro se acerca al promedio de la población.

La media de una muestra se aleja del promedio de la distribución muestral con una **dispersión conocida** (i.e., el error estándar de la media, o desviación estándar de las medias en la distribución muestral).

Además, **sin importar la forma de la distribución de la población** las tres ideas anteriores se sostienen para las medias y las proporciones.



No importa la forma de la distribución de la población, la distribución muestral **nos permite recuperar parámetros de la población**.

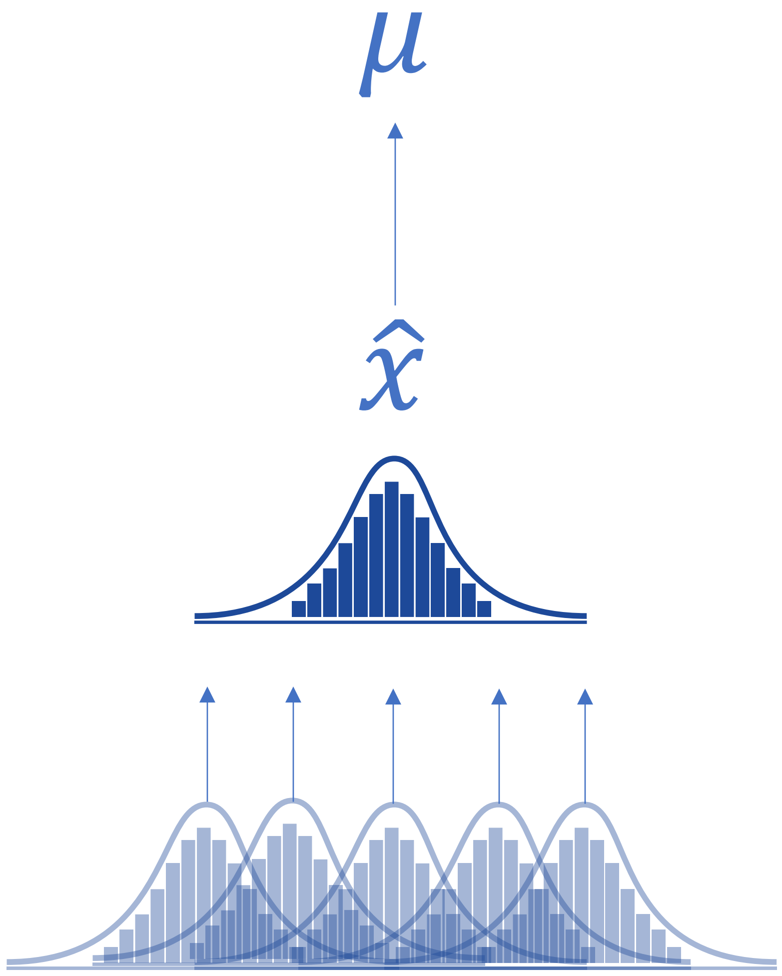
Distribución muestral y distribución normal

Uno de los apoyos centrales de la inferencia basada en diseño, es que la **distribución muestral de los estadígrafos toma una forma**.

Y, por tanto, podemos emplear una **distribución probabilística** para realizar inferencias (i.e., algo que represente a la forma de la distribución muestral). Las distribuciones probabilísticas las podemos pensar como histogramas de distribuciones ideales, que tienen un área bajo la curva que es igual a uno. Y con esta característica podemos saber cuál es la proporción de casos esperados bajo o sobre cada valor de la distribución.

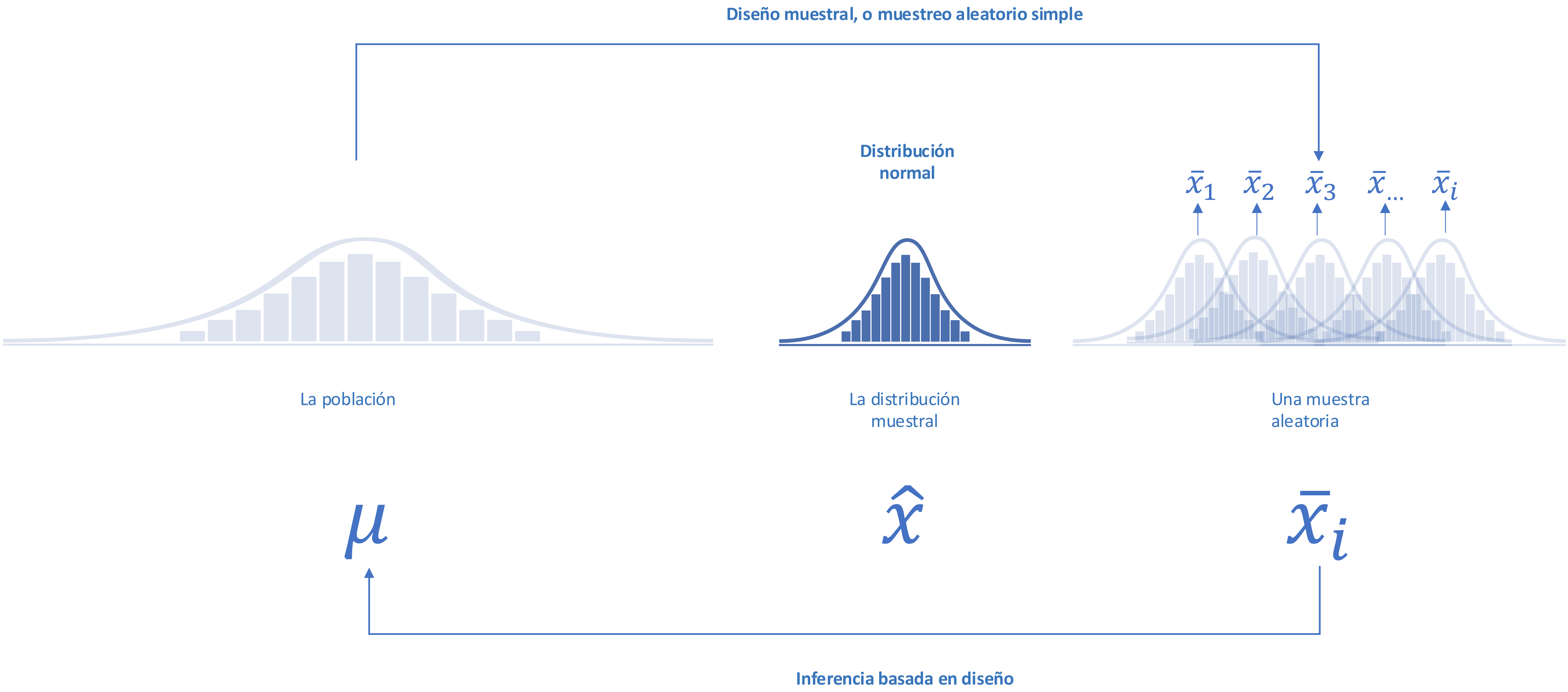
La **distribución normal estandarizada**, es una distribución de este tipo. Podemos emplear lo que sabemos de esta distribución para realizar inferencias con una muestral aleatoria.

Y podemos proceder de una forma muy parecida, con otras distribuciones probabilísticas (distribuciones z, t, y F).

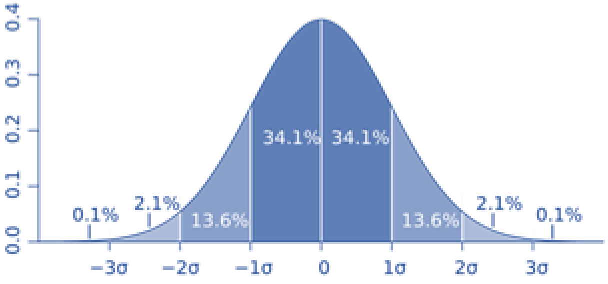


No importa la forma de la distribución de la población, la distribución muestral **nos permite recuperar parámetros de la población**.

La distribución muestral de la media conforma a una distribución normal

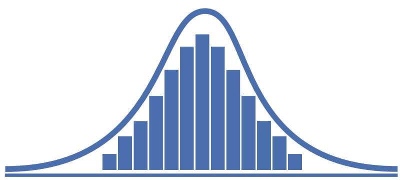


La distribución muestral de la media conforma a una distribución normal

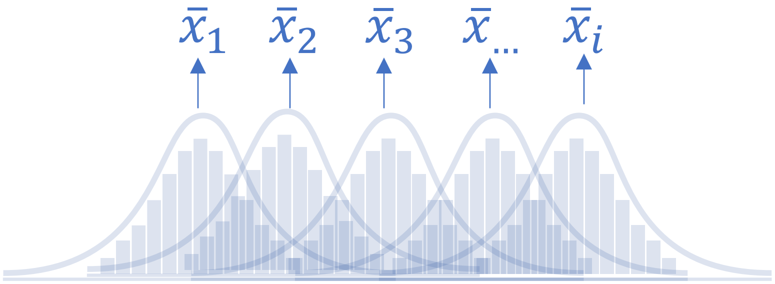


Distribución normal estandarizada

Distribución normal



La distribución muestral



Una muestra aleatoria

$$\mu \qquad \hat{x} \qquad \bar{x}_i$$

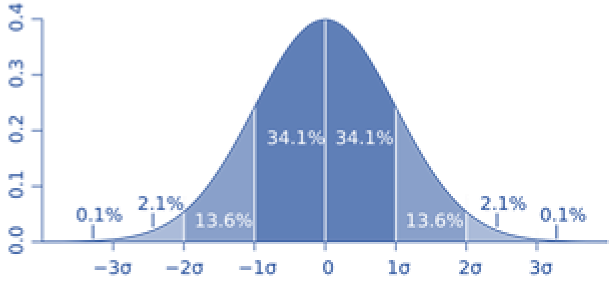
El parámetro poblacional, toma la posición de la media en la distribución normal estandarizada, donde esta es un valor cero. Recordemos que el promedio de las distribuciones muestrales converge a μ

Inferencia basada en diseño

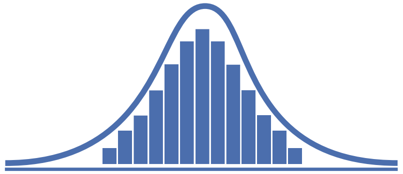
La distribución muestral de la media conforma a una distribución normal

Podemos transformar a la distribución muestral, a una distribución normal estandarizada

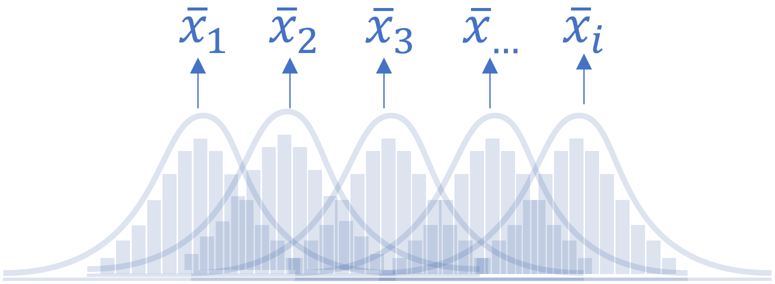
Distribución normal



Distribución normal estandarizada



La distribución muestral



Una muestra aleatoria

Podemos transformar los valores de una distribución muestral, a puntajes z.

$$\mu$$

$$\hat{x}$$

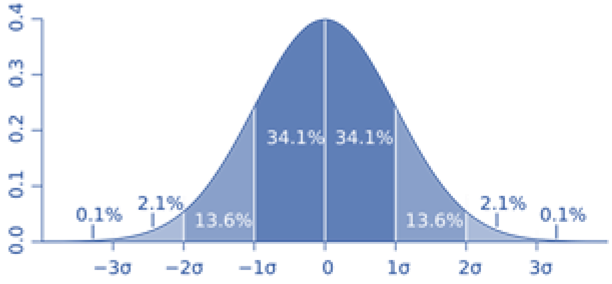
$$\bar{x}_i$$

Inferencia basada en diseño

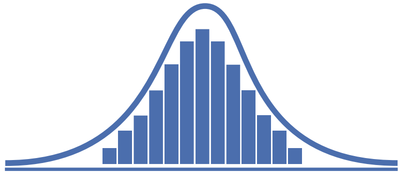
La distribución muestral de la media como distribución normal estandarizada

```
# r code
mutate(z_score = (raw_score – mean(raw_score))/sd(raw_score))
```

Distribución normal

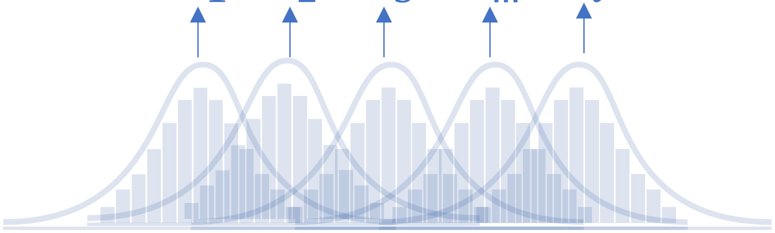


Distribución normal estandarizada



La distribución muestral

$\bar{x}_1$ $\bar{x}_2$ $\bar{x}_3$ $\bar{x}_{\dots}$ $\bar{x}_i$



Una muestra aleatoria

μ $\hat{x}$ $\bar{x}_i$

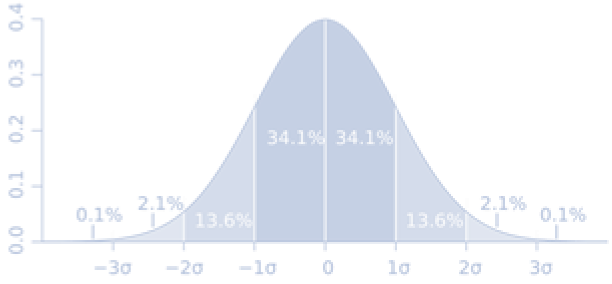
Inferencia basada en diseño

Podemos transformar los valores de una distribución muestral, a puntajes z.

Así como podemos transformar los valores de la distribución muestral a puntajes z, entonces por extensión, podemos hacer lo mismo con una de las muestras aleatorias.

La distribución muestral como distribución normal estandarizada

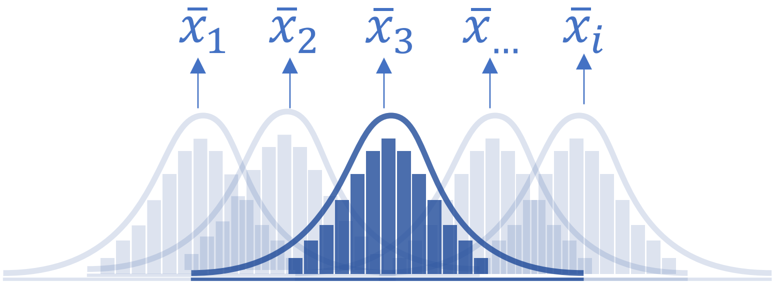
```
# r code
mutate(z_score = (raw_score – mean(raw_score))/sd(raw_score)
```



Distribución normal estandarizada

Distribución normal

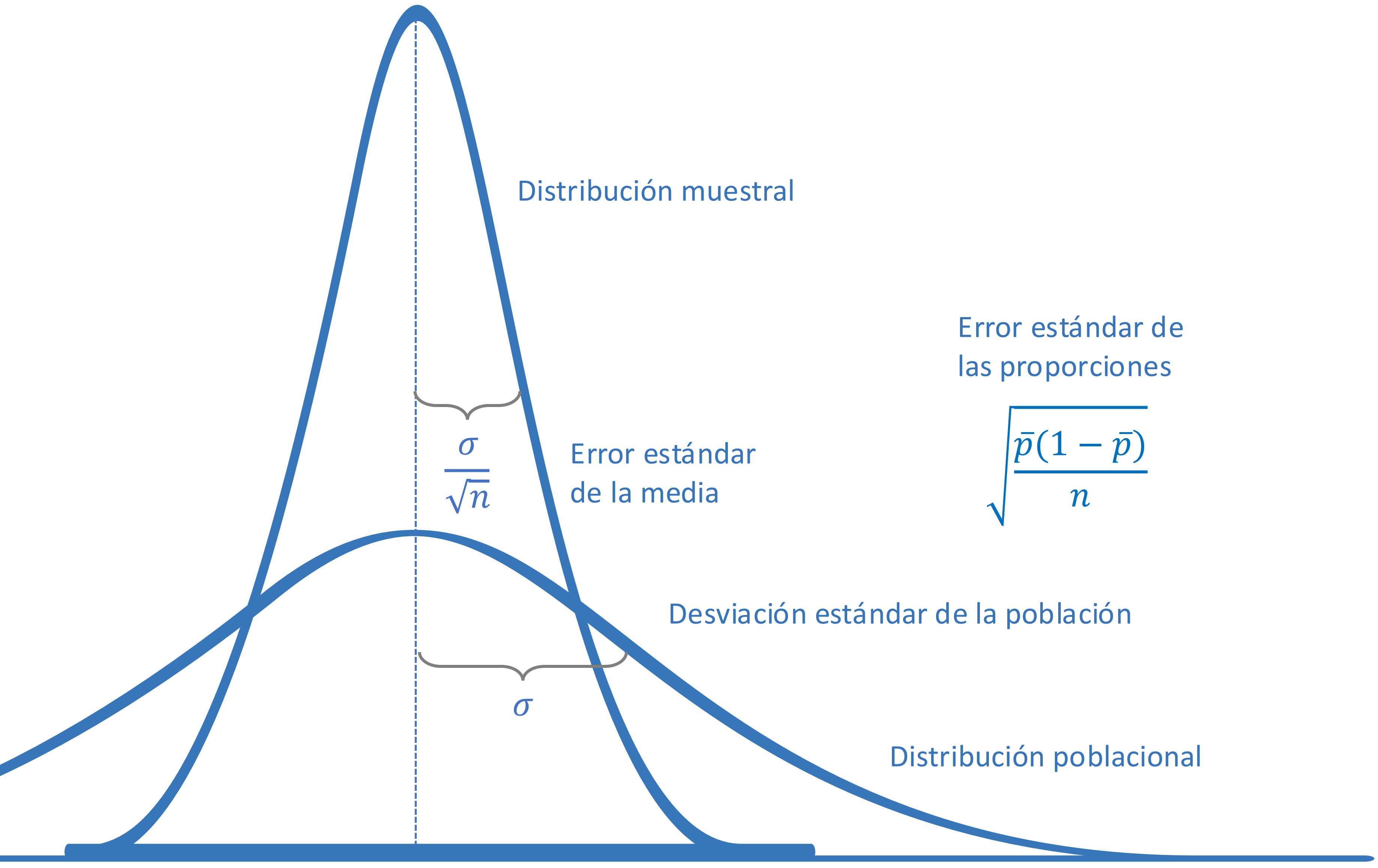
La distribución muestral



Una muestra aleatoria

Si transformamos los valores de una muestra aleatoria a puntajes z, entonces podemos inferir su posición relativa con respecto al parámetro de interés.

Empleando lo que sabemos de la distribución muestral, y lo que sabemos de la distribución normal estandarizada, podemos tomar una muestra aleatoria y conjeturar una expectativa razonable respecto a la posición del parámetro de la población.

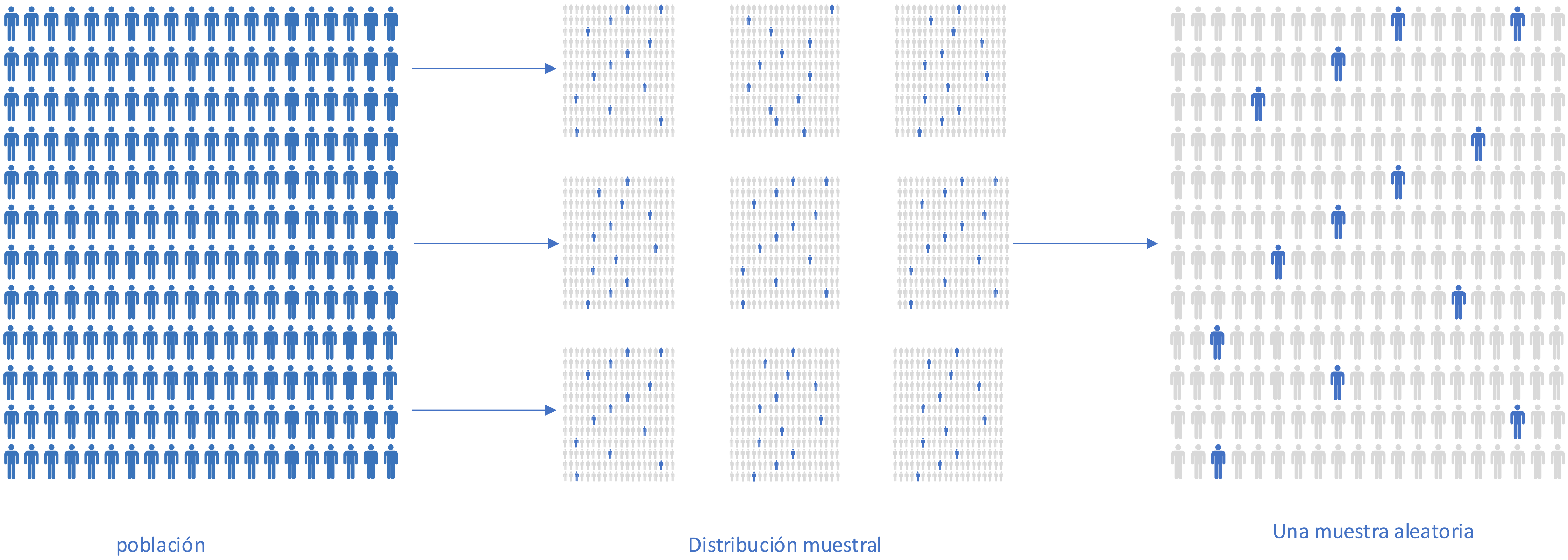


Inferencia

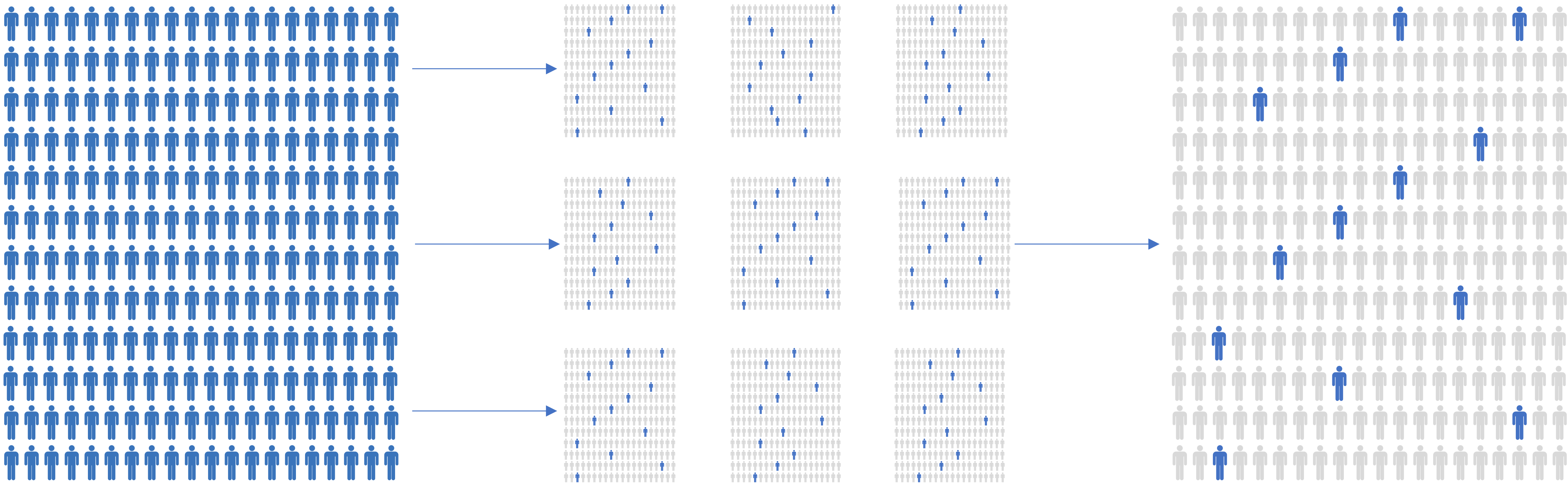
Sesgo de las muestras

Posición relativa de las muestras aleatorias respecto al parámetro de interés

De la población, a la distribución muestral, a una muestra aleatoria



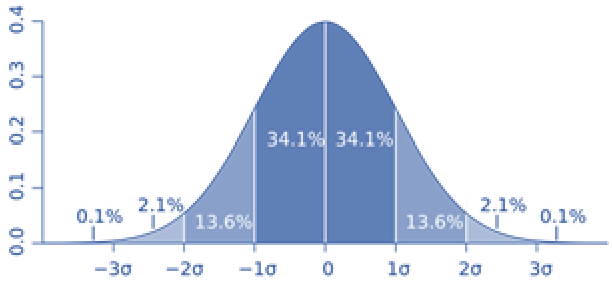
De la población, a la distribución muestral, a una muestra aleatoria

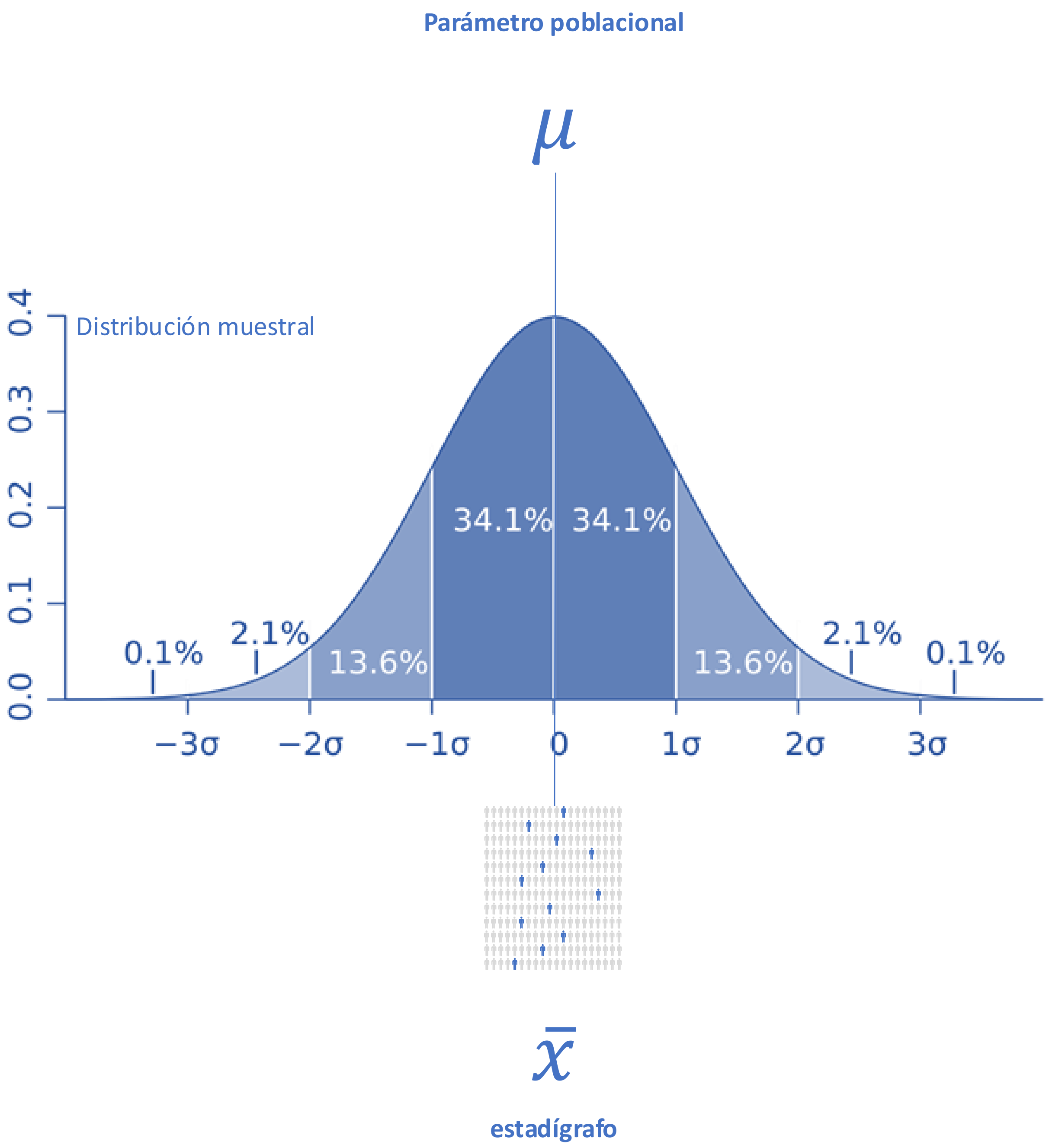


población

Distribución muestral

Una muestra aleatoria



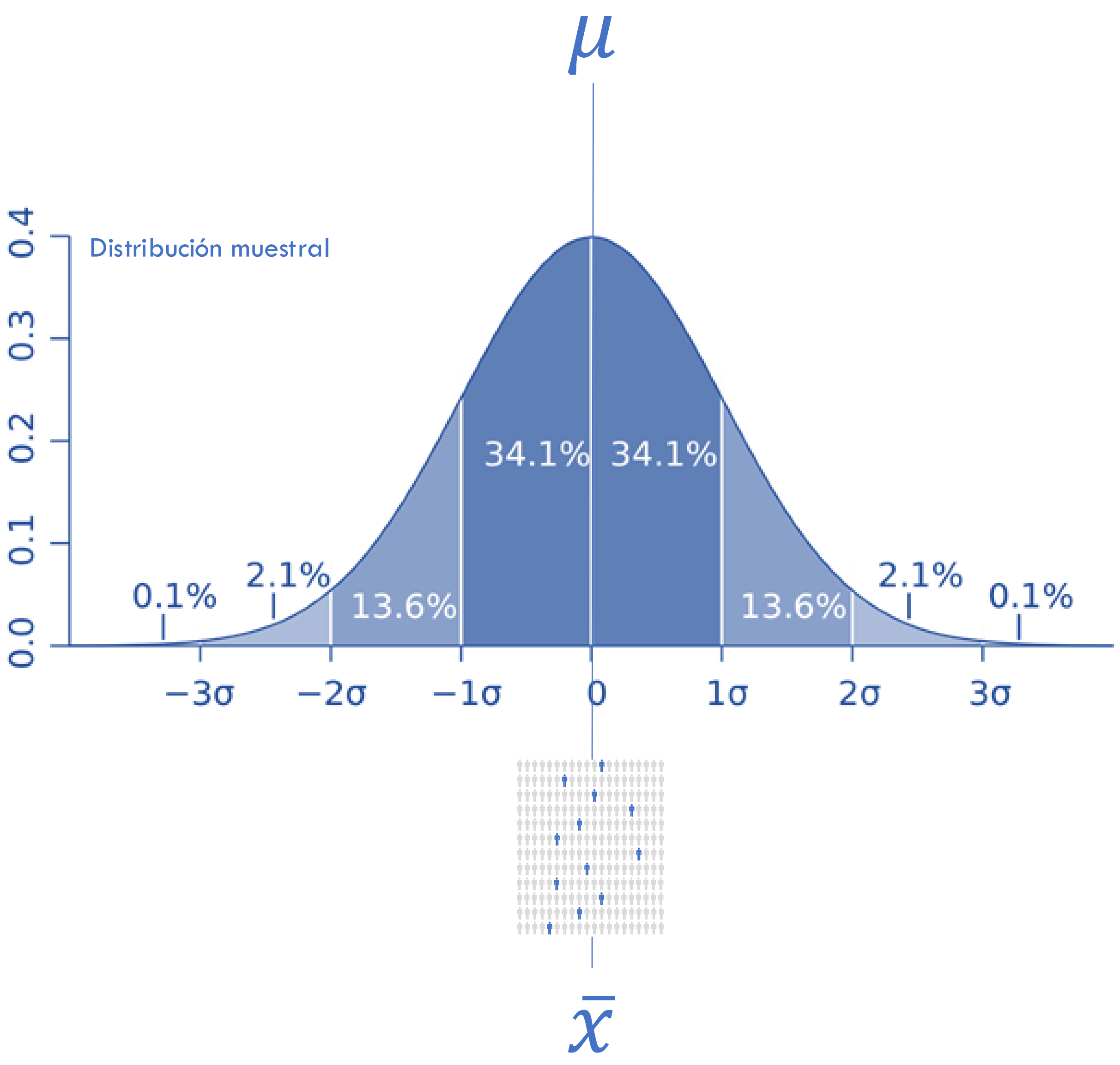


Sabemos que la media de la distribución muestral converge con el parámetro de la población.

Y que la media de una muestra aleatoria podría tomar cualquier posición con respecto a la distribución muestral.

Y que debido a que esta distribución es normal, entonces podemos expresar su desviación con respecto al parámetro normal empleando a la distribución normal estandarizada como referencia.

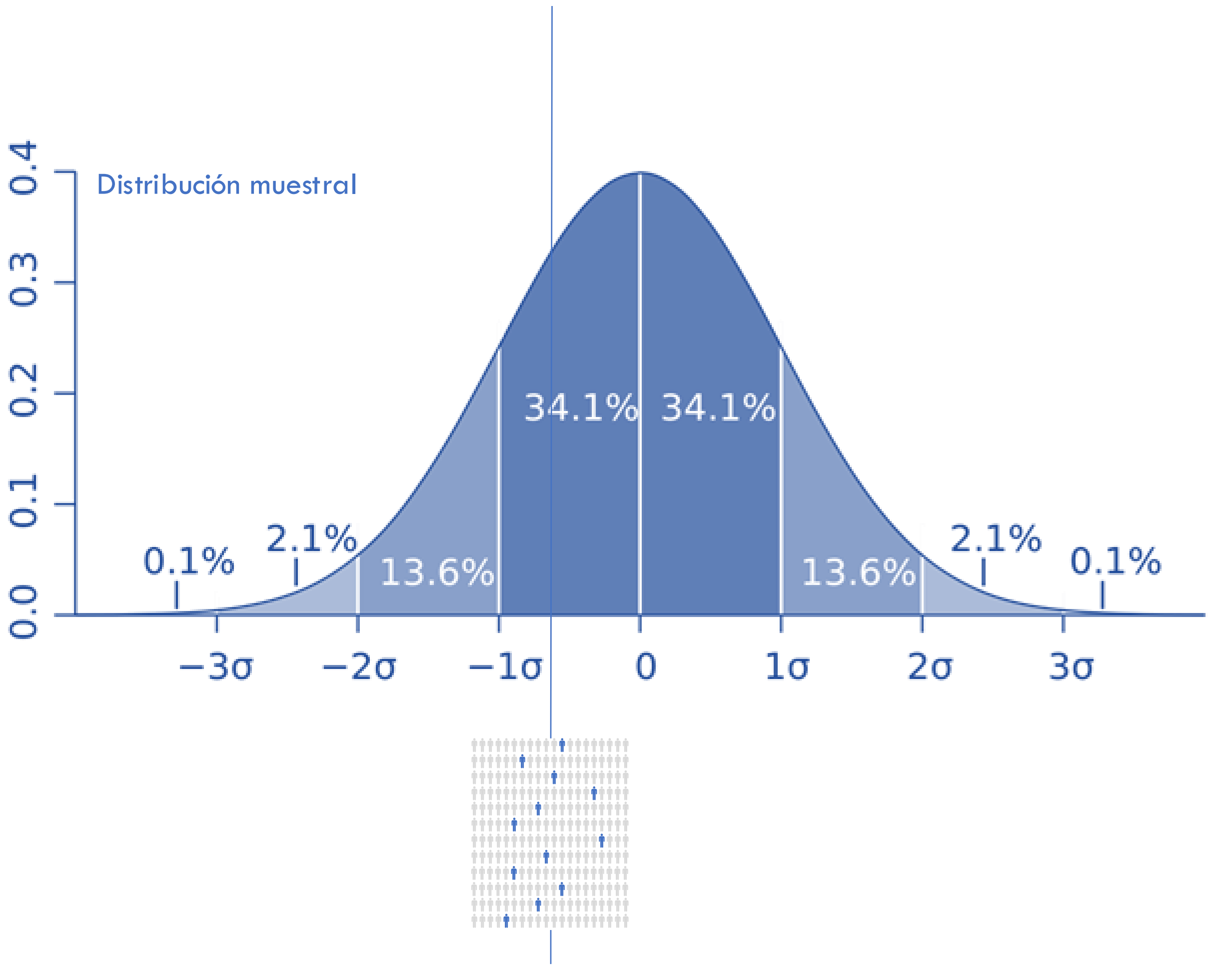
Parámetro poblacional



estadígrafo

Parámetro poblacional

μ

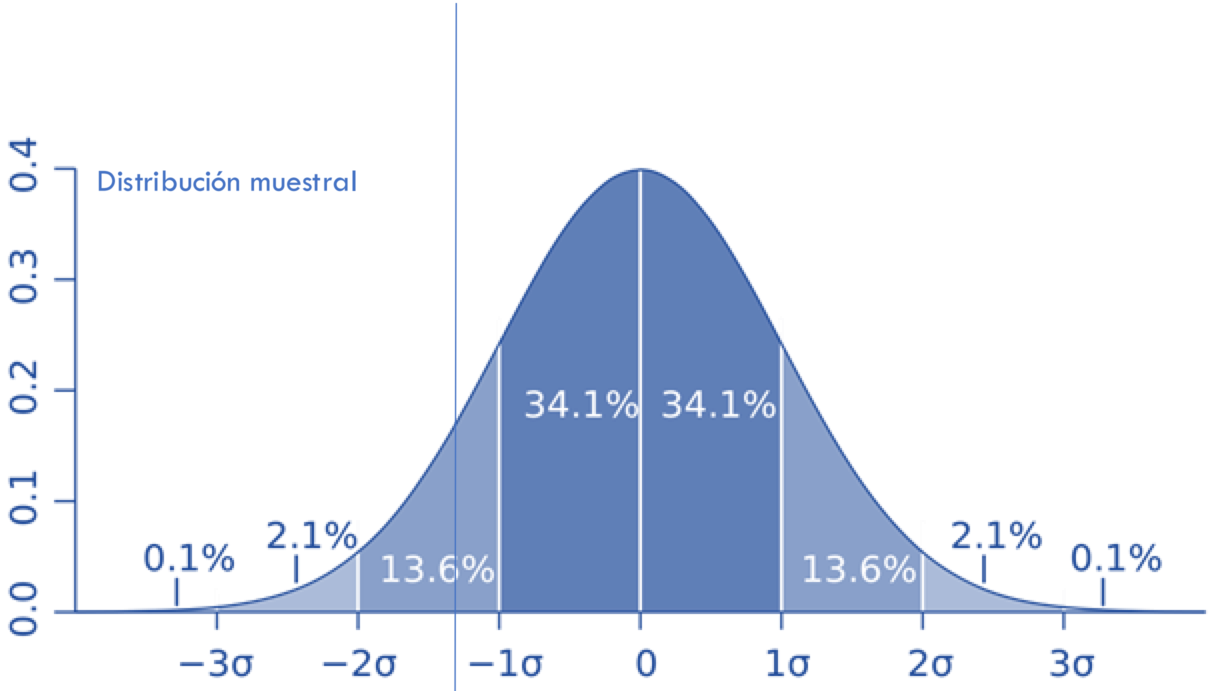


$\bar{x}$

estadígrafo

Parámetro poblacional

μ

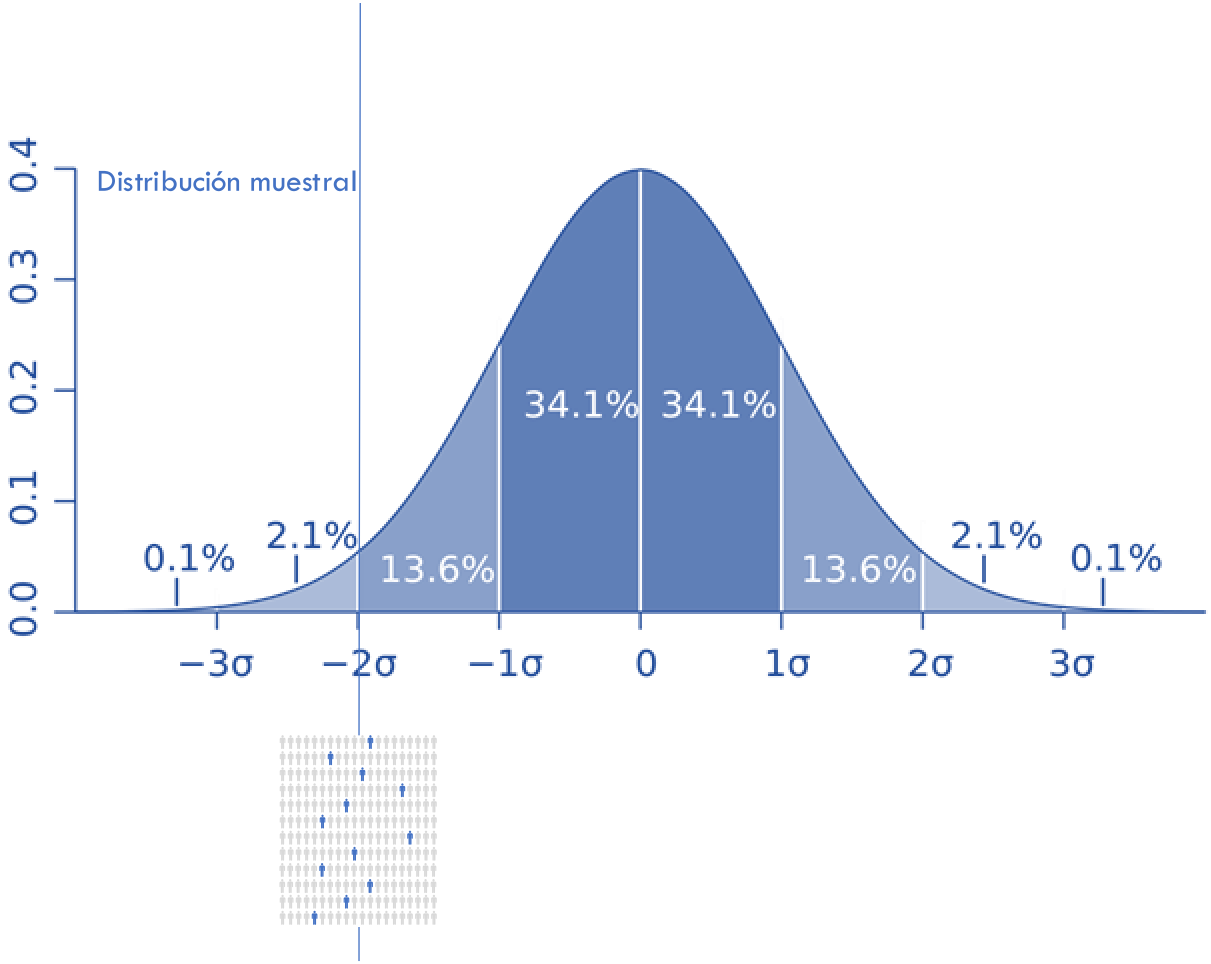


$\bar{x}$

estadígrafo

Parámetro poblacional

μ

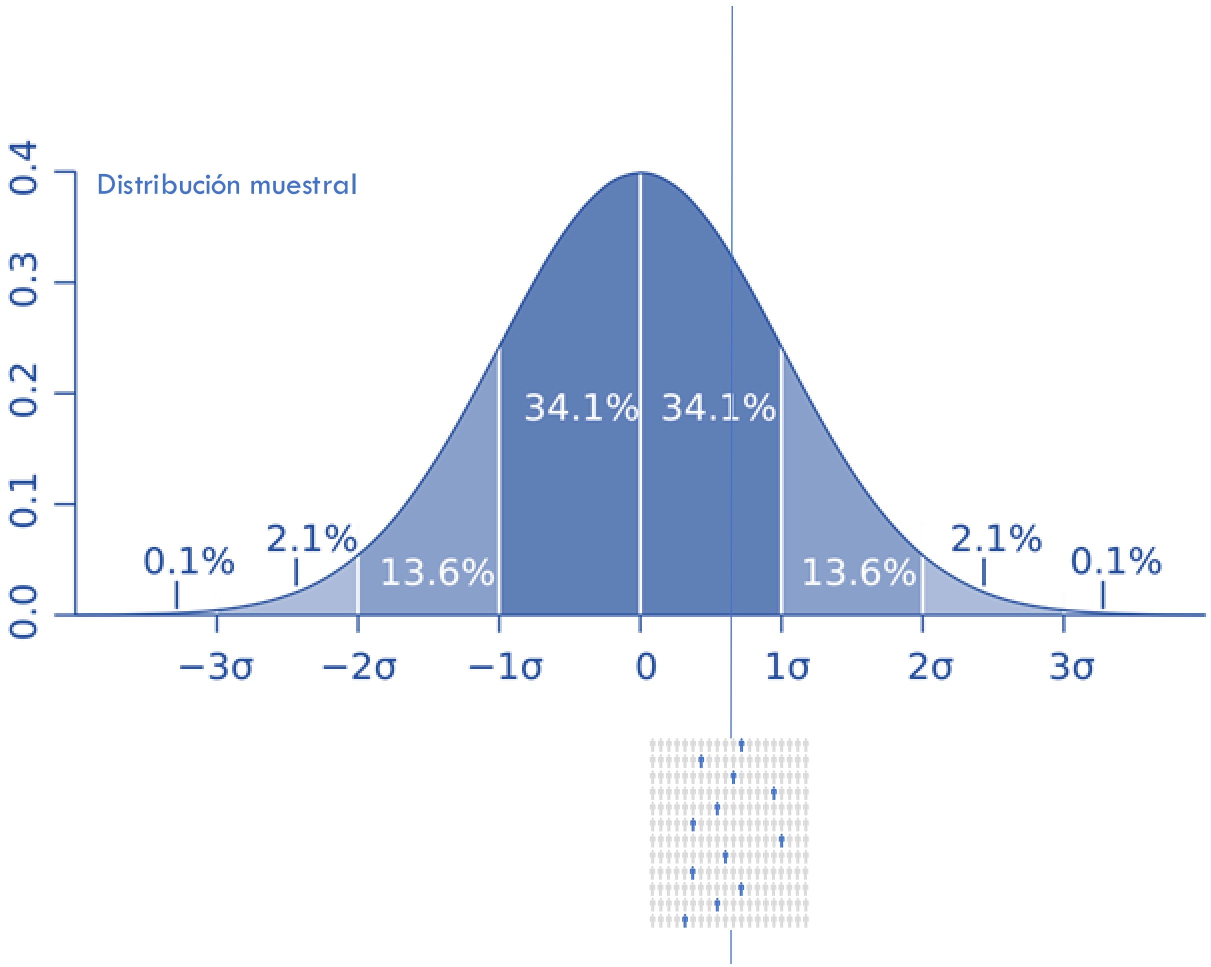


$\bar{x}$

estadígrafo

Parámetro poblacional

μ

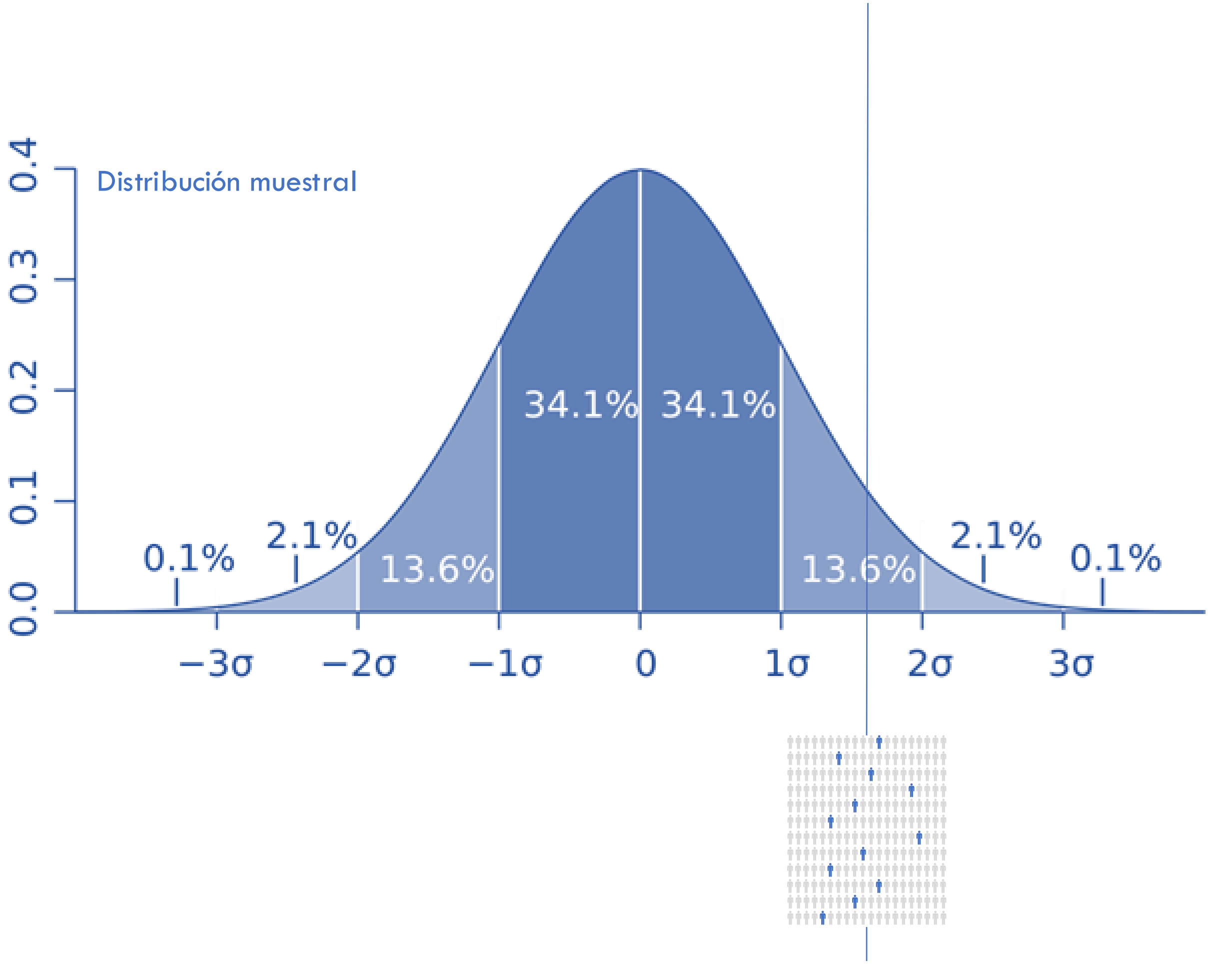


$\bar{x}$

estadígrafo

Parámetro poblacional

μ

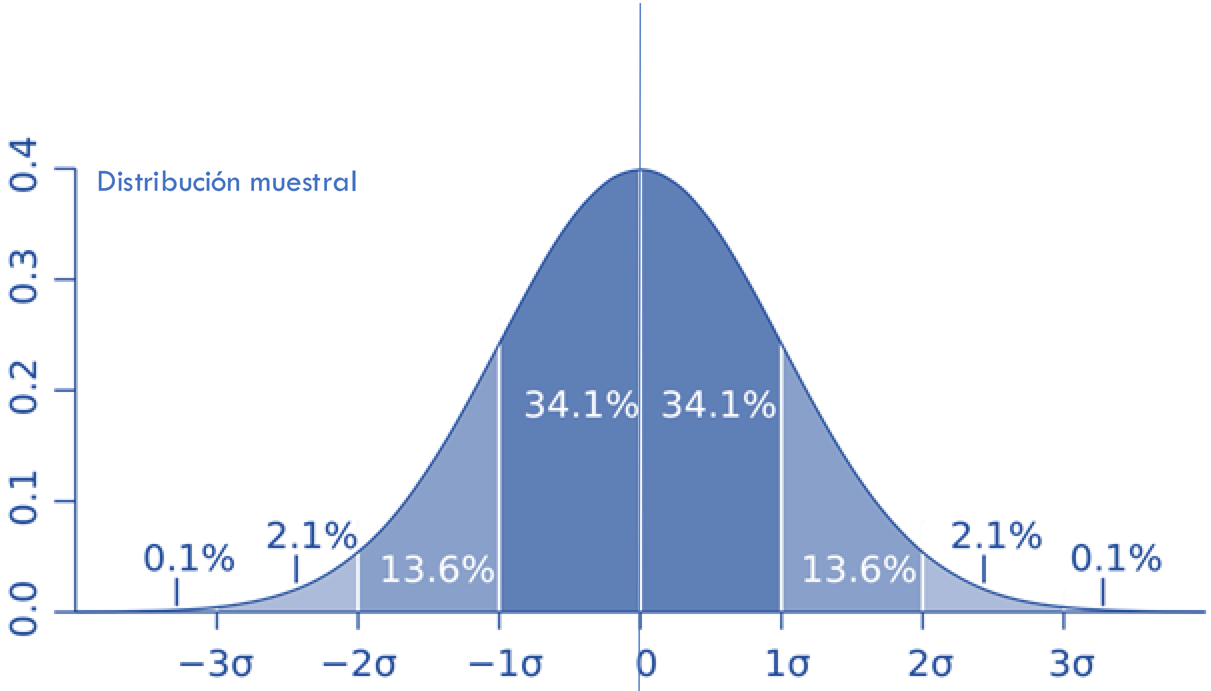


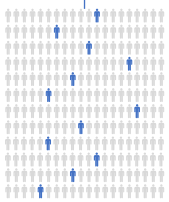
$\bar{x}$

estadígrafo

Parámetro poblacional

μ





$\bar{x}$

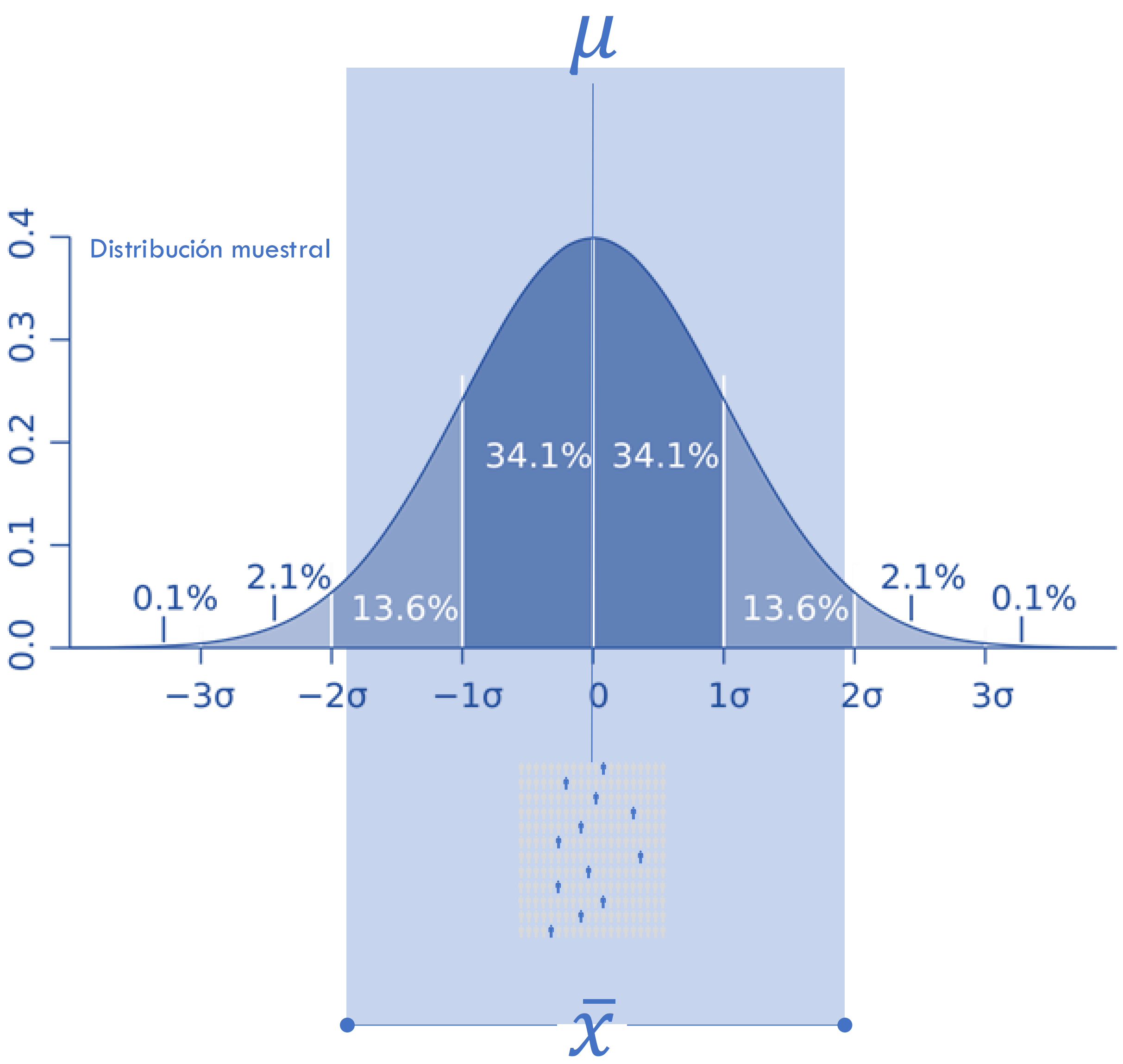
estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Parámetro poblacional



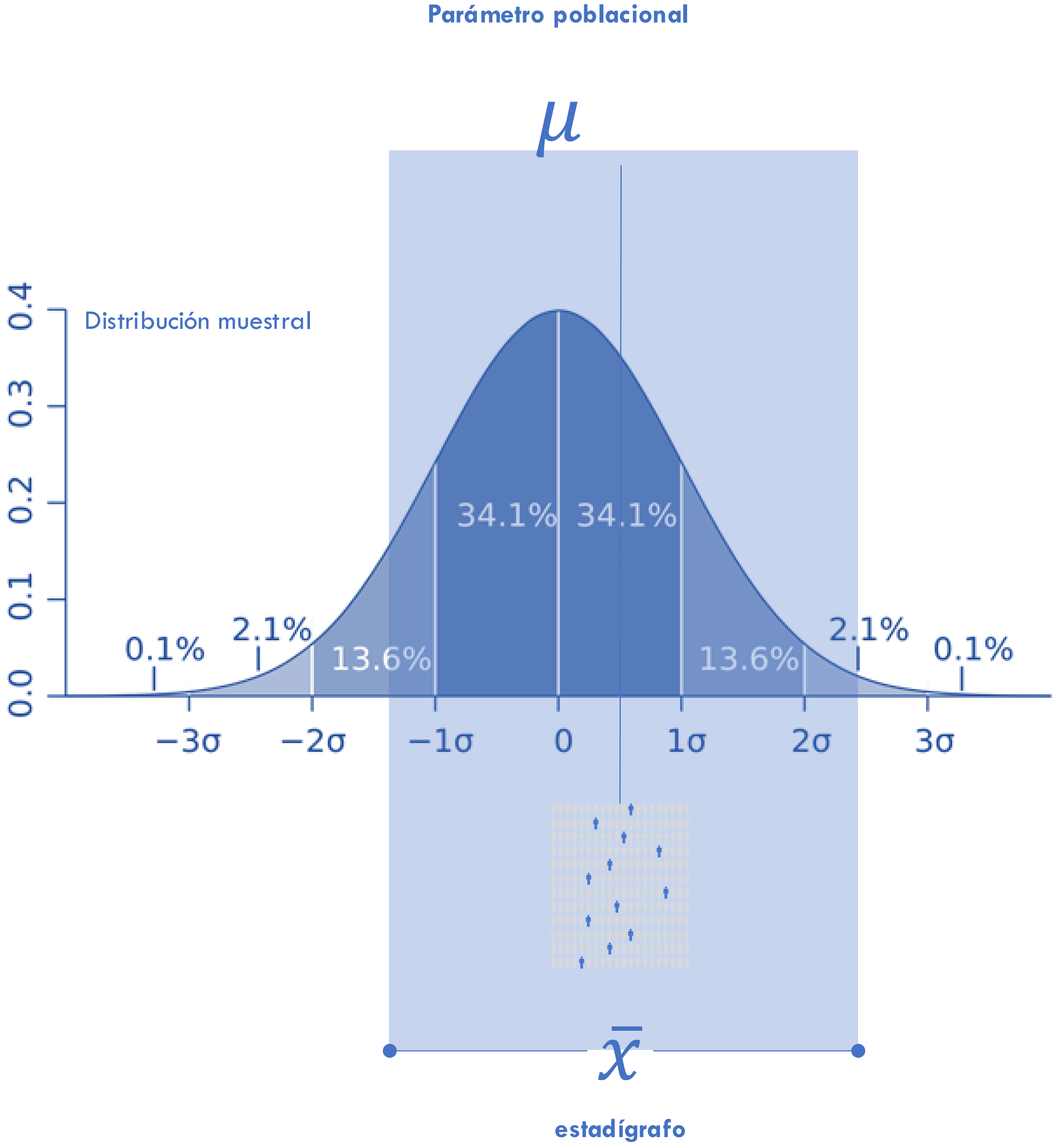
estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.



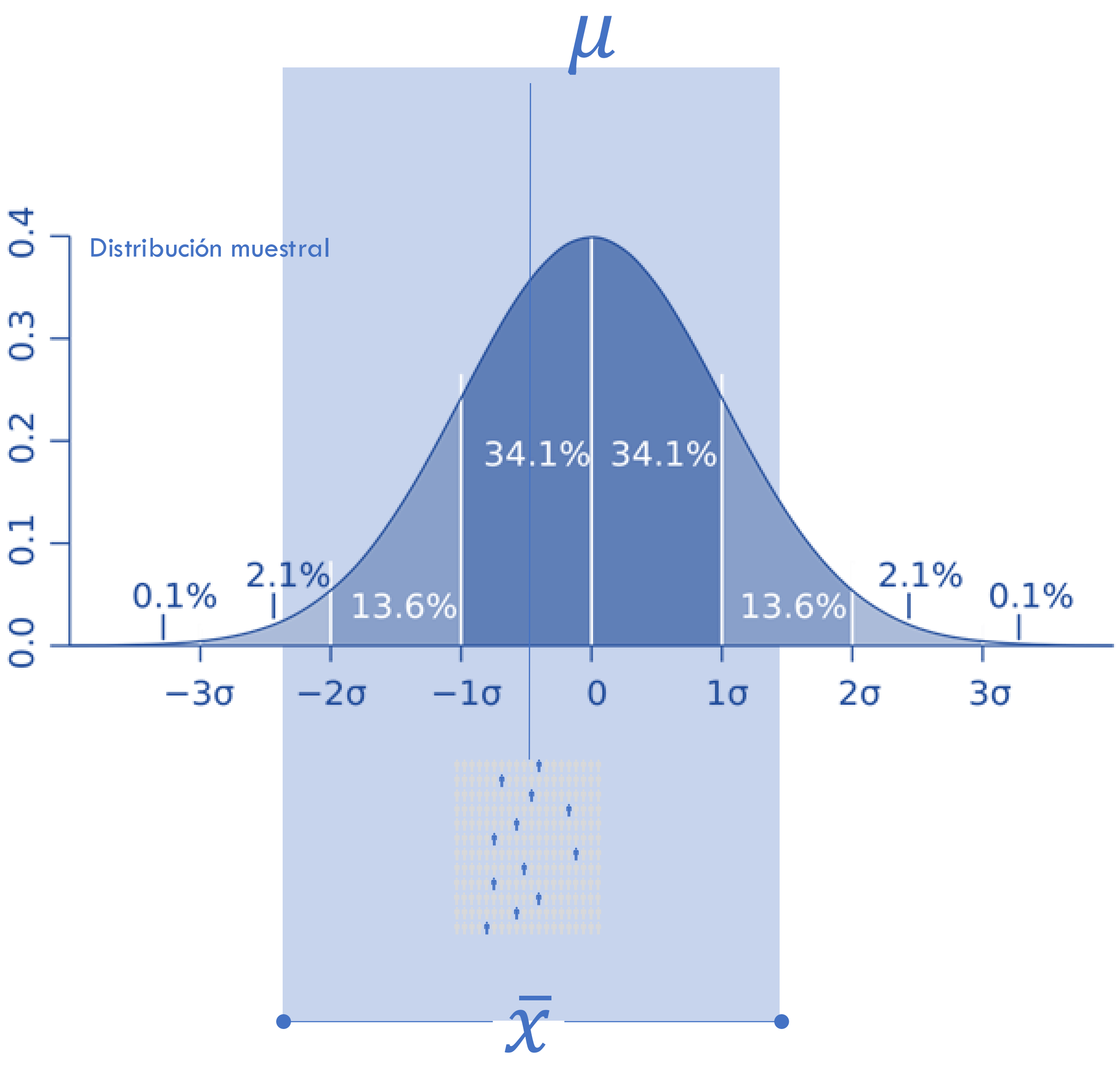
Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual “capture” al parámetro de la población, una mayoría de las veces.

Parámetro poblacional



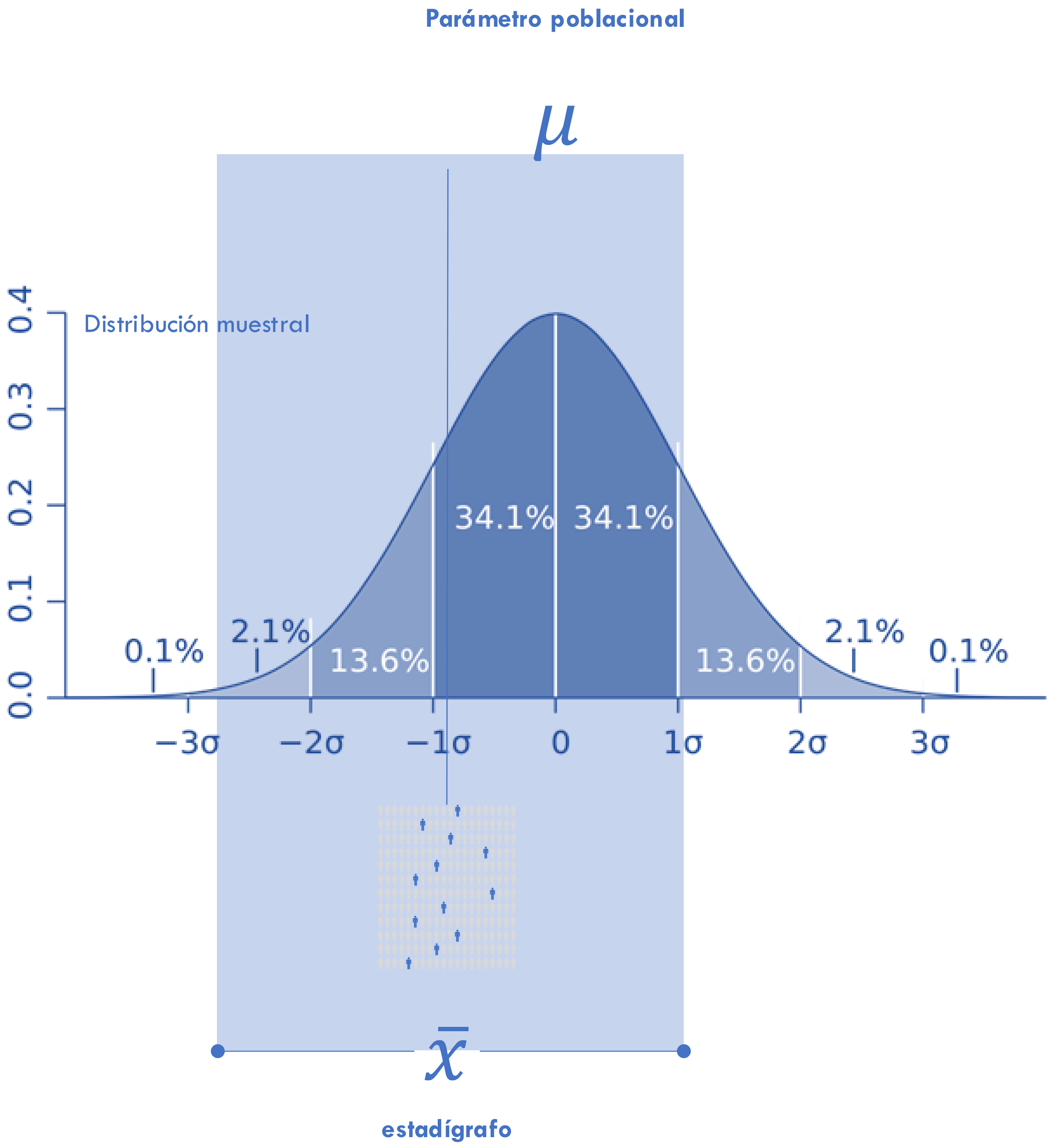
estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.



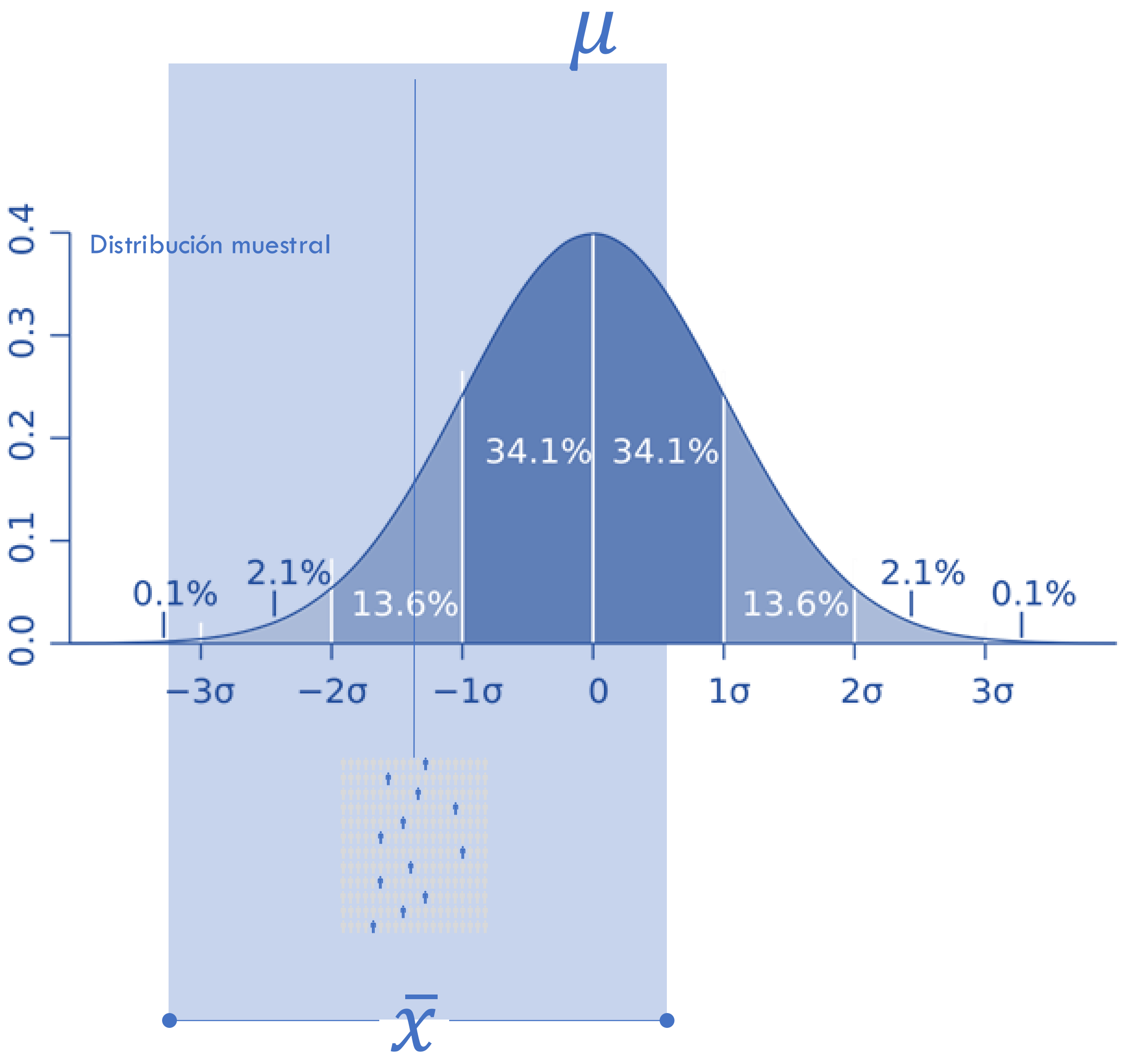
Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual “capture” al parámetro de la población, una mayoría de las veces.

Parámetro poblacional



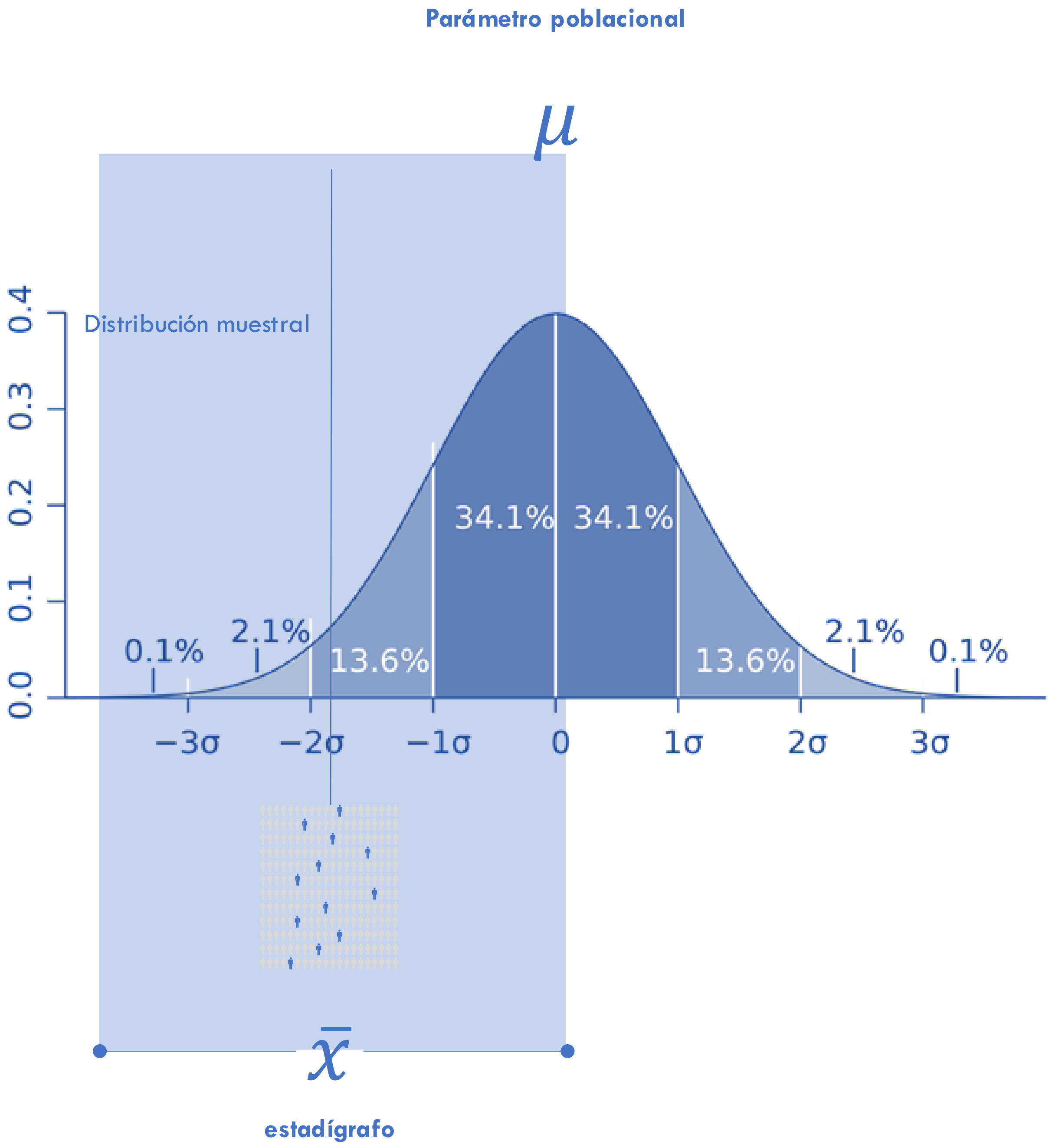
estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.



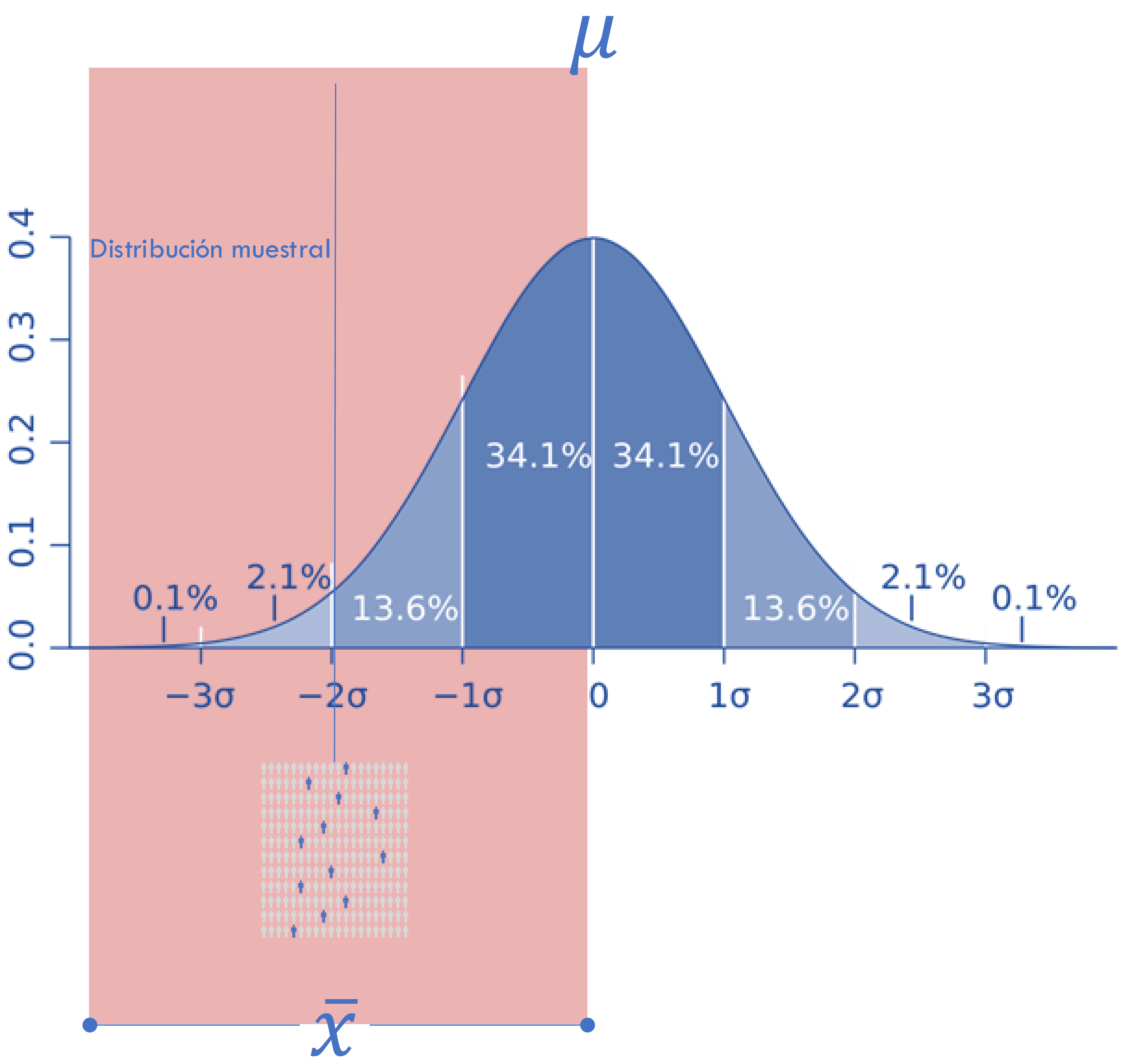
Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual “capture” al parámetro de la población, una mayoría de las veces.

Parámetro poblacional



estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

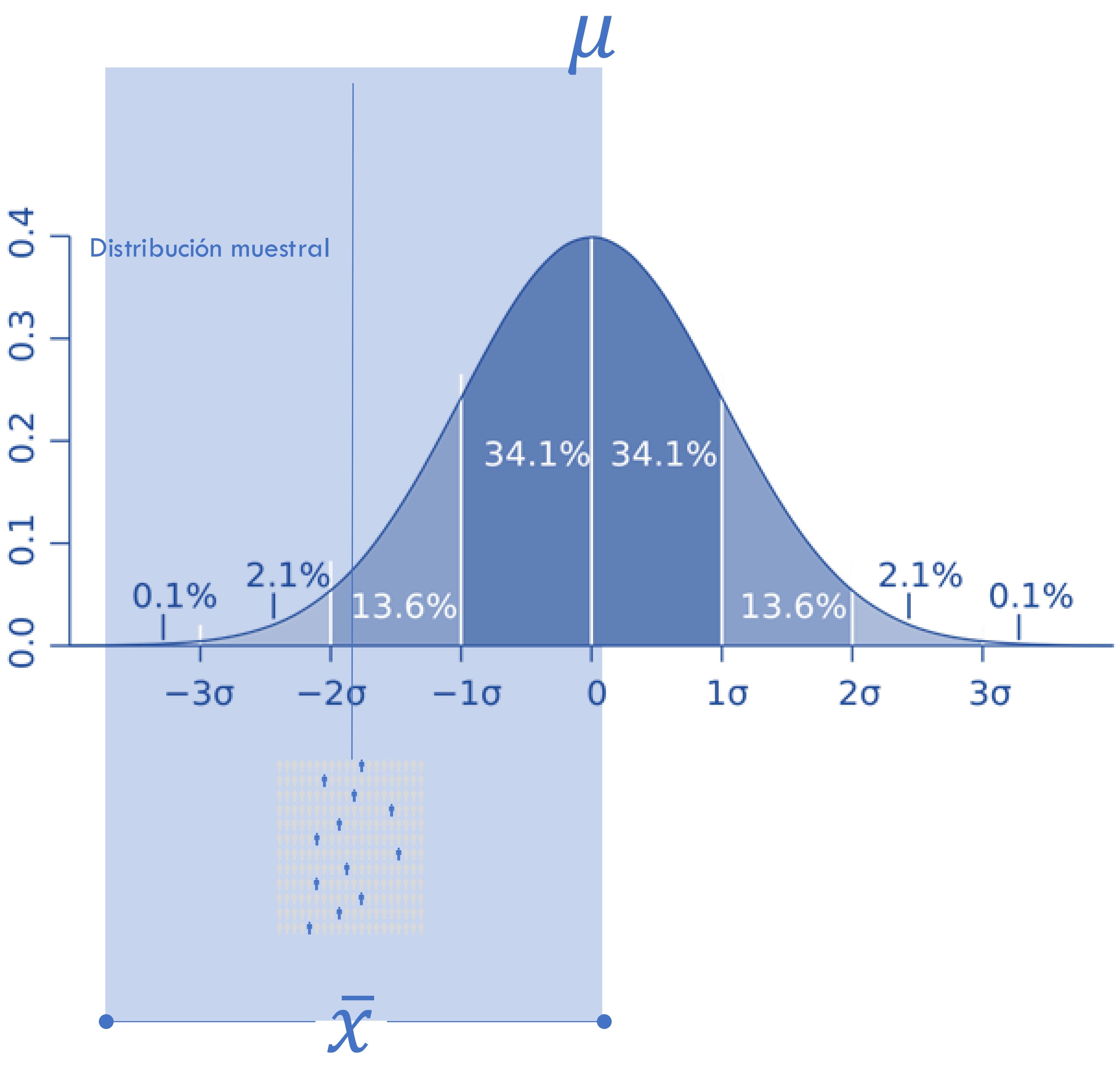
Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual “capture” al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.

Parámetro poblacional



estadígrafo

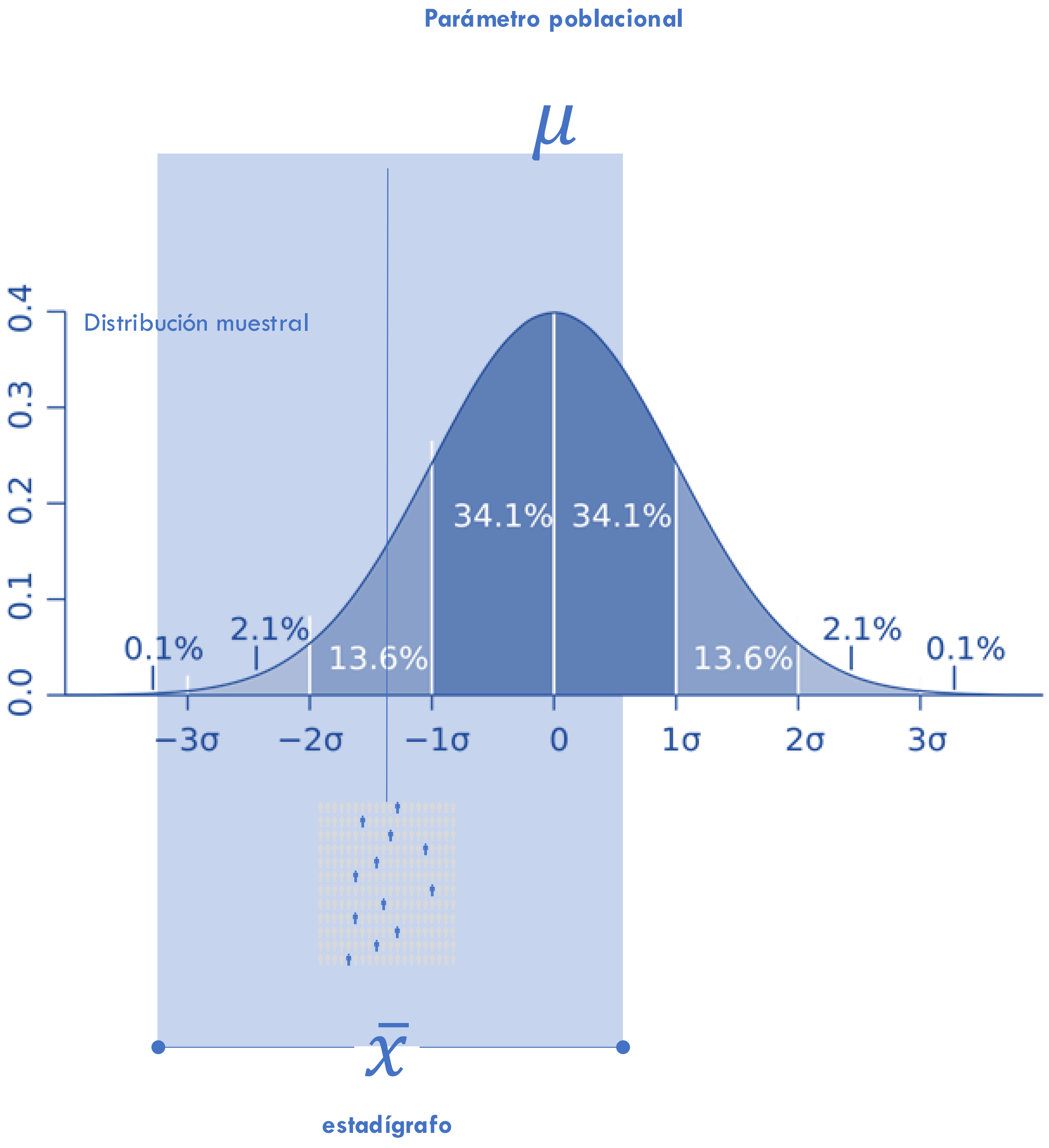
Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.



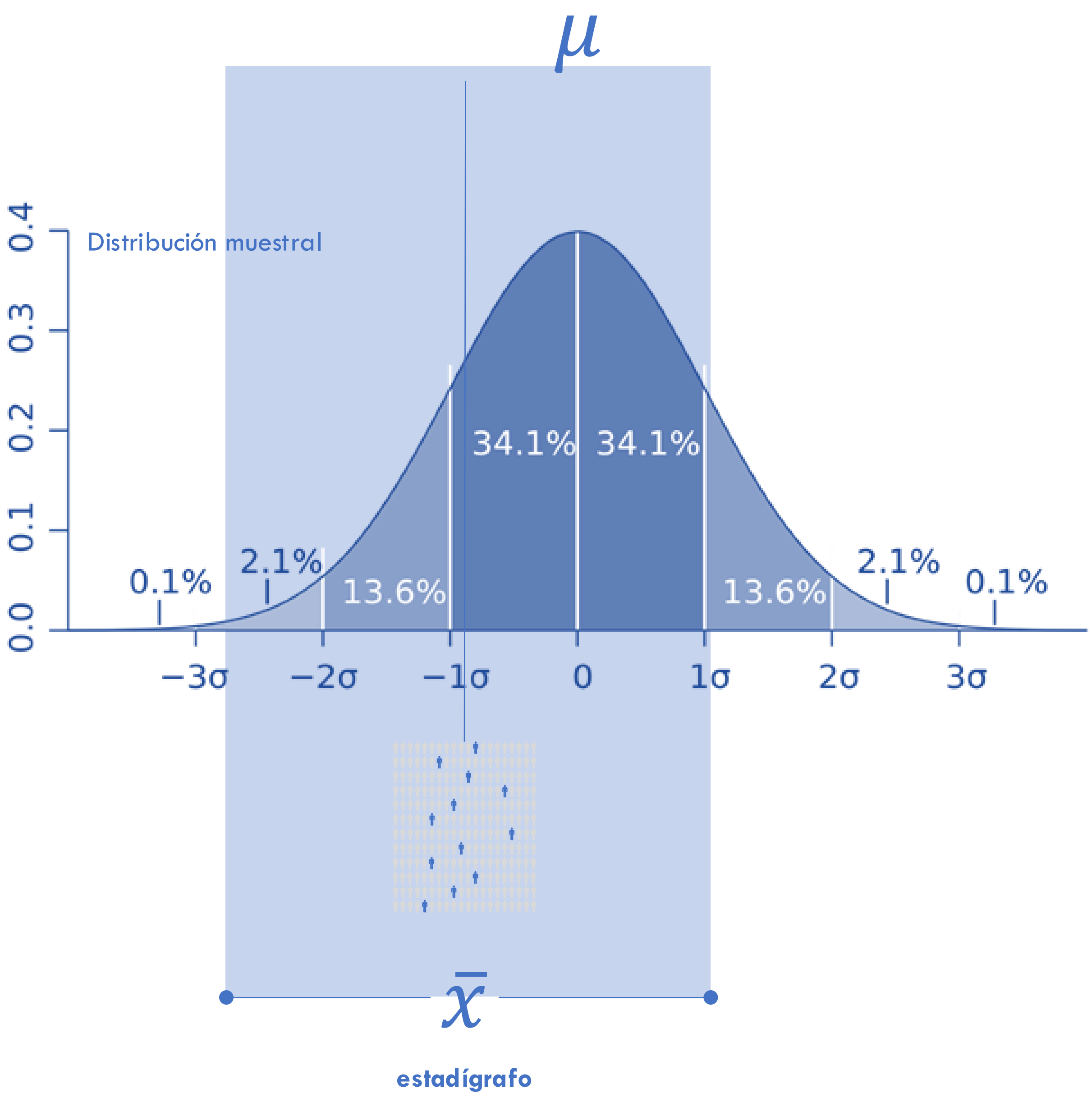
Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.



Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

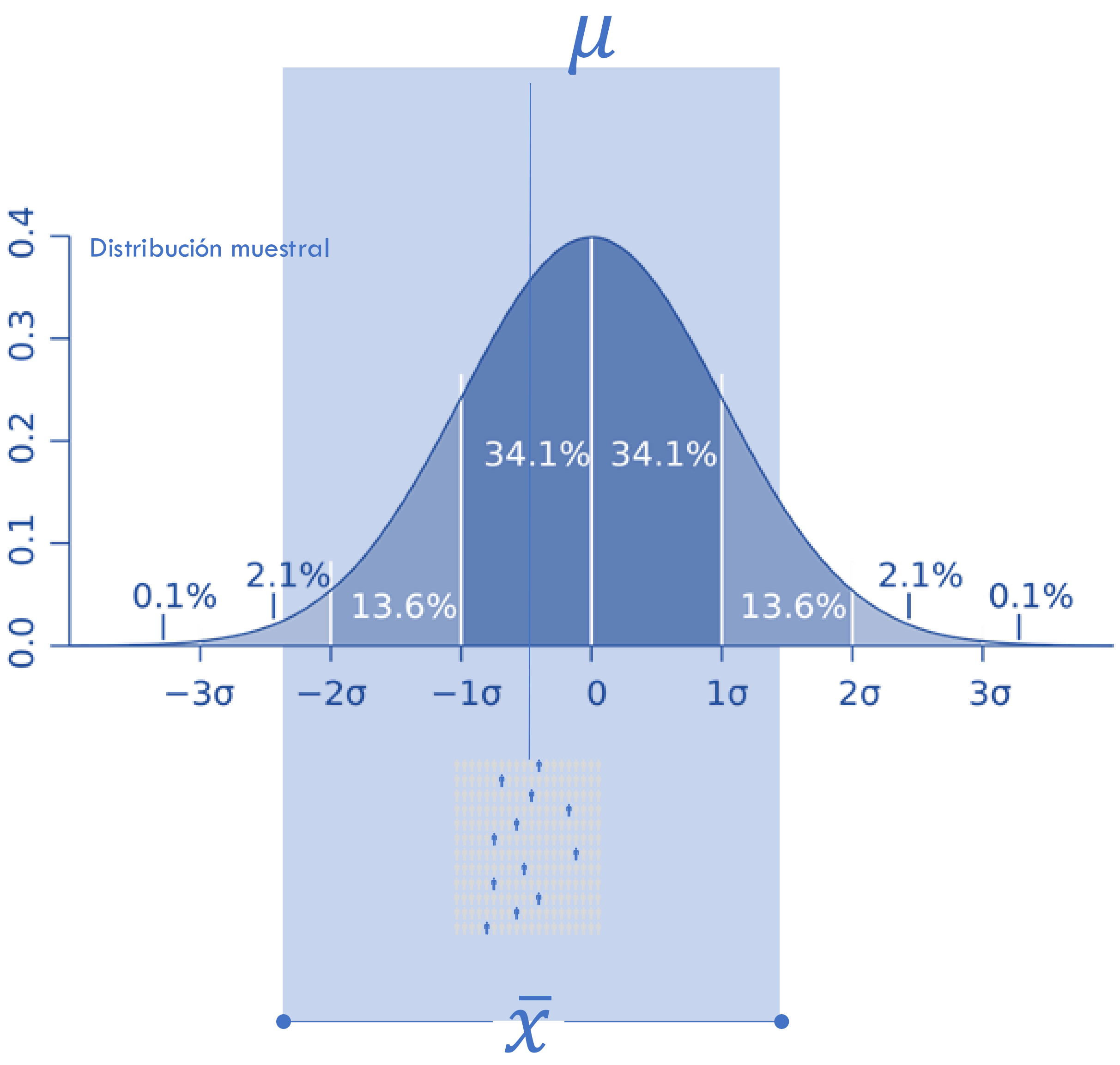
Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.

Parámetro poblacional



estadígrafo

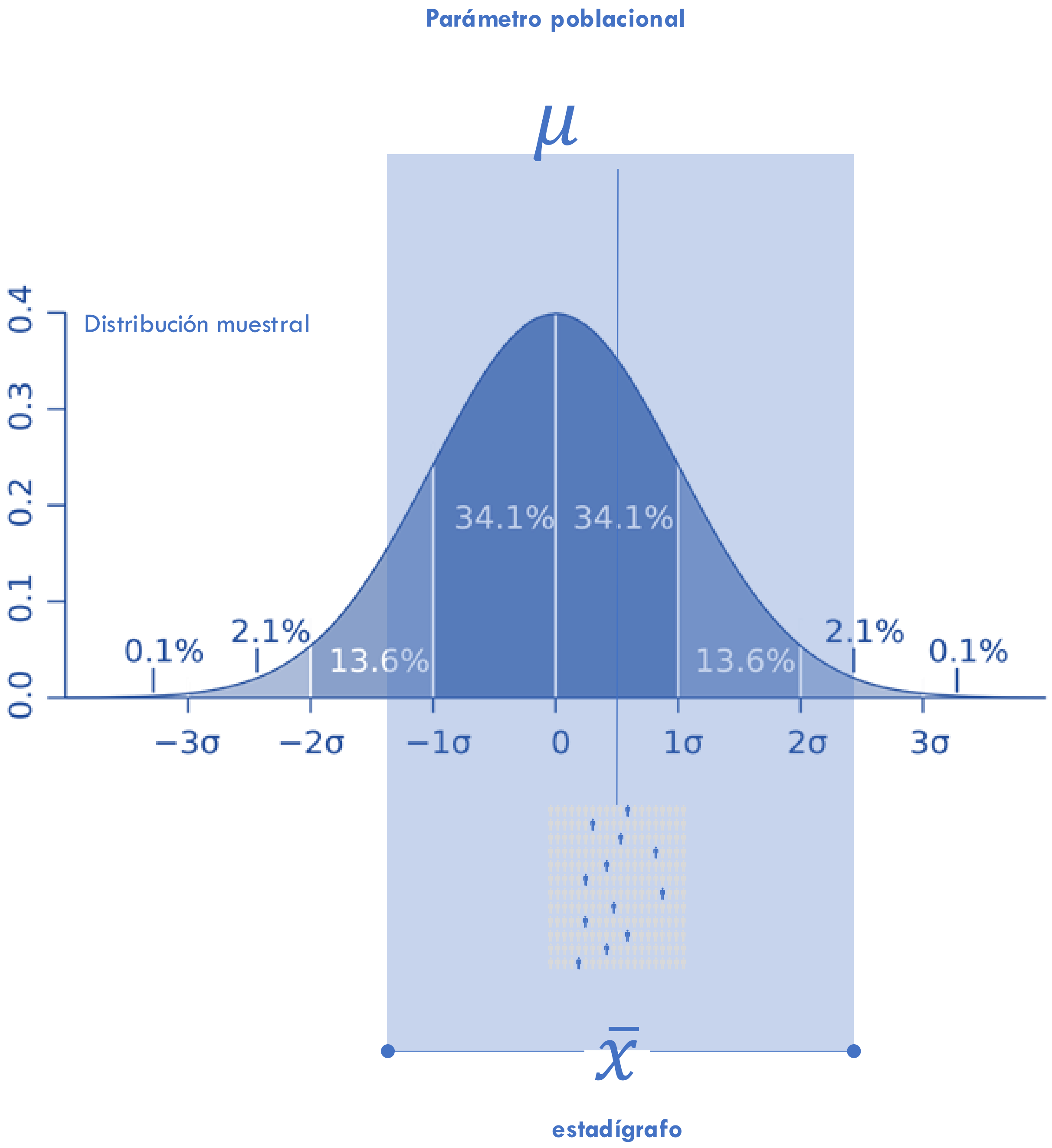
Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.



Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

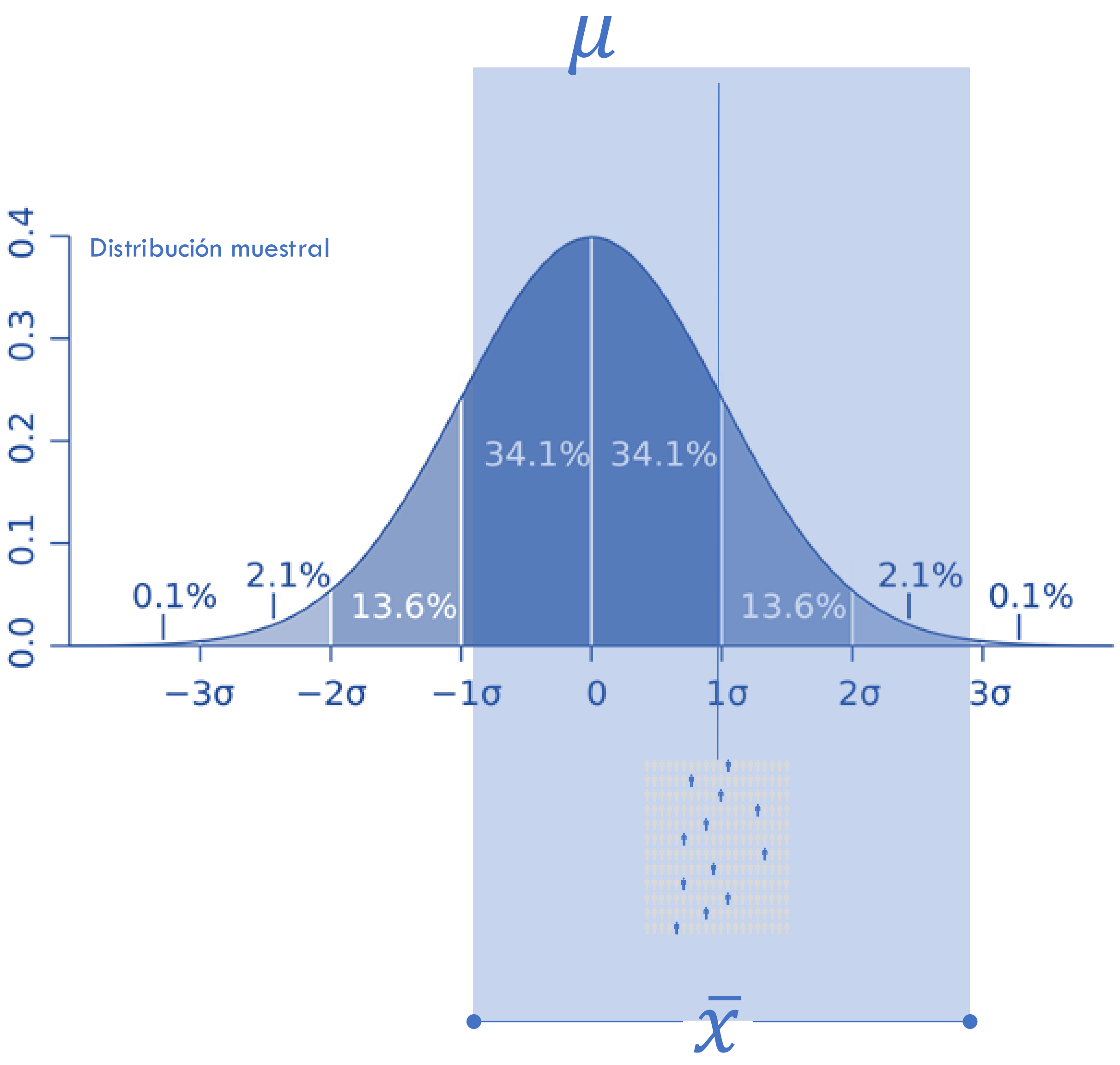
Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual “capture” al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.

Parámetro poblacional



estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

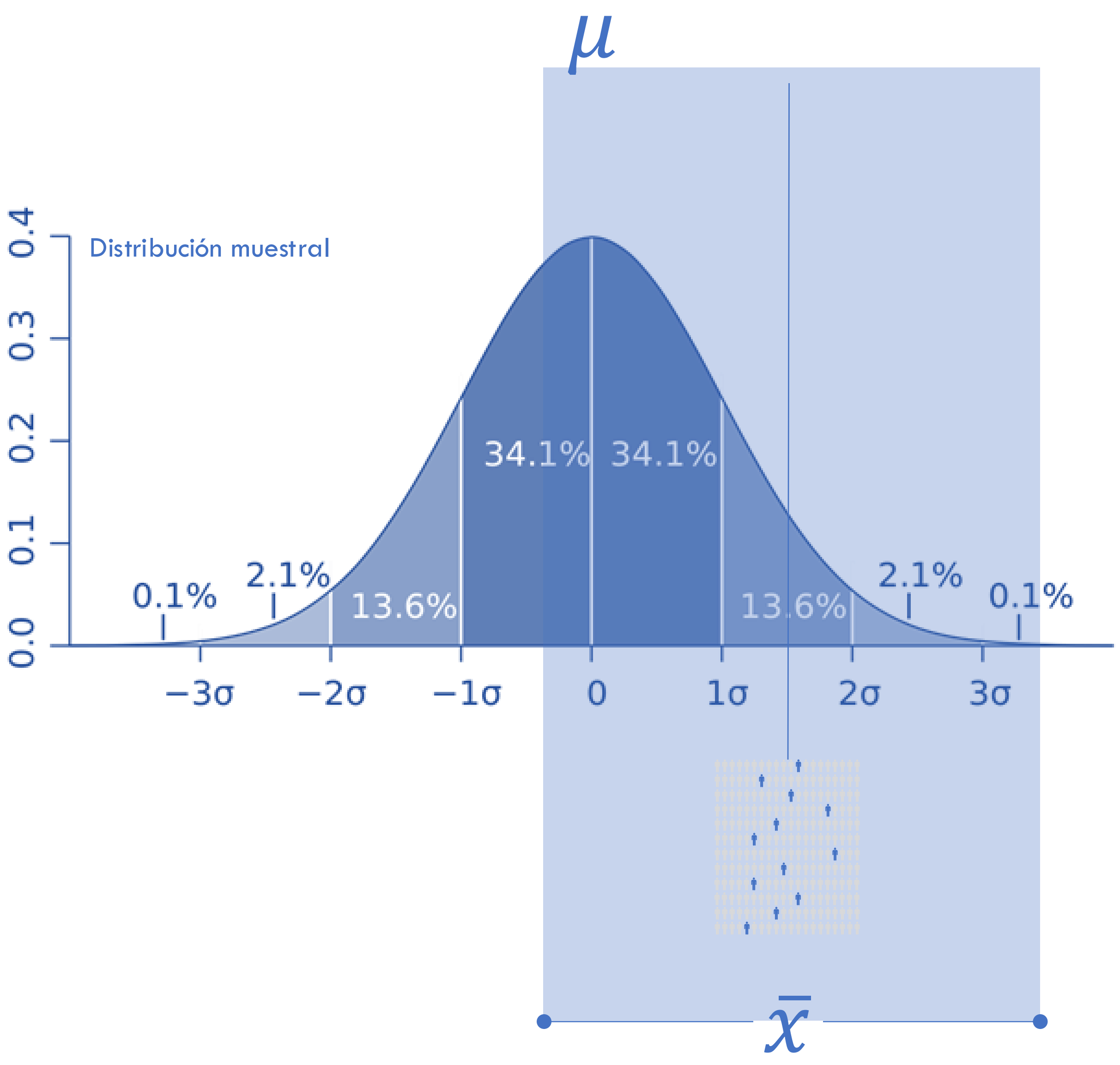
Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.

Parámetro poblacional



estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

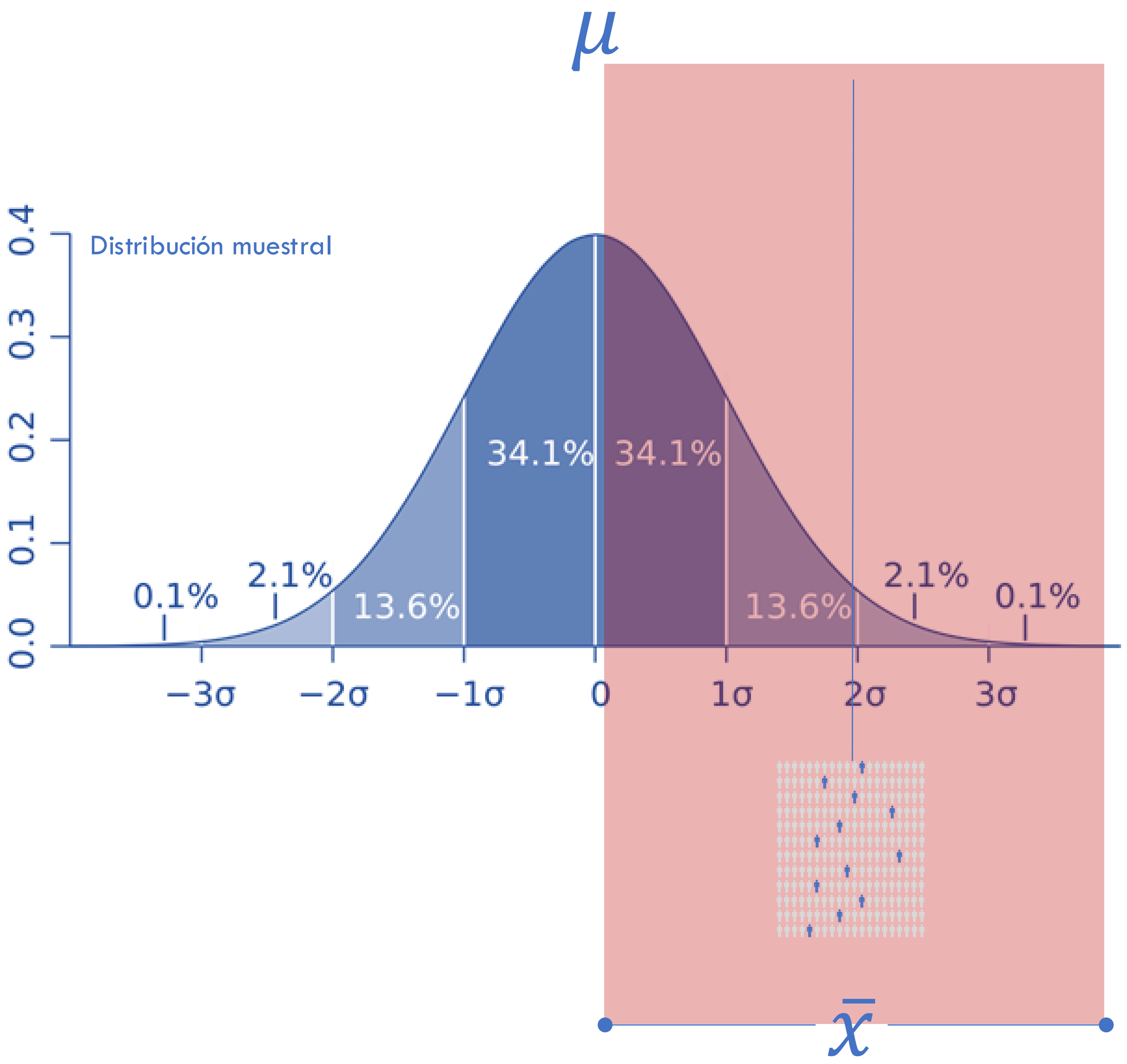
Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual “capture” al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.

Parámetro poblacional



estadígrafo

Tenemos incertidumbre respecto a cuál es la muestra aleatoria que tenemos de la distribución de muestras aleatorias posibles.

Tampoco sabemos, cuando la muestra aleatoria que tenemos coincide con el centro de la distribución.

Lo único que sabemos, es que la muestra aleatoria que tenemos es una de todas las muestras aleatorias posibles.

Lo que podemos hacer es establecer un rango de valores; un intervalo de confianza, el cual "capture" al parámetro de la población, una mayoría de las veces.

Si este rango es de 1.96 de Error estándar, entonces **esperamos que solo un 5% de las veces, una muestra aleatoria no capture el parámetro poblacional.**

Inferencia

Inferencia intervalar

Qué son los intervalos de confianza

Inferencia por medio de Intervalos de confianza

Una muestra aleatoria, puede contener el parámetro de interés, o puede no contenerlo. Dado que la distribución muestral es normal, entonces, podemos emplear a la distribución normal estandarizada, una **distribución probabilística,** para realizar inferencias.

Empleando la información de la distribución normal estandarizada, podemos crear intervalos, que un 95% de las veces contengan al parámetro poblacional.

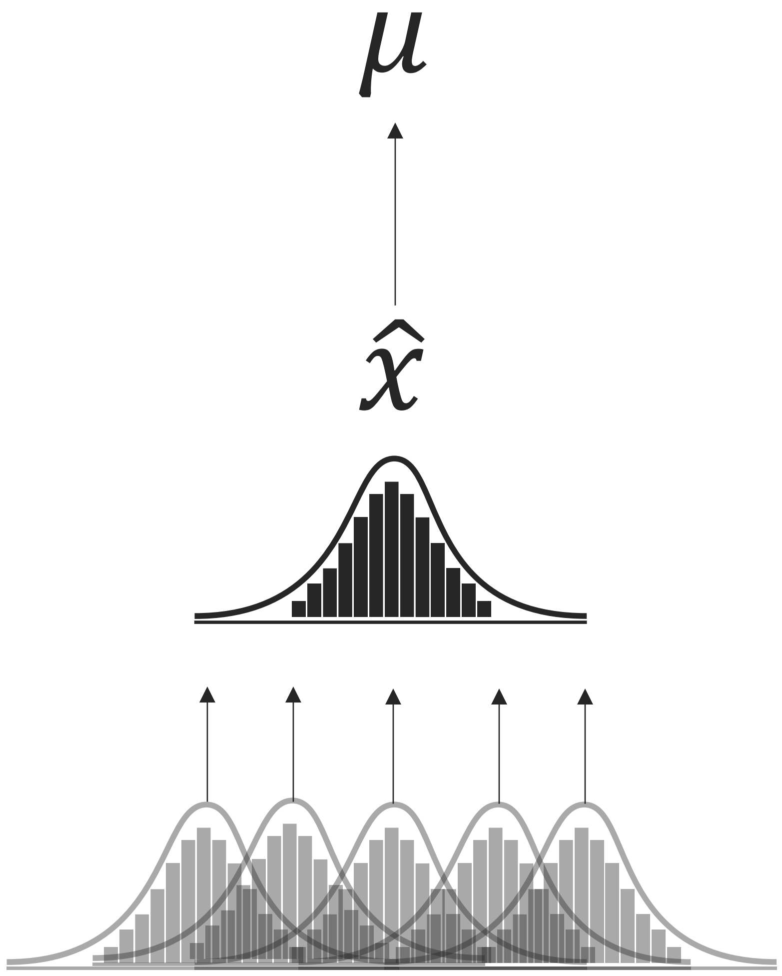
Medias

$$CI95\%\ [\bar{x}\ - 1.96\frac{\sigma}{\sqrt{n}}\ ;\ \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\ ]$$

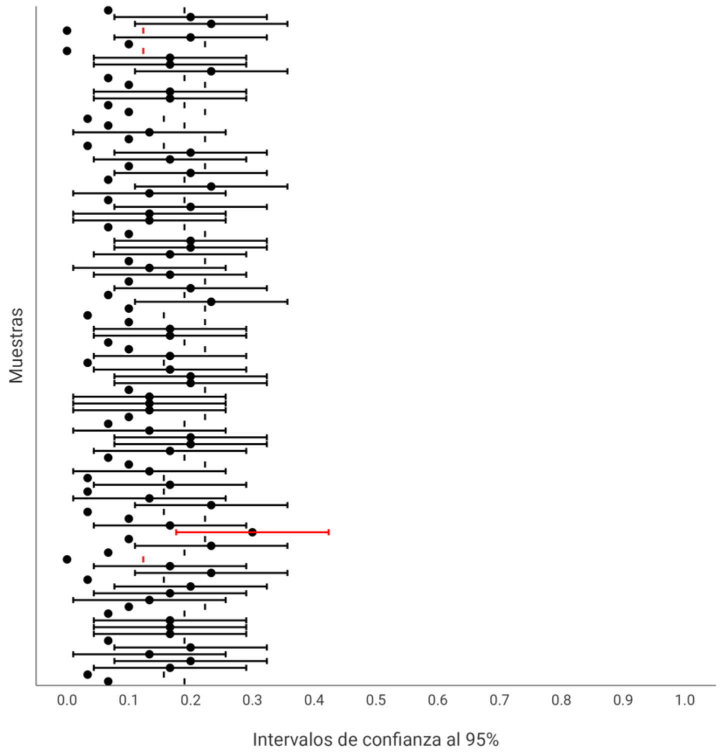
Proporciones

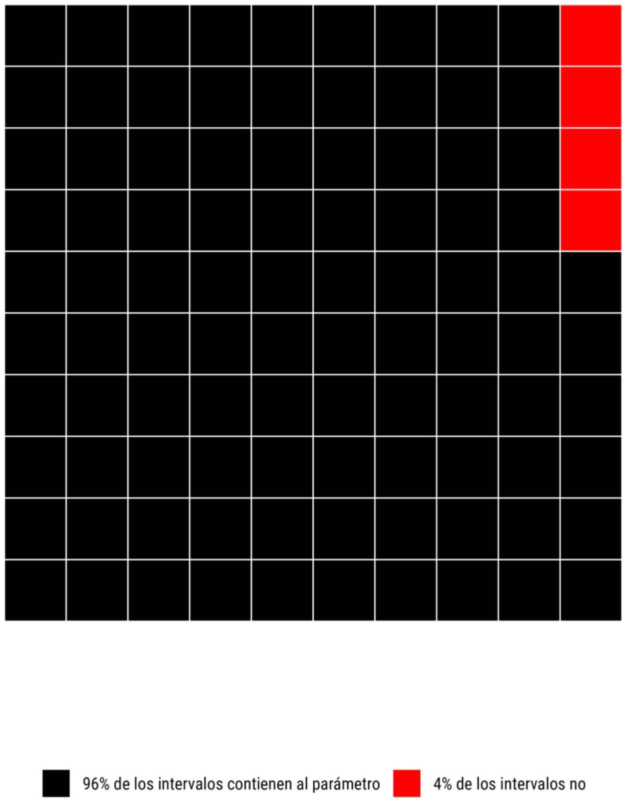
$$CI95\%\ [\bar{x}\ - 1.96\sqrt{\frac{\bar{p}(1-\bar{p})}{n}};\ \bar{x} + 1.96\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}]$$

Realizar inferencias poblacionales con este método, consiste en indicar que el parámetro de la población (e.g., medias, proporciones), podría encontrarse en un rango de valores.



Distribución Muestral de 30 observaciones



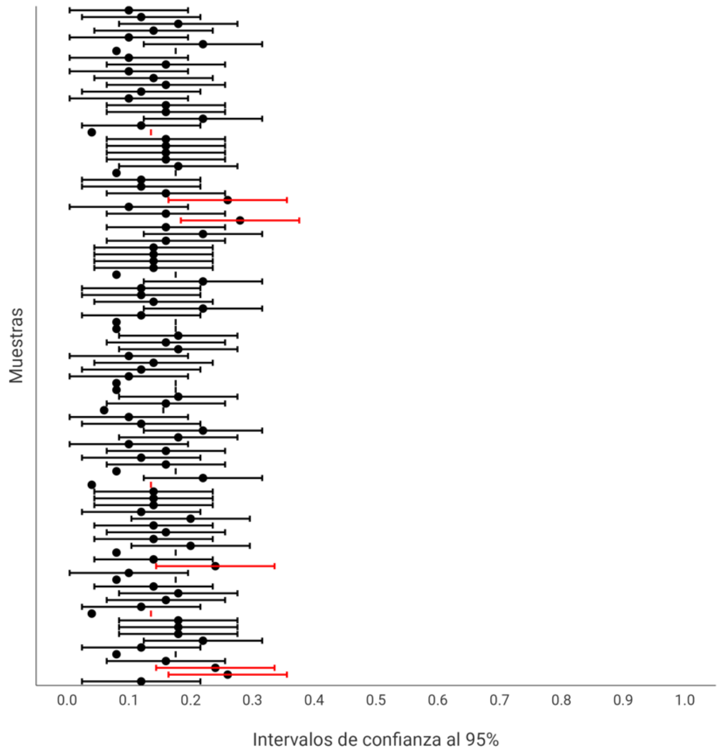


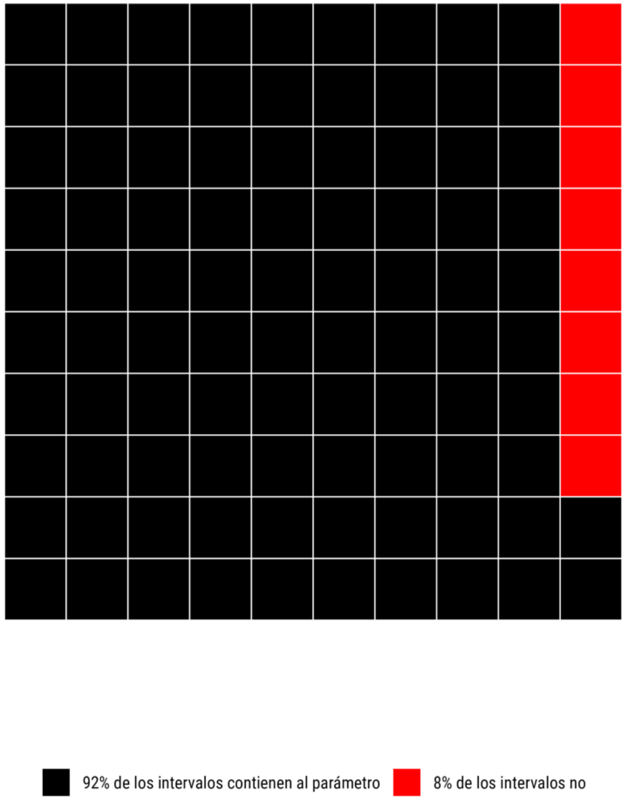
Generamos una distribución muestral de 30 observaciones, con 100 replicas. Calculamos el límite inferior y superior de un intervalo de 95% de confianza alrededor de la media de cada muestra.

Luego, clasificamos cada uno de estos intervalos según si contiene o no, al parámetro de la población (Pr(y = 1) = .14).

Un intervalo, contiene al parámetro de la población, si el límite inferior es menor al parámetro de la población, y si el límite superior es mayor al parámetro poblacional.

Distribución Muestral de 50 observaciones

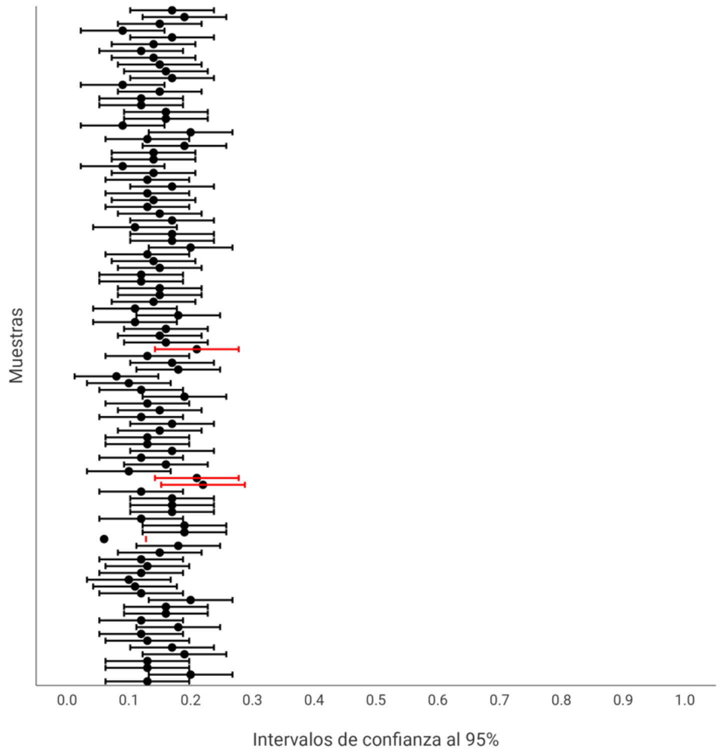


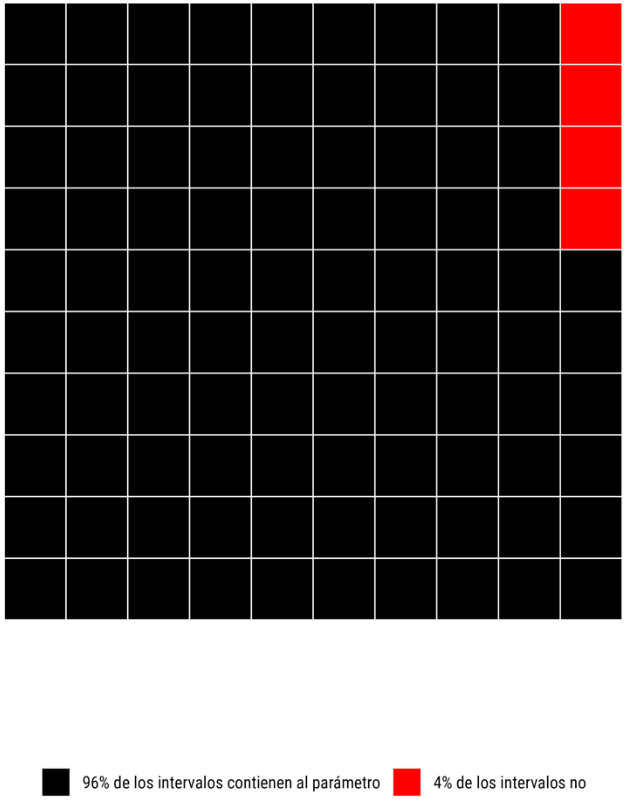


A medida que empleamos muestras de mayor tamaño, por ejemplo, de 50 observaciones, en vez de 30 observaciones, los intervalos de confianza son más estrechos. En otras palabras, el rango intervalar es más pequeño.

En estas simulaciones, intervalos pueden capturar al parámetro poblacional unas 92% a 96%. No obstante, si aumentáramos la cantidad de muestras de la distribución muestral, por ejemplo, a 10000 muestras, la expectativa es que el 95% de los intervalos generados capturaría al parámetro de la población.

Distribución Muestral de 100 observaciones



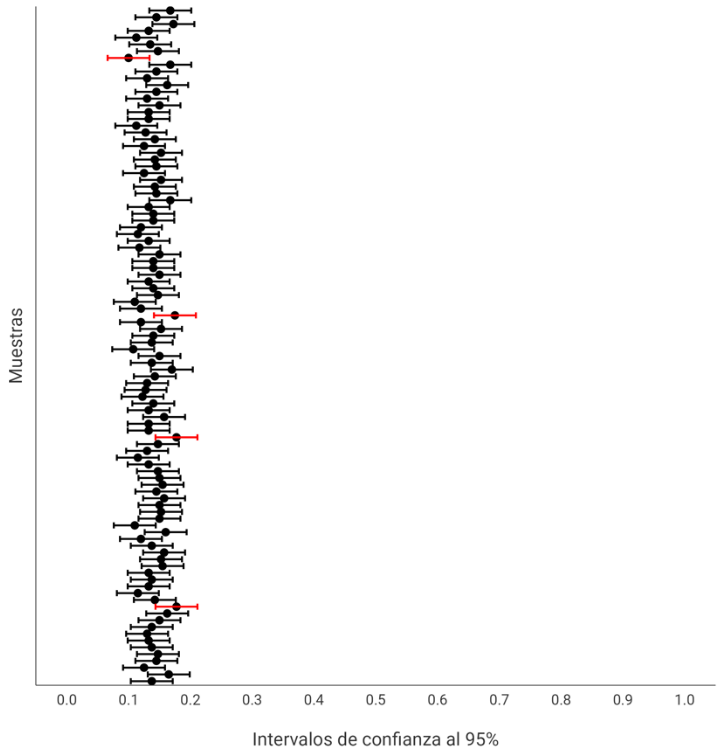


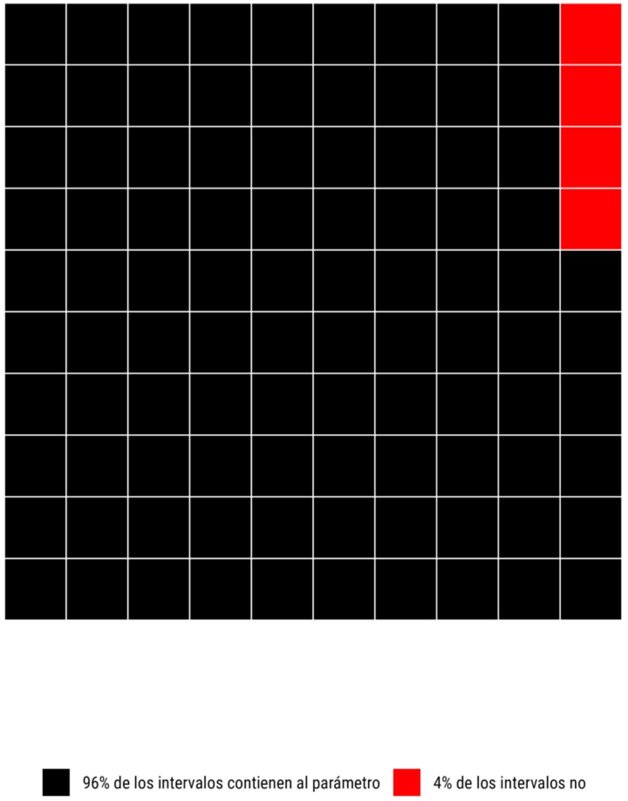
Nótese que, aunque la distribución muestral tuviera 30, 50, 100 o 500 casos, dado que la distribución muestral es normal, estos intervalos capturan al parámetro poblacional solo un 95% de las veces.

En otras palabras, el tamaño muestral no influye sobre la cantidad de veces en que un intervalo de 95% de confianza va a capturar al parámetro poblacional.

En el caso de las medias, a partir de muestras de 30 observaciones, la distribución muestral, toma una forma normal. De esta forma, a partir de este tamaño, podemos construir intervalos de esta forma.

Distribución Muestral de 400 observaciones



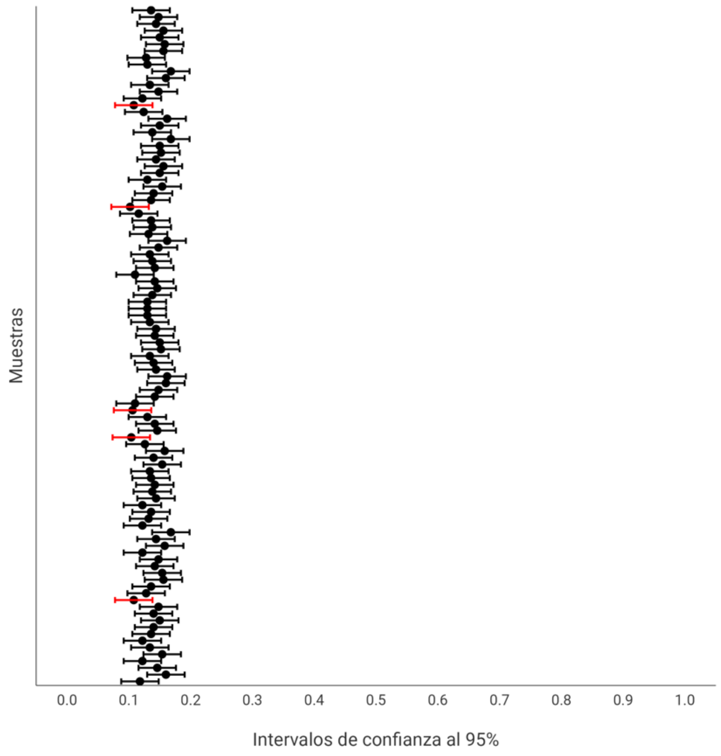


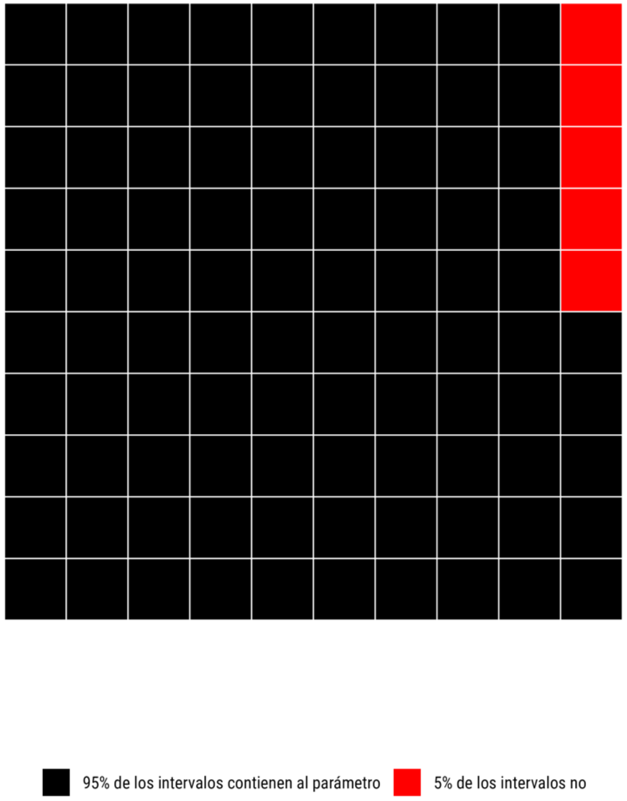
Para crear intervalos de confianza de 95% empleamos el puntaje z = 1.96.

Bajo -1.96 se encuentran un 2.5% de los valores, de esta distribución probabilística. Sobre 1.96 se encuentra otro 2.5% de todos los valores posibles.

Multiplicamos este valor, llamémosle valor z crítico, por la desviación estándar de la distribución muestral. Dicho de otra forma. Multiplicamos el valor crítico, por el error estándar de la media. Con esta operación, obtenemos el margen de error.

Distribución Muestral de 500 observaciones





El margen de error es luego, sumado y restado sobre el estimado de cada muestra aleatoria generada.

Con lo anterior, generamos dos valores: el límite inferior del intervalo, y el límite superior del intervalo.

En términos generales, estamos empleando la siguiente formula:

$$CI95\%$$

$$[\bar{x} \ - 1.96\ ES\ ;\ \bar{x} + 1.96\ ES\ ]$$

Dónde, ES es el error estándar (i.e., la desviación estándar de la distribución muestral).

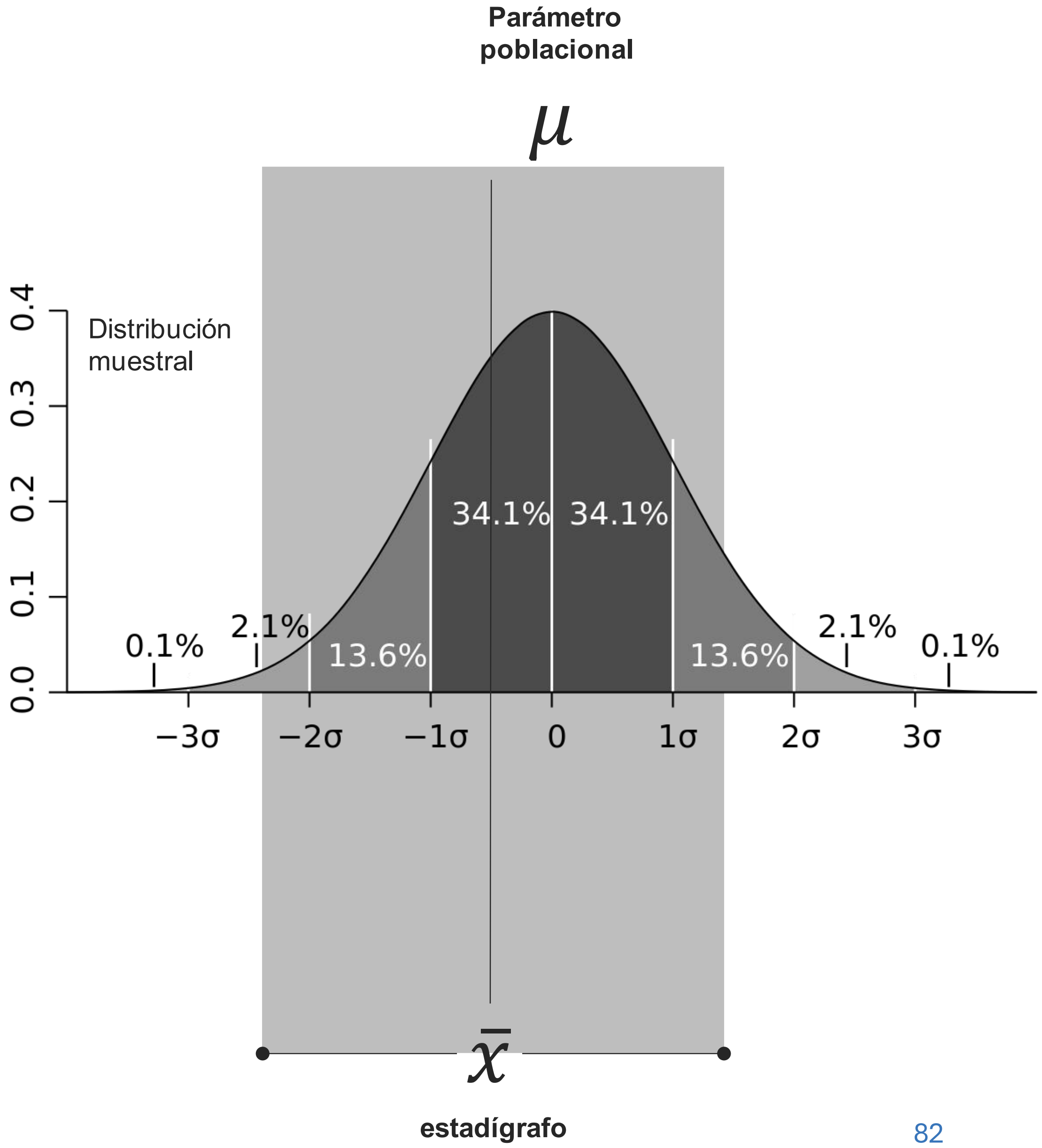
Intervalo de confianza

Un intervalo es un rango de valores que se extienden entre dos límites, uno superior y otro inferior.

Hablamos de **intervalo de confianza**, al intervalo que se construye alrededor de los estadígrafos, para garantizar que **el parámetro poblacional fuera contenido en una proporción de las veces**, es decir con una cierta probabilidad, si generásemos una distribución muestral.

El intervalo de confianza más tradicional es el de 95%. También es posible construir intervalos de 99%. Intervalos de 99% son a veces empleados para muestras grandes (e.g., 1000 observaciones). Mientras que los intervalos de 95% se emplean con muestras de menos de 1000 casos.

En esta clase hemos visto como crear intervalos de confianza para medias. También es posible crear estos intervalos para otros estadígrafos.



CI para	Estadísgrafo	Margen de error	Cuando se usa tradicionalmente
Media poblacional (μ)	$\bar{x}$	$\pm z \frac{\sigma}{\sqrt{n}}$	x es normal n ≥ 30 σ es conocido
Media poblacional (μ)	$\bar{x}$	$\pm t_{n-1} \frac{s}{\sqrt{n}}$	n < 30 σ no es conocido
Proporcional poblacional (p)	$\bar{p}$	$\pm z \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$	Cada proporción p, o 1-p, es mayor o igual a 10 observaciones.
Diferencia de dos medias poblacionales ($\mu_1 - \mu_2$)	$\bar{x}_1 - \bar{x}_2$	$\pm z \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	Ambas distribuciones son normales; de tamaño 30 o mayor; y/o con varianzas conocidas.
Diferencia de dos medias poblacionales ($\mu_1 - \mu_2$)	$\bar{x}_1 - \bar{x}_2$	$\pm t_{n_1+n_2-2} \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$	Muestras de menos de 30 casos; y escenarios de varianzas no conocidas
Diferencia de dos proporciones poblacionales ($p_1 - p_2$)	$\bar{p}_1 - \bar{p}_2$	$\pm z \sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}$	p y 1-p, son proporciones generadas por muestras mayores a 10 observaciones, en ambos grupos.

Ref: https://www.dummies.com/article/academics-the-arts/math/statistics/surveying-statistical-con¦dence-intervals-188336/

Inferencia

Interpretación de los CI95%

Cómo interpretar los intervalos de confianza

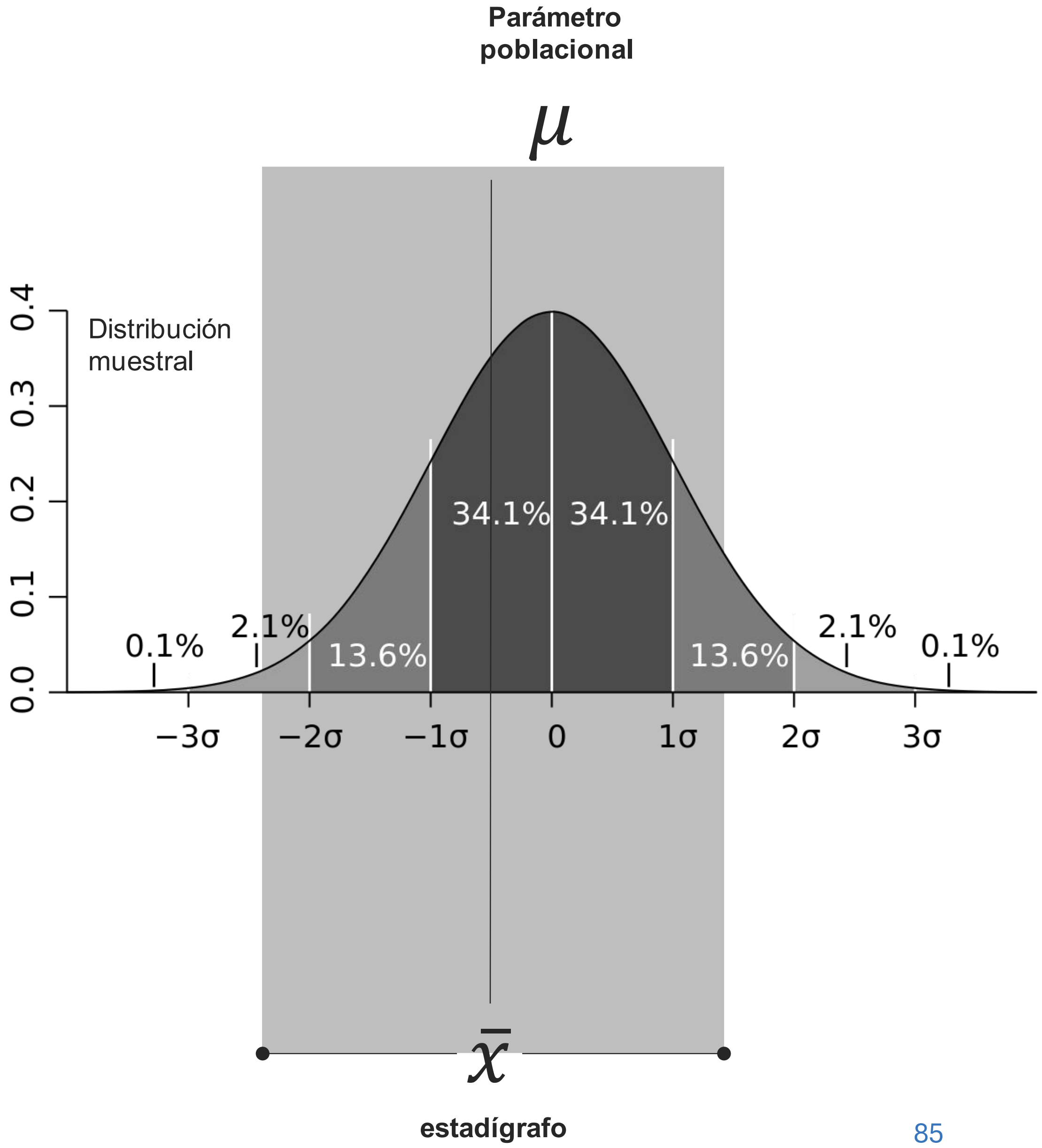
Intervalo de confianza y su interpretación errónea

Si un intervalo de confianza de 95% ha sido creado para estimar el valor de un parámetro de la población, la probabilidad que el parámetro de interés ***se encuentre entre los límites del intervalo de confianza, es igual a 95 de 100****. [...] Esto es una mala comprensión de lo que implica un intervalo de confianza.*

— Huck, 2009, p147

Como vimos en las simulaciones, y en la selección de diferentes muestras aleatorias sobre datos reales (ver anexos), **estos intervalos "pueden" contener o no**, al parámetro de la población.

El "95% de las veces" es una expectativa respecto a los intervalos de confianza sobre las medias en la distribución muestral.

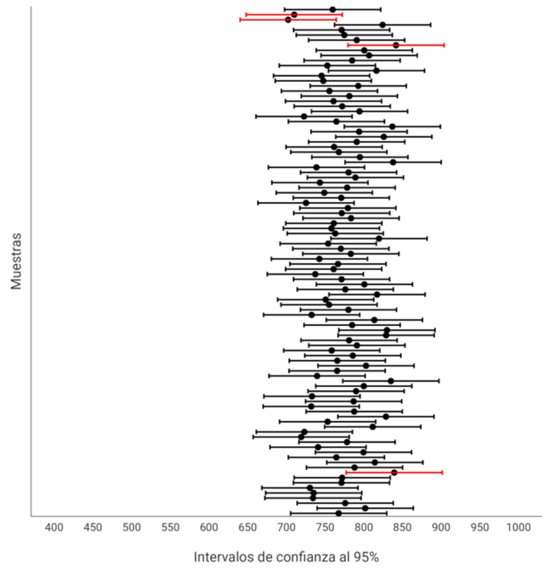


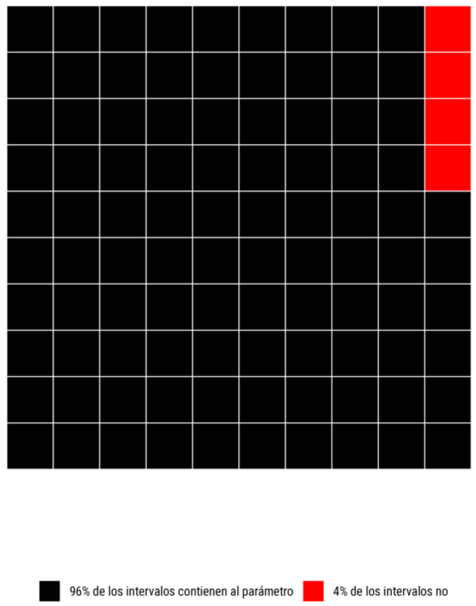
Intervalo de confianza y su interpretación

Dado el teorema del límite central, si las muestras aleatorias fueran recogida infinitas veces se posee la expectativa de que **la distribución muestral resultante,** puede contener al parámetro de la población 95% de las veces, una vez construido los intervalos de confianza sobre las medias de cada muestra.

La probabilidad de contener 95% de las veces al parámetro, es una afirmación respecto a los intervalos de la distribución muestral, no respecto a una media de una sola muestra.

Uno de los aspectos críticos a distinguir en la interpretación de los CI95%, es el objeto de la afirmación anterior. Sería un error categorial (Ryle, 1949), asumir que el 95% de probabilidad es respecto a la media de una sola muestra. La interpretación correcta de un intervalo es que tenemos un solo intervalo de 95% confianza, de una distribución muestral.





Cuando se produce una distribución muestral de n≥30 tamaño, podemos construir intervalos de 95% de confianza alrededor de todas las medias de cada muestra. De todos los intervalos, 95% de estos intervalos van a capturar al parámetro poblacional. **Cuando estimamos resultados con una sola muestra aleatoria, tenemos solo uno de estos intervalos posibles.**

Intervalo de confianza y su forma de presentación

EXCERPTS 6.4–6.5 • *Confidence Intervals Around a Mean and a Percentage*

ABC scores were obtained from 46 children, the mean score was 51 and 95% CI was 46–56 (SD 18).

Source: Fernell, E., Hedvall, A., Norrelgen, F., Eriksson, M., Höglund-Carlsson, L., Barnevik-Olsson, M., et al. (2010). Developmental profiles in preschool children with autism spectrum disorders referred for intervention. *Research in Developmental Disabilities*, *31*(3), 790–799.

Of the 18–30 year olds, 49.8% (95% confidence interval: 48.5%–51.2%) were men and 50.2% (48.8%–51.5%) were women.

Source: Cavazos-Rehg, P. A., Spitznagel, E. L., Krauss, M. J., Schootman, M., Bucholz, K. K., Cottler, L. B., et al. (2010). Understanding adolescent parenthood from a multisystemic perspective. *Journal of Adolescent Health*, *46*(6), 525–531.

Los puntajes ABC fueron obtenidos de una muestra de 46 niños, con un promedio de 51 y un intervalo de confianza al 95% 46-56 (DS = 18).

De los jóvenes de 18 a 30 años, 49.8% (intervalo de confianza al 95%: 48.5%-51.2%) son hombres, y 50.2% (48.8%-51.5%) son mujeres.

Intervalo de confianza y su forma de presentación

In Table 3, we present the school level estimates. The predictor with the highest size is again authoritarianism, with a standardised effect in the constrained model of 0.84 ($SE = 0.03$, CI 95% [0.78, 0.89], $p < .001$). For Guatemala, this effect is smaller than the rest of the countries, reaching a standardised effect of 0.59 ($SE = 0.07$, CI 95% [0.46, 0.72], $p < .001$). Thus, schools with higher levels of authoritarianism present higher levels of tolerance of corruption in comparison to other schools. In addition, schools with higher civic knowledge present lower levels of tolerance of corruption, reaching a standardised effect of -0.19 ($SE = 0.04$, CI 95% [−0.26, −0.11], $p < .001$) in the constrained model. Inspecting the country specific results, we found that Guatemala presents a larger coefficient than the rest of the compared countries, where its standardised effect is of −0.41 ($SE = 0.09$, CI 95% [−0.59, −0.23], $p < .001$), whereas Paraguay and the Dominican Republic present non-significant effects on this estimate. However, all country estimates overlap with the confidence interval of the

Carrasco, D., Banerjee, R., Treviño, E., & Villalobos, C. (2020). Civic knowledge and open classroom discussion: explaining tolerance of corruption among 8th-grade students in Latin America. *Educational Psychology*, *40*(2), 186–206. https://doi.org/10.1080/01443410.2019.1699907

También es posible construir intervalos de confianza no solo sobre medias, o sobre porcentajes. Los podemos crear alrededor de coeficientes de modelos. En este caso, se reportan los intervalos de confianza alrededor de coeficientes estandarizados de un modelo mixto, o de regresión multinivel. En este caso, se reporta el estimado primero, al interior del paréntesis, se reporta al error estándar, y al intervalo de confianza del estimado.

Inferencia

Marcos de inferencia

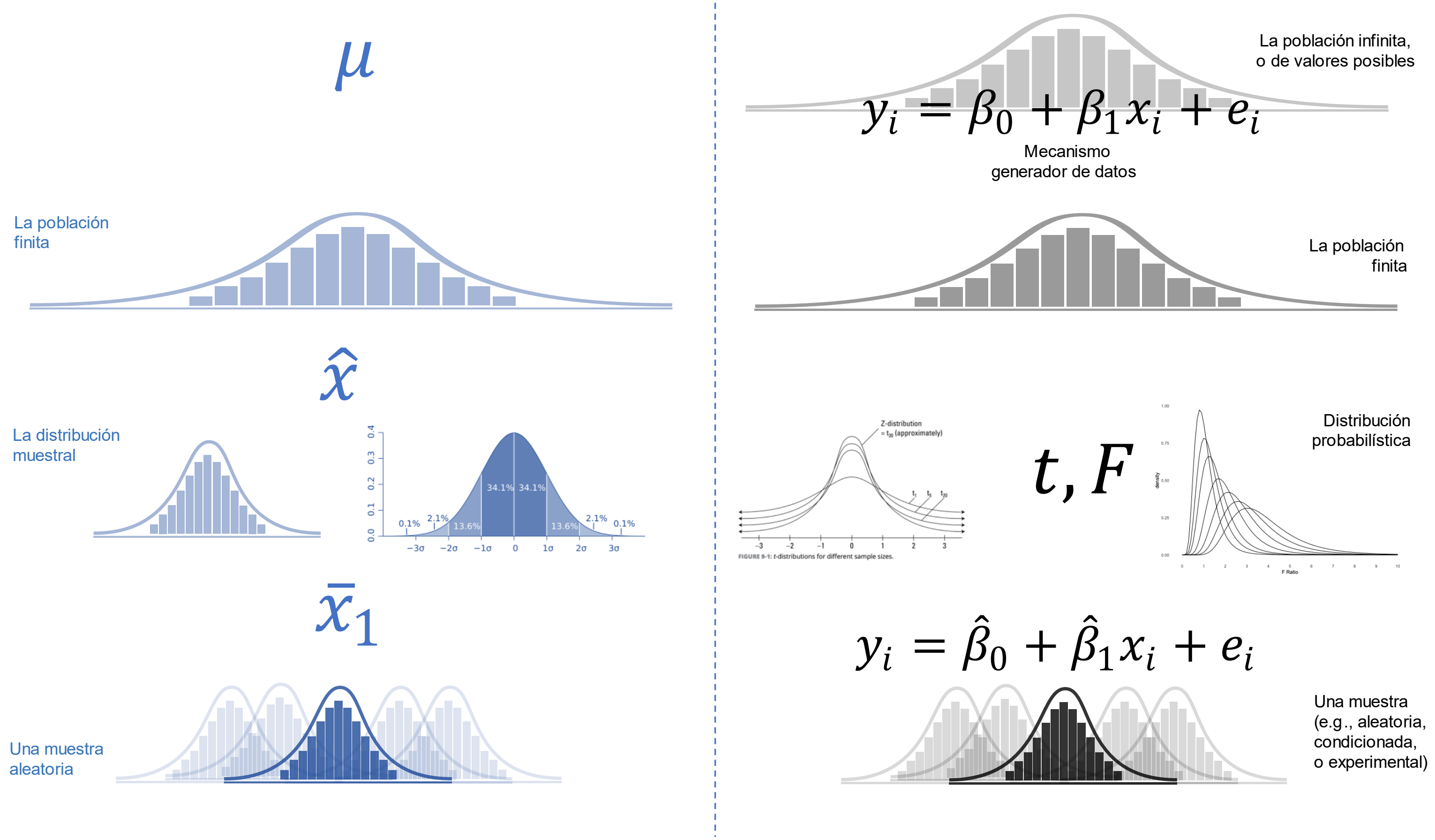
Inferencia basada en modelo e inferencia basada en diseño

Tipos de inferencia (Sterba, 2009)

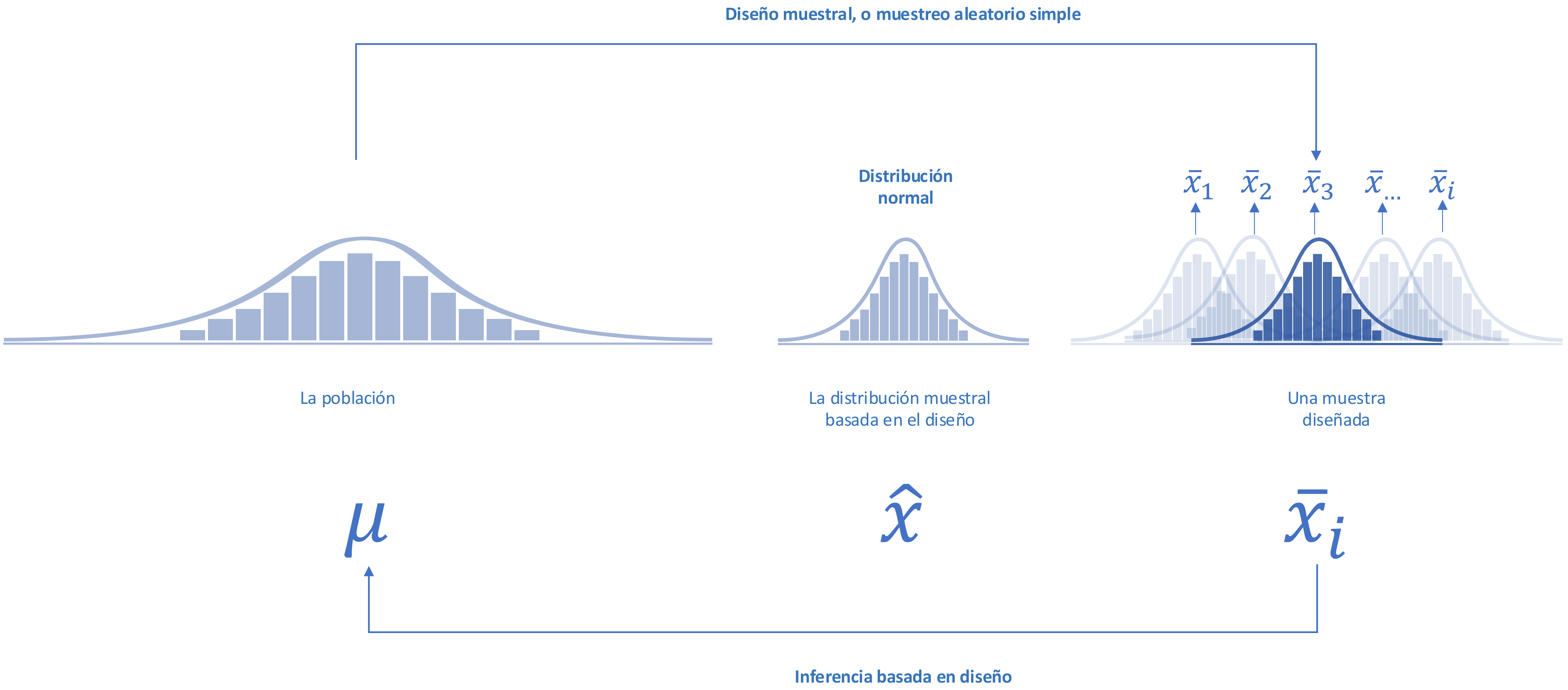
Sterba (2009) describe dos marcos conceptuales con los que se realiza inferencias empleando cifras.

Uno, es la **inferencia basada en diseño**. Este marco, es el marco fundamental de la creación de encuestas, y muestras probabilísticas. Los estudios de gran escala emplean este marco para el diseño de sus estudios. En este marco el objeto de la inferencia es la población finita.

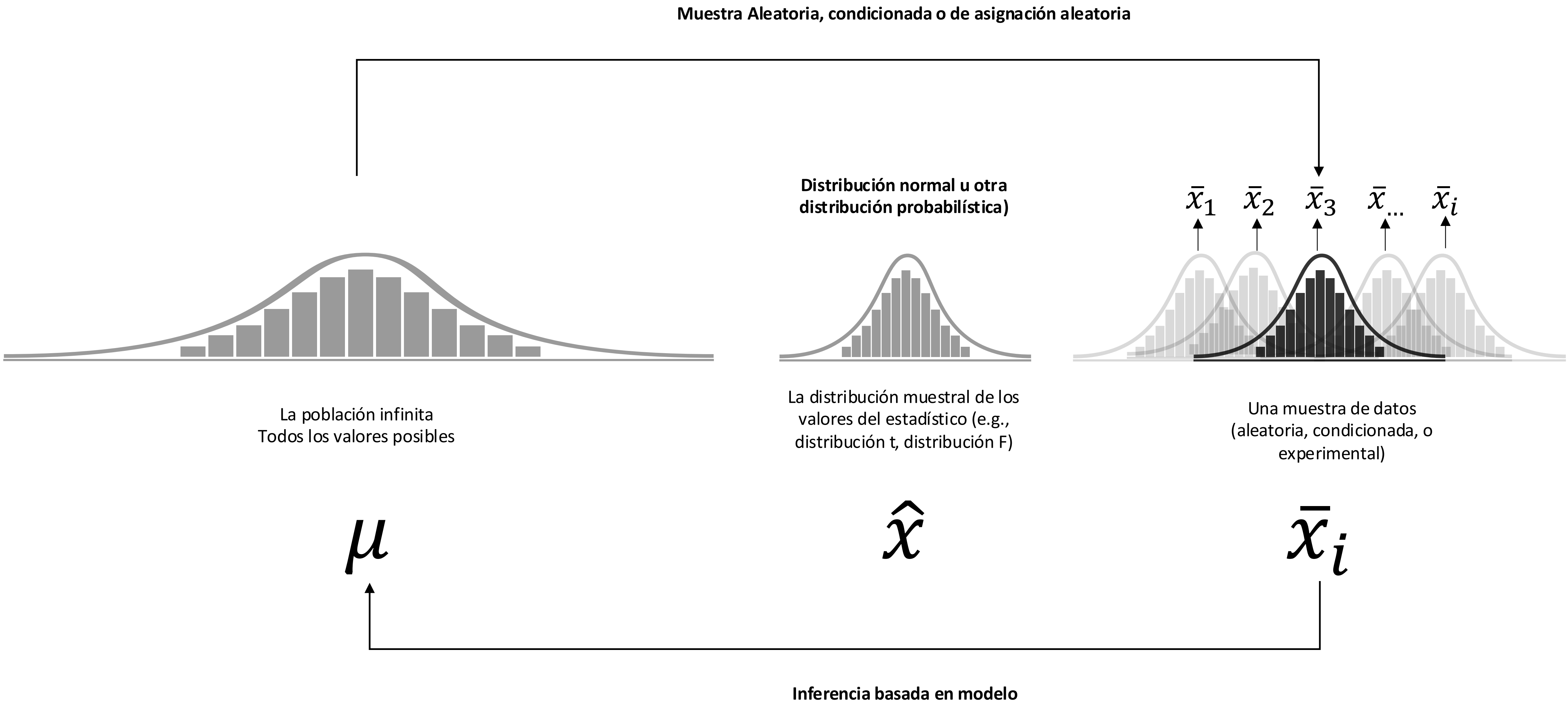
El segundo marco, es la **inferencia basada en modelo**. Este segundo marco es comúnmente empleado para sacar conclusiones empleando cifras, que provienen de la aplicación de modelos estadísticos. En este segundo marco, el objeto de la inferencia es la población infinita.



Inferencia basada en diseño



Inferencia basada en modelo



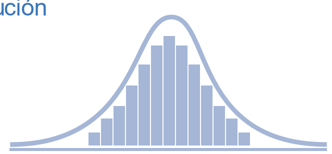
Tipos de inferencia (Sterba, 2009)

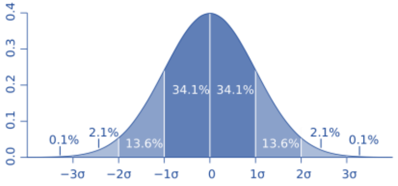
$$\mu$$

La población
finita

$$\hat{x}$$

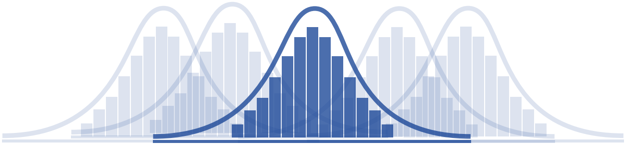
La distribución
muestral





$$\bar{x}_1$$

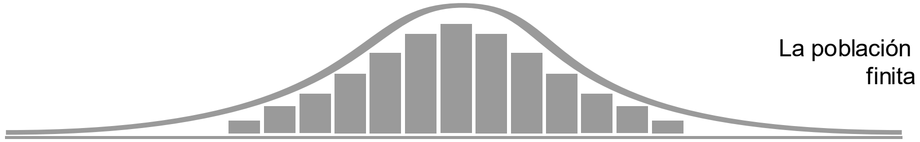
Una muestra
aleatoria



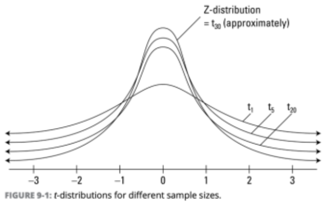
La población infinita,
o de valores posibles

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

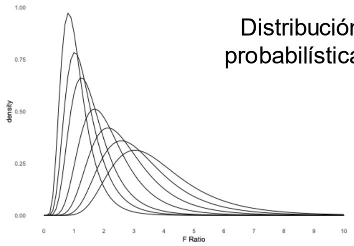
Mecanismo
generador de datos



La población
finita

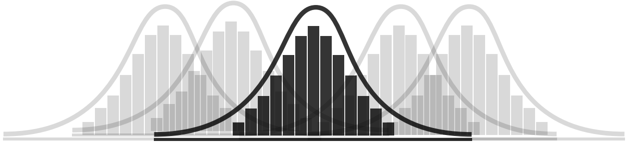


$$t, F$$



Distribución
probabilística

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$



Una muestra
(e.g., aleatoria,
condicionada,
o experimental)

Población finita
Distribución de valores de una población seleccionable.

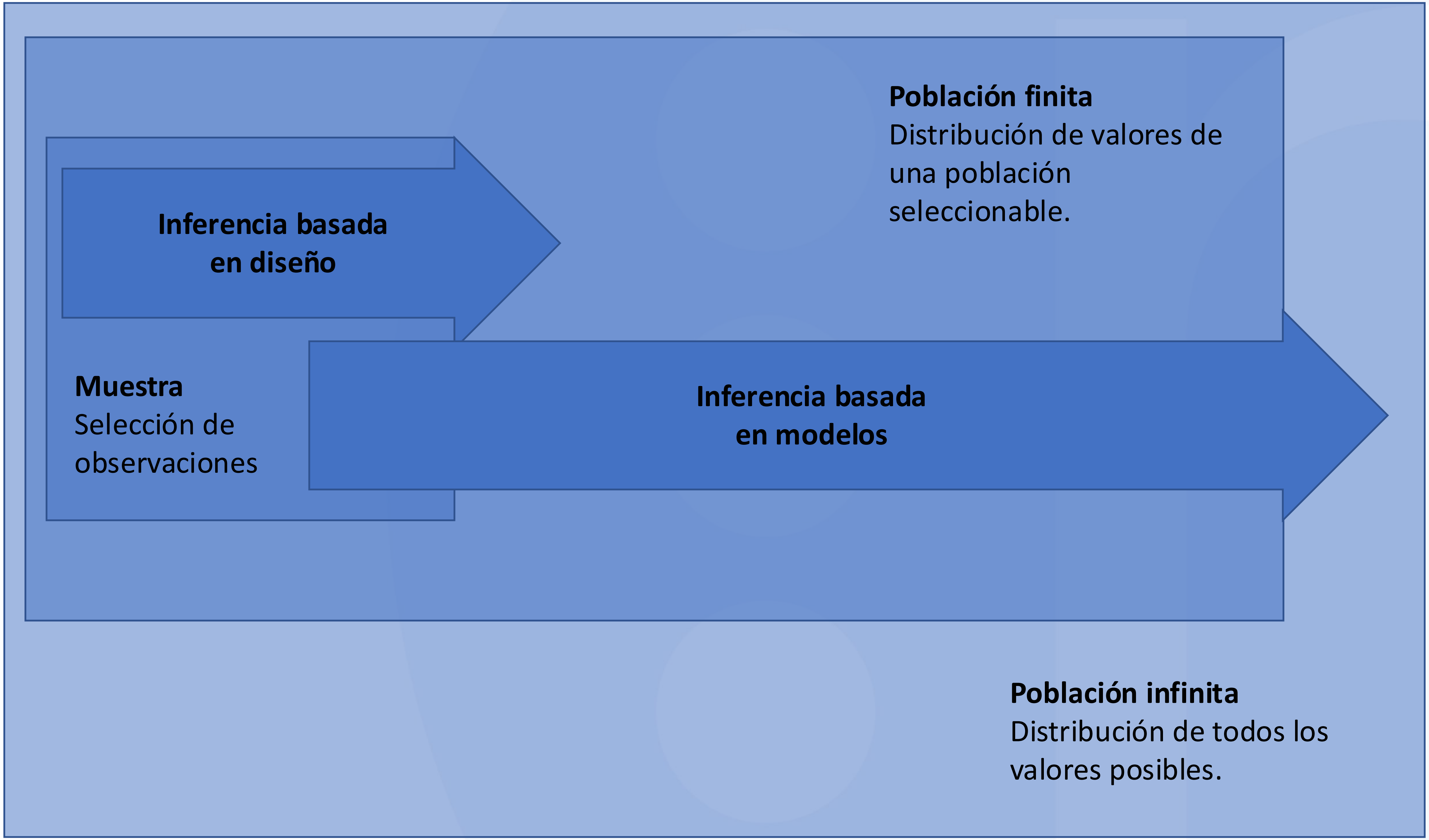
Muestra
Selección de observaciones

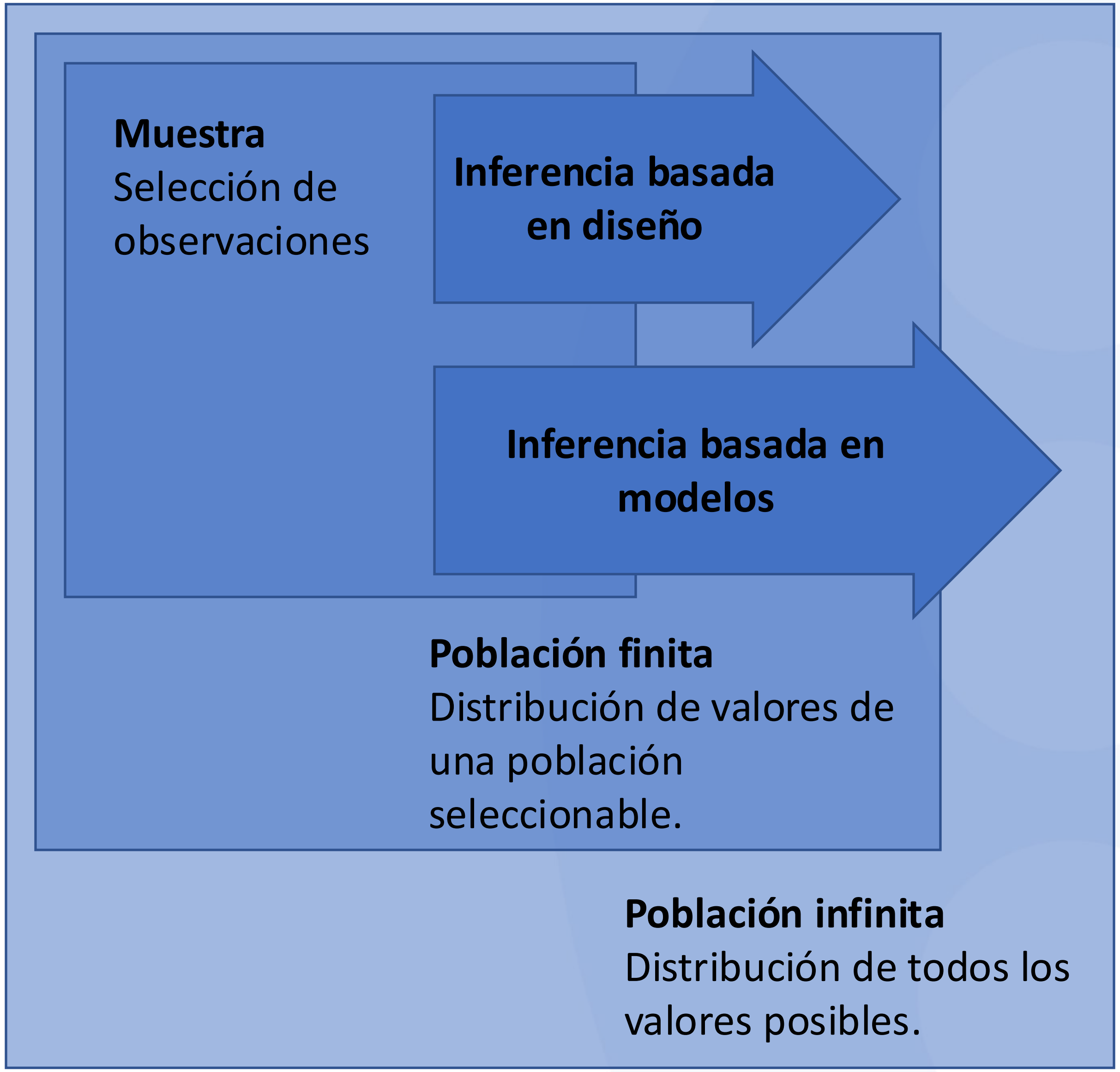
Población finita
Distribución de valores de una población seleccionable.

Inferencia basada en diseño

Muestra
Selección de observaciones


Población finita
Distribución de valores de una población seleccionable.
Muestra
Selección de observaciones
Inferencia basada en modelos
Población infinita
Distribución de todos los valores posibles.


Población finita
Distribución de valores de una población seleccionable.
Inferencia basada en diseño
Muestra
Selección de observaciones
Inferencia basada en modelos
Población infinita
Distribución de todos los valores posibles.



Nota: Tambien es posible realizar inferencias hibridas, basadas en modelos y en diseño. Este es el punto central del atículo de Sterba (2009).

Dos formas de inferencia

En esta clase hemos visto el ejercicio de realizar **inferencias basadas en la idea de obtener muestras, respecto a una población finita.** Con estas muestras, queremos estimar al parámetro poblacional, por medio del cálculo de estadígrafos. Debido a **cómo se comporta una distribución muestral de un estadígrafo**, podemos estimar al parámetro poblacional, con el estadígrafo.

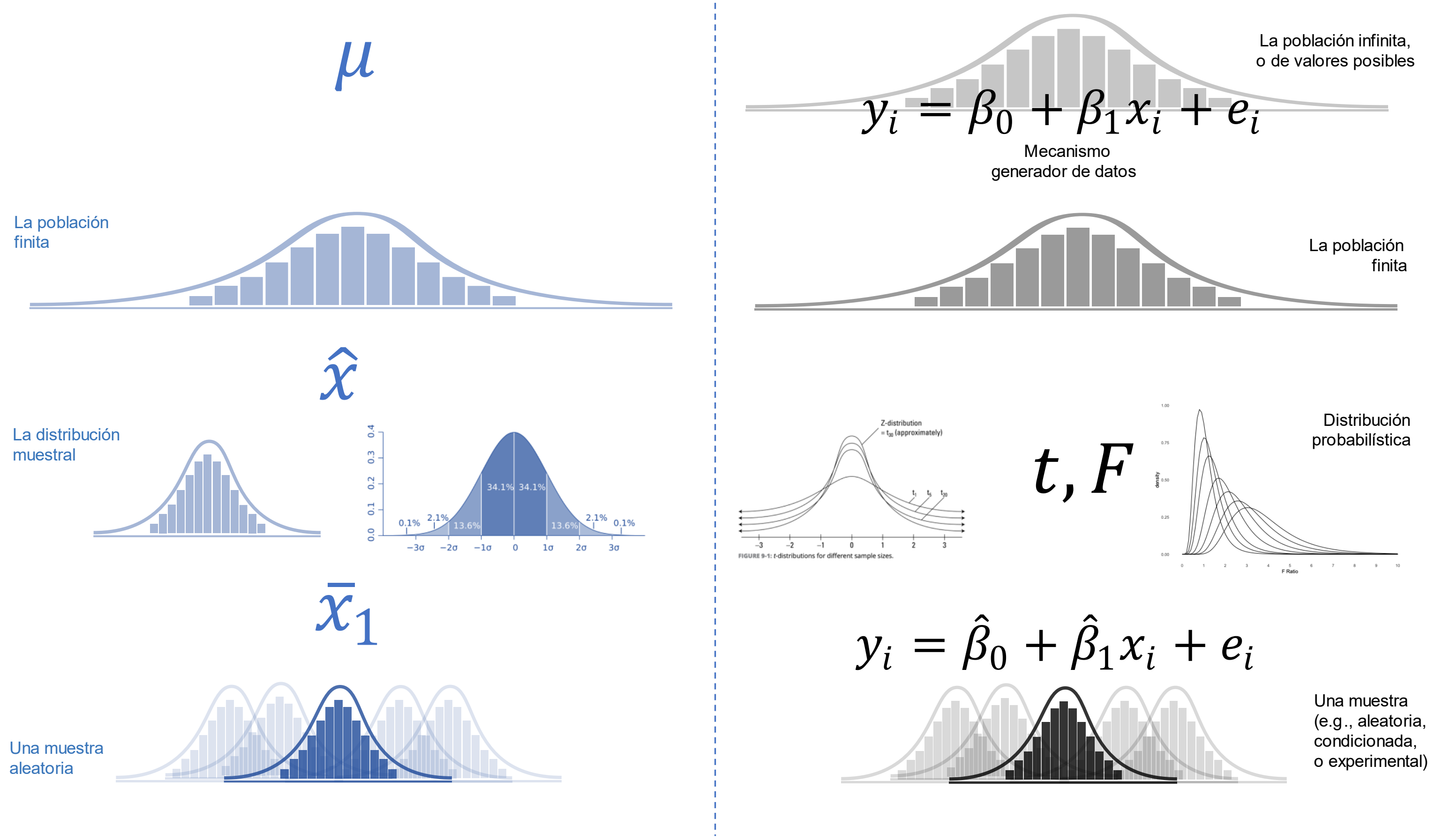
Existe una forma adicional de realizar inferencias, llamada **inferencia basada en modelos**. En este caso, hacemos conclusiones no respecto a una población finita, sino a una **población infinita**; respecto a la distribución total de valores posibles un parámetro (Sterba, 2009). En este caso, la muestra observada, y la población finita serían solo realizaciones de un **mecanismo generador de datos** (Rabe-Hesketh et al., 2012)

Tipos de inferencia (Sterba, 2009)

La inferencia basada en diseño es una muy buena herramienta para contestar preguntas respecto a atributos de la población finita (i.e., medias, porcentajes, y otras cifras).

En contraste, la inferencia basada en modelo tiende a ser empleada para estimados más complejos, los cuales serán empleados en argumentos que tienen más pasos (i.e., relaciones entre variables, heterogeneidad de relaciones, efectos experimentales, y evaluación de programas).

El primer marco depende del diseño en gran medida. El segundo, depende de que el "modelo" sea el mecanismo generador de datos.

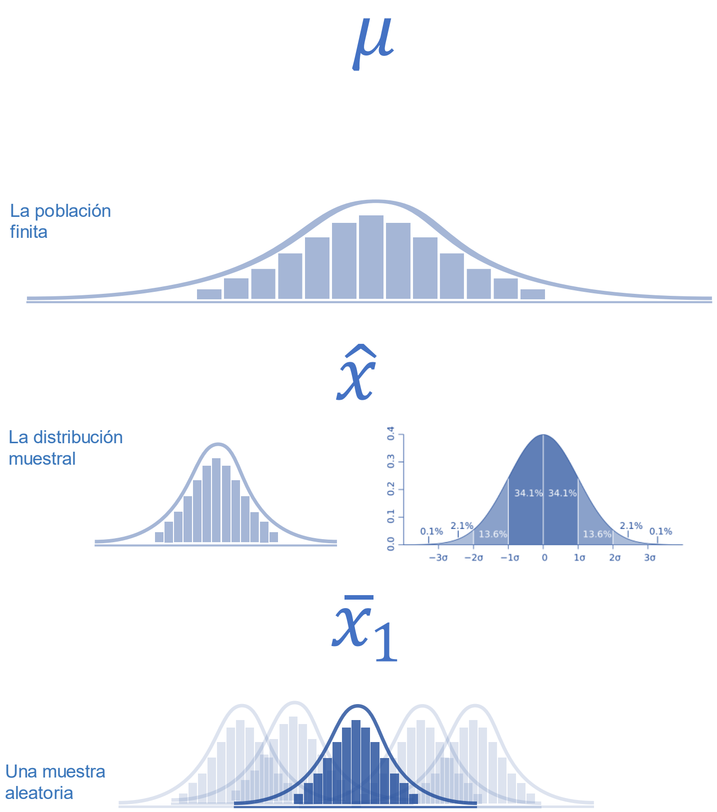


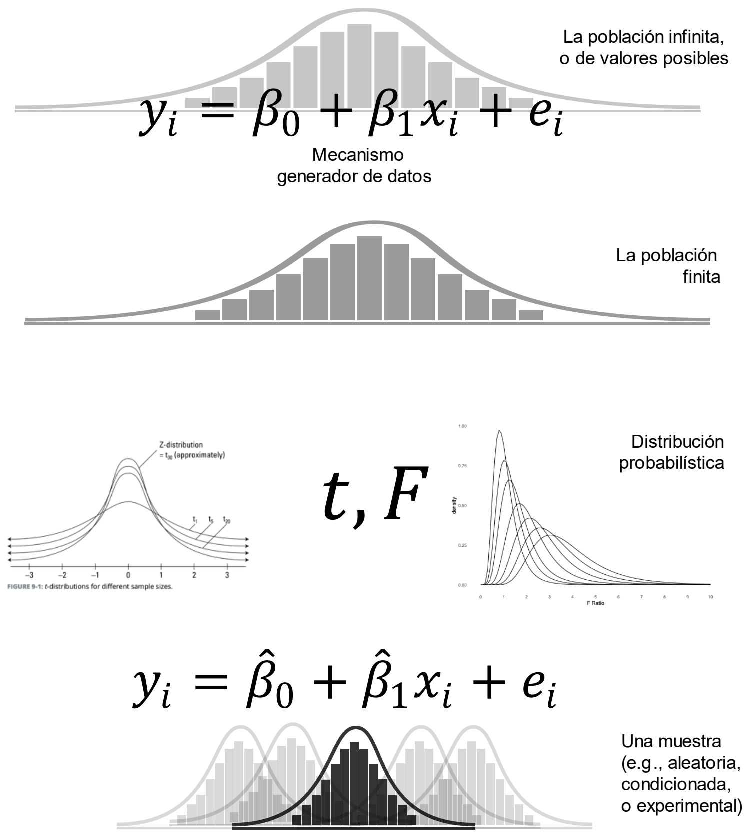
Model based inference

En la inferencia basada en modelo, el investigador tiene que:

- Plantear un modelo razonable (modelo estructural) (e.g., alguna forma de regresión, o regresiones).
- Plantear un modelo probabilístico (e.g., el modelo de errores)
- El modelo debe incluir al mecanismo de selección de las observaciones (i.e., estratificaciones, anidaciones, desproporciones, e incluso atriciones ver Enders (2011)).

Las inferencias realizadas, dependen de que tan correcto sea el modelo para expresar al mecanismo generador de los datos.





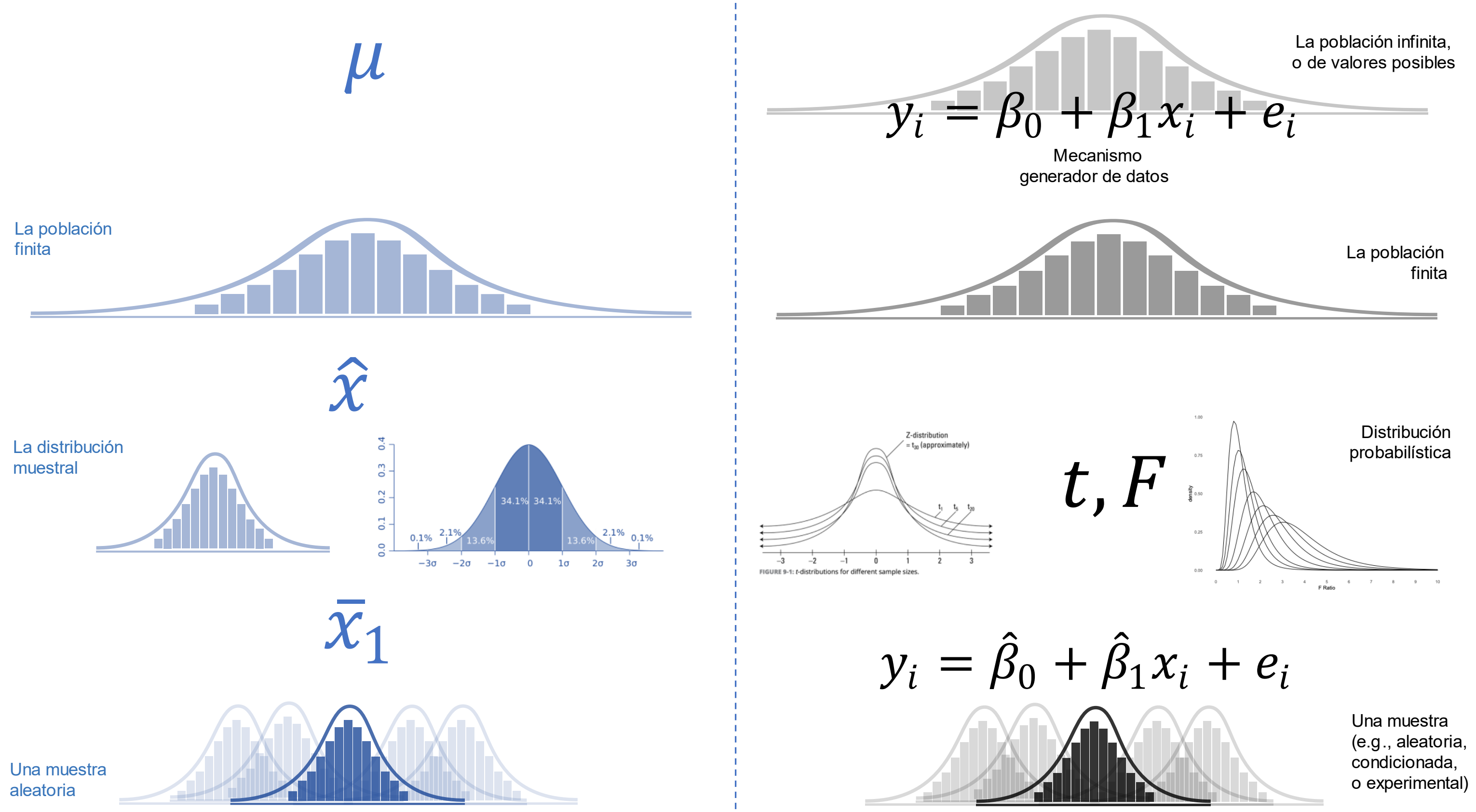
Una inferencia basada en modelo de forma pura aplicaría un modelo de regresión (o regresión generalizada), incluyendo una serie de términos que expresaran que sabemos del proceso que genera a los datos.

lm(y ~ 1 + x + w + xz + wz + ...)

Design based inference

Hasta ahora, solo hemos descansado en la idea de que la inferencia basada en modelo aplica para los muestreos aleatorios simples. En el caso de diseños muestrales estratificados, y de observaciones anidadas, y con un mecanismo de selección conocido, estos aspectos son incorporados a la estimación de puntos estimados y errores que acompañan los estimados (Sterba, 2009; Binder & Roberts, 2006).

La calidad de las inferencias realizadas y sus límites, están condicionados por el diseño muestral (i.e., el marco muestral, las características del diseño).



Una inferencia basada en diseño de forma pura es producir descriptivos, empleando un objeto que incluya al diseño:

```
srvyr::as_survey_design()
```

Limitaciones de los marcos de inferencia

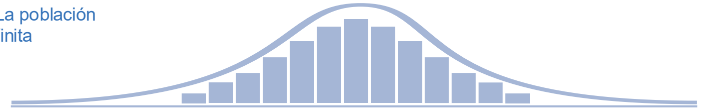
what are the limitations of the model-based and design-based frameworks, and how can these be overcome?

We showed that the model-based framework's central limitation lies in the need to tediously condition on all complex sampling features in model specification,

…and the design-based framework's central limitation lies in the inability to address analytic/causal hypotheses and account for nonsampling errors.

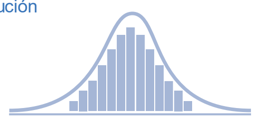
μ

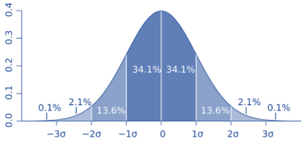
La población finita



$\hat{x}$

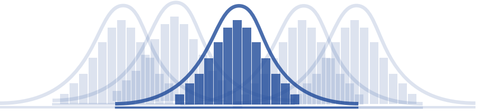
La distribución muestral





$\bar{x}_1$

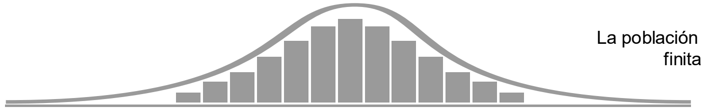
Una muestra aleatoria



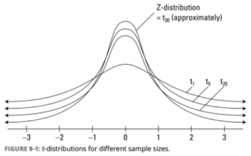
La población infinita, o de valores posibles

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

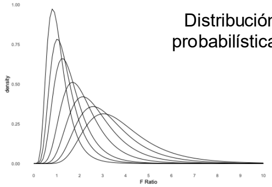
Mecanismo generador de datos



La población finita

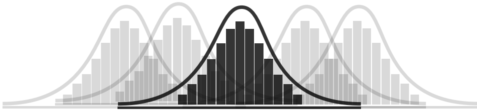


t, F



Distribución probabilística

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$



Una muestra (e.g., aleatoria, condicionada, o experimental)

Muchas gracias!

Carrasco, D., PhD,
Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile
https://dacarras.github.io/

Referencias

Binder, D., & Roberts, G. (2006). Approaches for analyzing survey data: a discussion. ASA Section on Survey Research Methods, 2771–2778. http://www.amstat.org/Sections/Srms/Proceedings/y2006/Files/JSM2006-000480.pdf

Cobb, G. W. (2007). The Introductory Statistics Course: A Ptolemaic Curriculum. Technology Innovations in Statistics Education, 1(1), 1–16. http://www.escholarship.org/uc/item/6hb3k0nz%5Cnpapers3://publication/uuid/426B647B-BAD9-4278-8F02-990BD513D5FA

Carrasco, D., Banerjee, R., Treviño, E., & Villalobos, C. (2020). Civic knowledge and open classroom discussion: explaining tolerance of corruption among 8th-grade students in Latin America. Educational Psychology, 40(2), 186–206. https://doi.org/10.1080/01443410.2019.1699907

Devlin, D., Guo, X., Kunin, D., & Xiang, D. (2018). Seeing Theory. Brown University. https://seeing-theory.brown.edu/doc/seeing-theory.pdf

Enders, C. K. (2011). Analyzing Longitudinal Data With Missing Values. Rehabilitation Psychology, 56(4), 267–288. https://doi.org/10.1037/a0025579

Huck, S. W. (2009). Statistical Misconceptions. Routledge.

Rabe-Hesketh, S., & Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition (3rd ed.). Stata Press.

Ross, K. N. (2005). Sample design for educational survey research. In Quantitative research methods in education planning. International Institute for Educational Planning (IIEP). UNESCO. http://www.unesco.org/iiep/PDF/TR_Mods/Qu_Mod3.pdf

Sterba, S. K. (2009). Alternative Model-Based and Design-Based Frameworks for Inference From Samples to Populations: From Polarization to Integration. In Multivariate Behavioral Research (Vol. 44, Issue 6, pp. 711–740). https://doi.org/10.1080/00273170903333574

Stigler, S. M. (2016). The seven pillars of statistical wisdom. Harvard University Press.

Inferencia empleando experimentos

Métodos de investigación cuantitativa

UAH-EMAPE

Enfoques Metodológicos para el Análisis de

Políticas Educativas

Septiembre 28

2024

Profesor invitado

Carrasco, D., PhD,

Centro de Medición MIDE UC

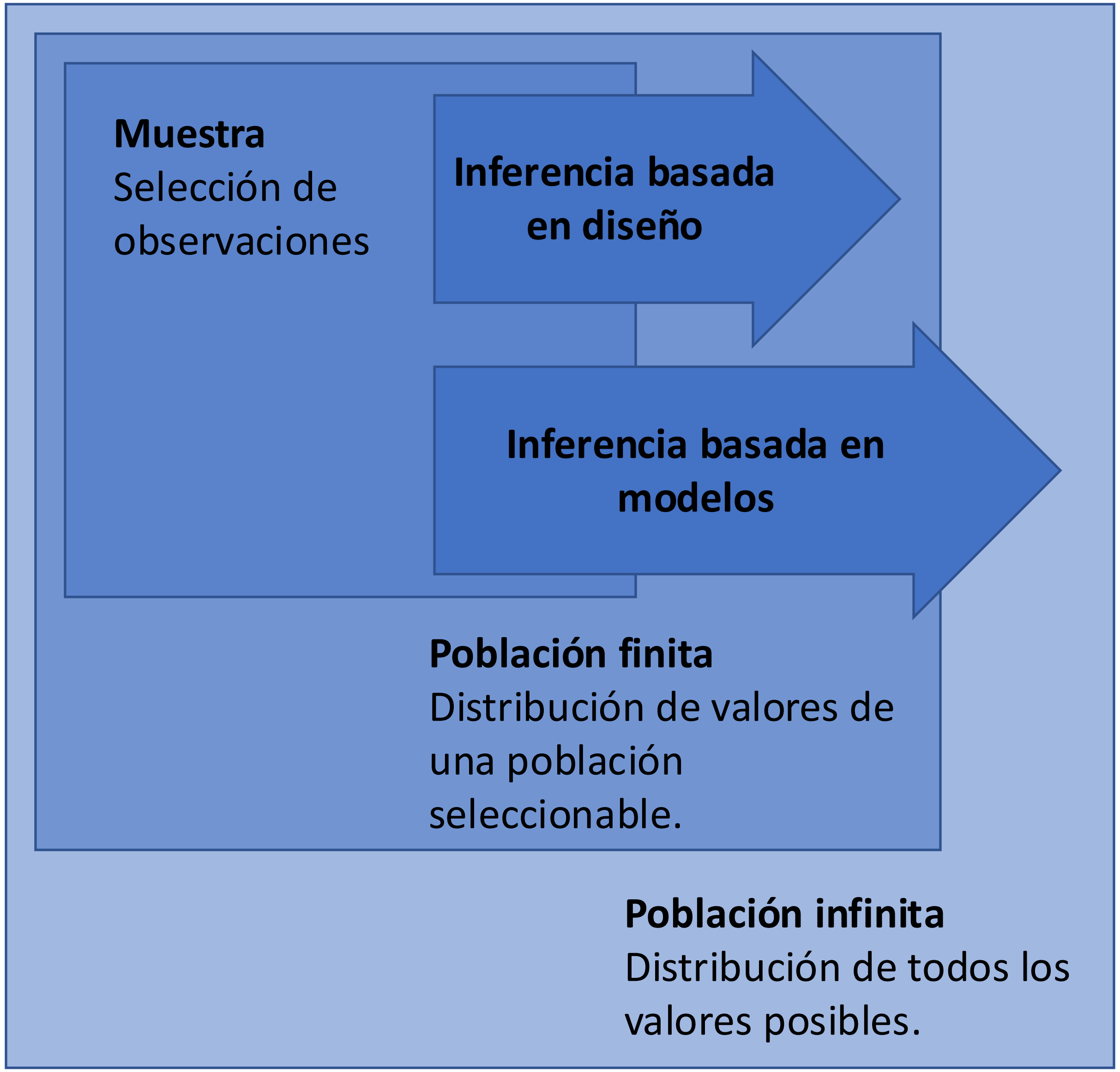
Pontificia Universidad Católica de Chile

Centro UC Medición - MIDE

Metodología Cuantitativa

Inferencias

Dos tipos de inferencia: basada en diseño, y basada en modelo



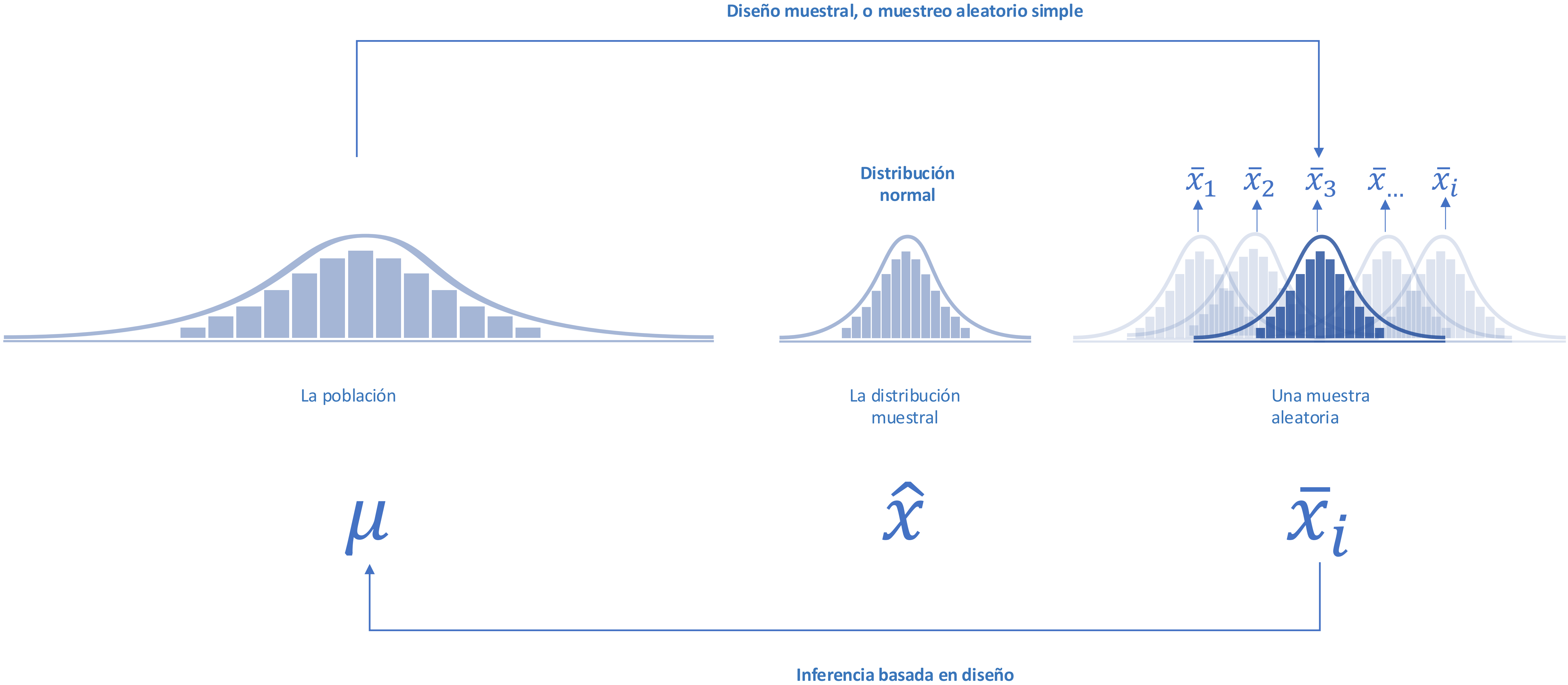
Nota: Tambien es posible realizar inferencias hibridas, basadas en modelos y en diseño. Este tipo de inferencias excede los contenidos del curso. Este es el punto central del atículo de Sterba (2009).

Dos formas de inferencia

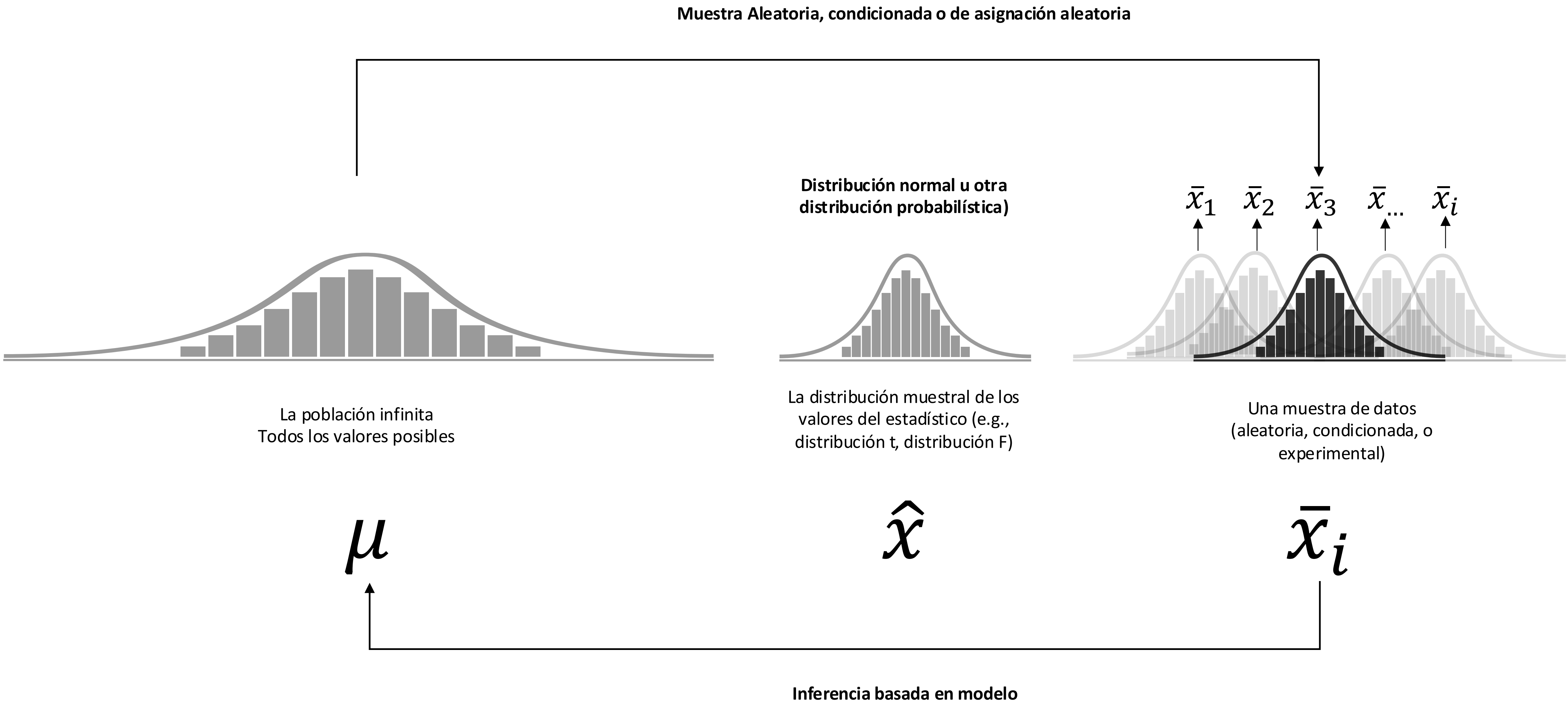
En clases anteriores, hemos visto el ejercicio de realizar **inferencias basadas en la idea de obtener muestras, respecto a una población finita.** Con estas muestras, queremos estimar al parámetro poblacional, por medio del cálculo de estadígrafos. Debido a **cómo se comporta una distribución muestral de un estadígrafo**, podemos estimar al parámetro poblacional, con el estadígrafo.

Existe una forma adicional de realizar inferencias, llamada **inferencia basada en modelos**. En este caso, hacemos conclusiones no respecto a una población finita, sino a una **población infinita**; respecto a la distribución total de valores posibles un parámetro (Sterba, 2009). En este caso, la muestra observada, y la población finita serían solo realizaciones de un **mecanismo generador de datos** (Rabe-Hesketh et al., 2012)

Inferencia basada en diseño

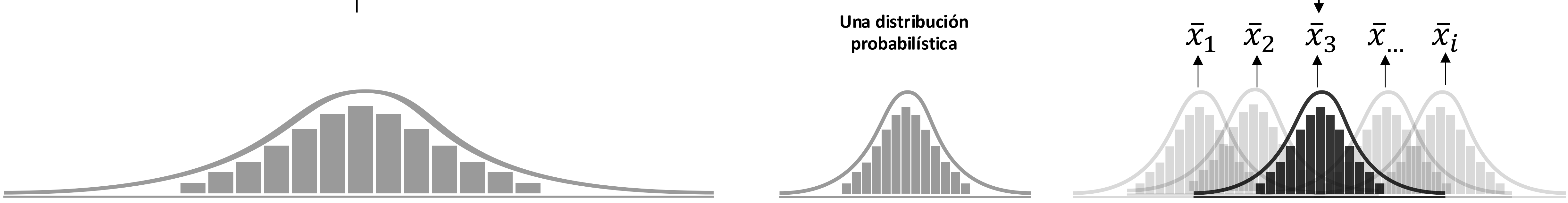


Inferencia basada en modelo



Inferencia basada en modelo (con regresión)

Muestra Aleatoria, condicionada o de asignación aleatoria



La población infinita
Todos los valores posibles

La distribución muestral de los valores del estadístico (e.g., distribución t, distribución F)

Una muestra de datos (aleatoria, condicionada, o experimental)

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

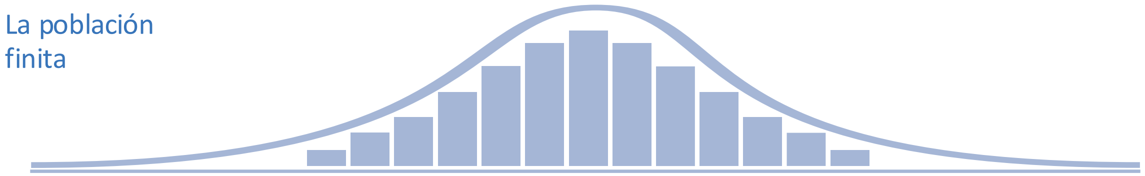
$$t, F$$

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$

Inferencia basada en modelo

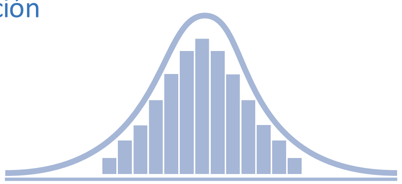
μ

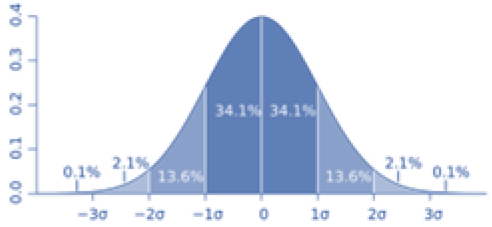
La población finita



$\hat{x}$

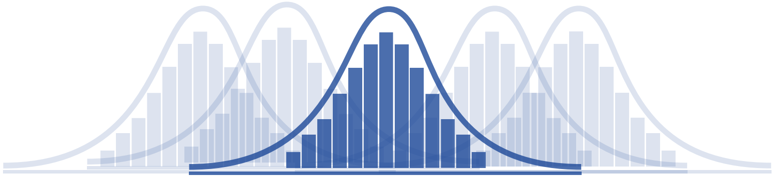
La distribución muestral





$\bar{x}_1$

Una muestra aleatoria

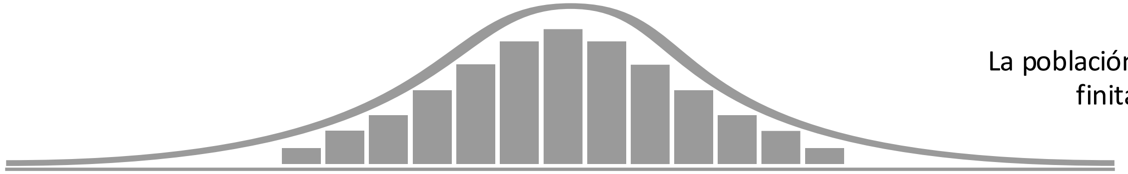


La población infinita, o de valores posibles

μ

Mecanismo generador de datos

La población finita



$\hat{x}$

Distribución probabilística

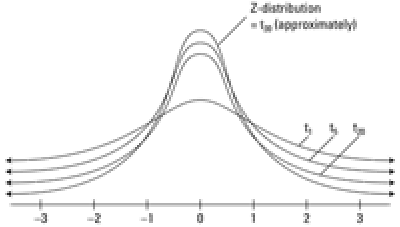
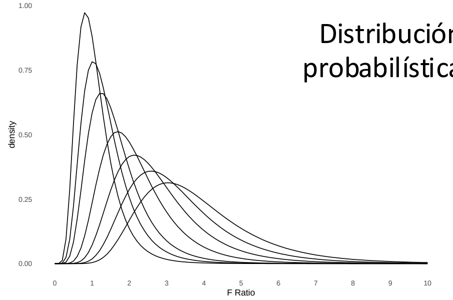
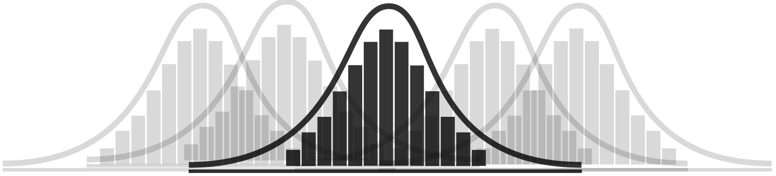


FIGURE 9-1: t-distributions for different sample sizes.



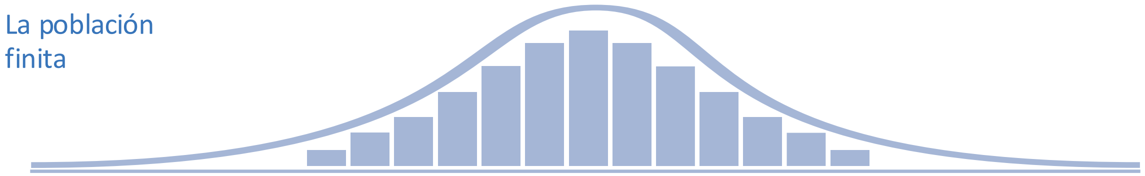
$\bar{x}_1$

Una muestra (e.g., aleatoria, condicionada, o experimental)



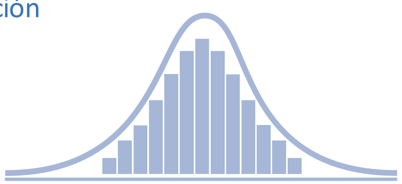
$$\mu$$

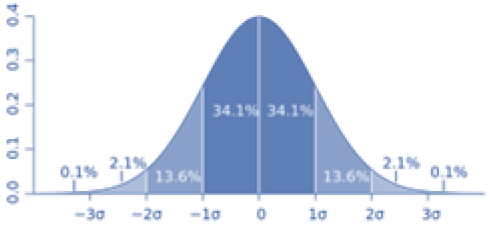
La población finita



$$\hat{x}$$

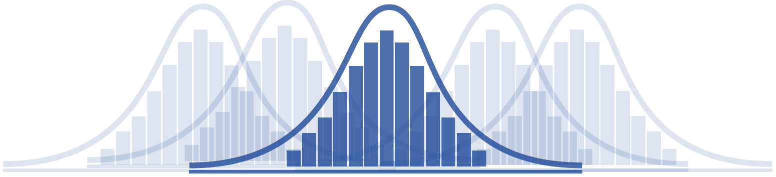
La distribución muestral





$$\bar{x}_1$$

Una muestra aleatoria

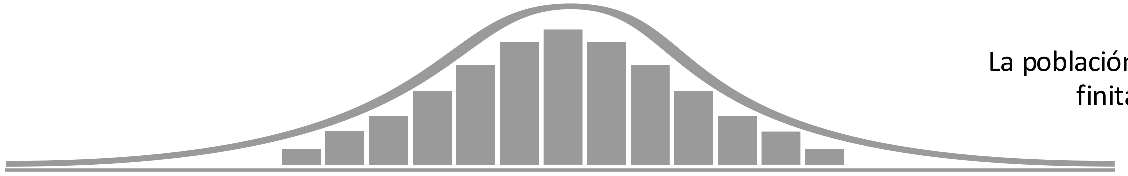


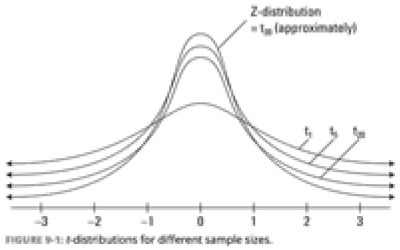
La población infinita, o de valores posibles

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

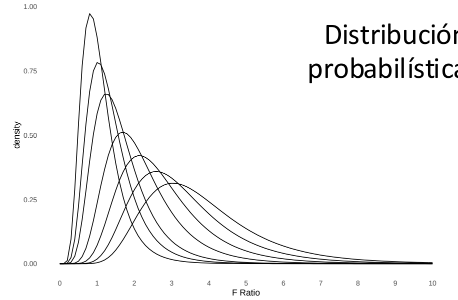
Mecanismo generador de datos

La población finita



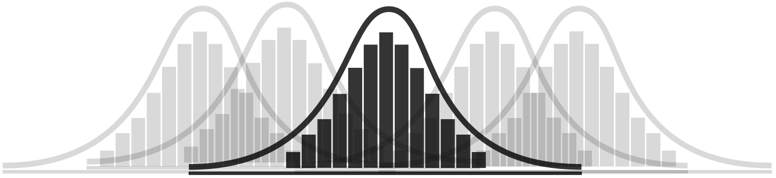


$$t, F$$



Distribución probabilística

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$

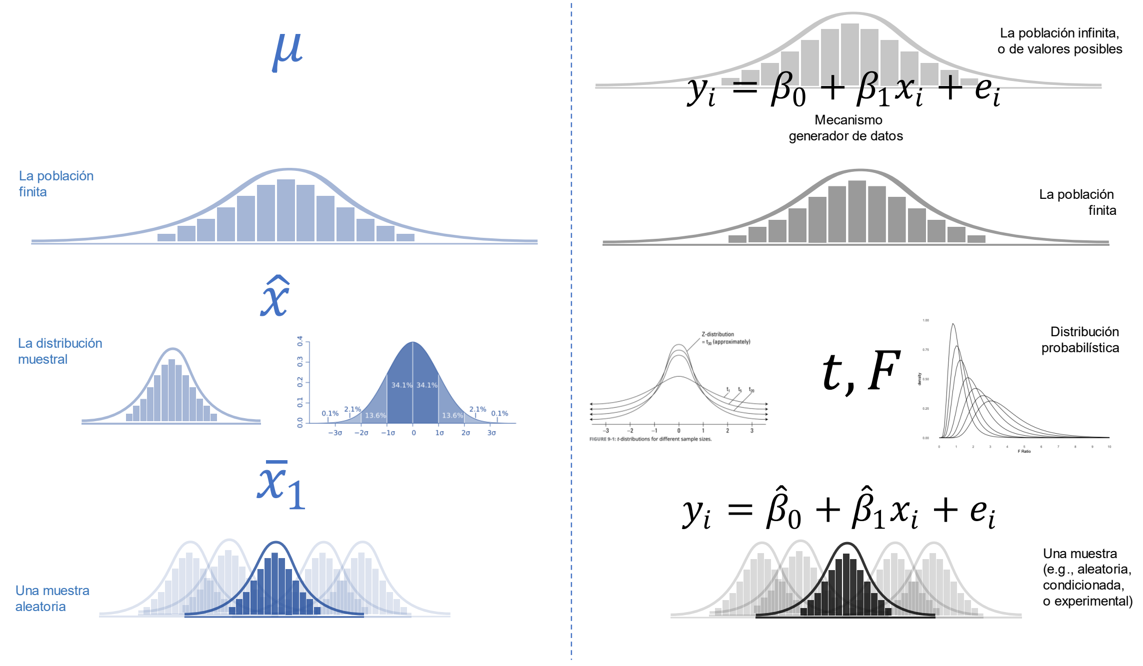


Una muestra (e.g., aleatoria, condicionada, o experimental)

Inferencias basadas en diseño y en modelo

La inferencia basada en **diseño** no requiere de un modelo. Se apoyo en las propiedades de la **distribución muestral**, para la construcción de intervalos de confianza, y la realización inferencias intervalares. Dado que la población finita (i.e., el conjunto seleccionable de observaciones existe), es posible generar muestras sobre estos datos.

Por su parte, la inferencia basada en **modelo** asume que un modelo (i.e., un **mecanismo generador de datos**) puede producir los datos de la población finita. Y, teniendo una muestra aleatoria, una muestra condicionada, o una muestra experimental, podemos ajustar los datos a un modelo, y realizar inferencias respecto a los parámetros del mecanismo generador de datos.



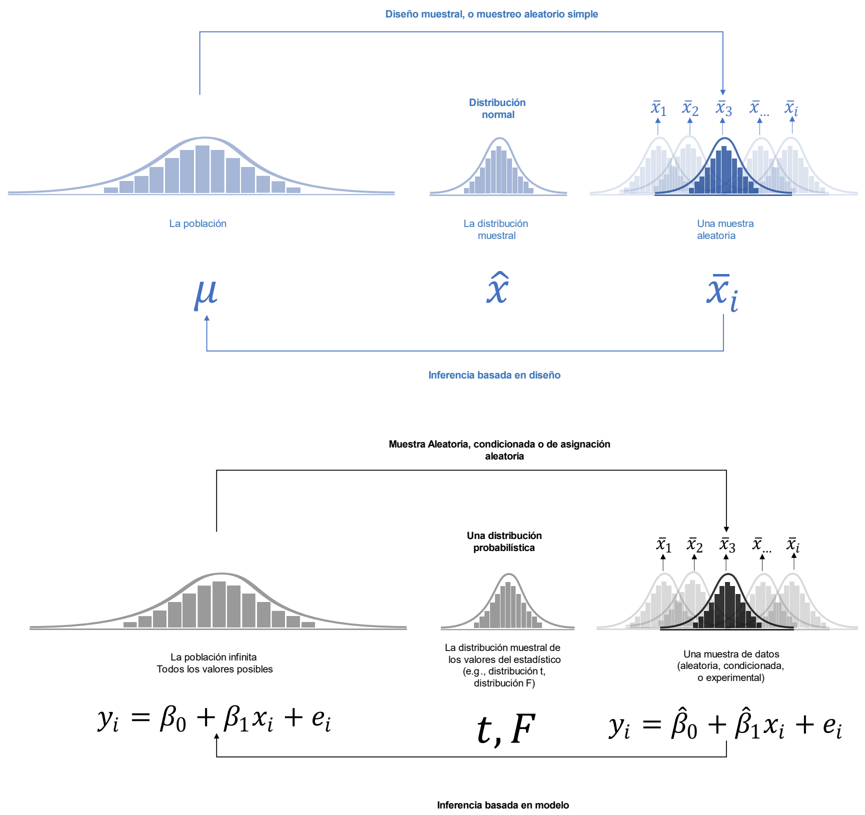
El foco de la inferencia basada en **diseño** es la **población finita y sus parámetros**. En el caso de la inferencia basada en **modelos**, el foco de la inferencia son los **parámetros del mecanismo generador de datos.**

Inferencia basada en modelo

A continuación vamos a revisar un ejemplo de inferencia basada en modelo. En particular, vamos a emplear un ejemplo de un estudio experimental. Y sobre estos datos vamos a realizar un contraste de hipótesis.

Vamos a realizar todos los pasos indicados por Huck (2012, p131), indicando:

1) Hipótesis nula
2) Hipótesis alternativa
3) Definición un nivel de significancia
4) Calcular resultados
5) Comparación a criterio (e.g., t, F)
6) Realizar una conclusión



Cobb (2007) plantea que hay fundamentalmente dos tipos de inferencia. Las que descansan en la distribución muestral (basadas en diseño), y las que descansan en la distribución de asignaciones aleatorias. Esta segunda distribución es propia de los estudios experimentales.

Metodología Cuantitativa

Contraste de hipótesis

Ejemplo con datos experimentales

Datos experimentales

Fondef ID21I10056 (study 2)

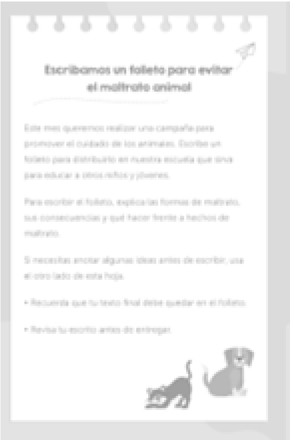
Modelo conceptual. Los escritores expertos manejan conocimiento del género en que escriben (cómo se escriben las cosas), y el tópico en que escriben (lo que se dice sobre las cosas) (McCutchen, 2011).

Cuando se evalúa a escritores inexpertos, o en formación (e.g., estudiantes) la tarea con la que se los evalúa requiere de que estos escritores manejen tanto al genero, como al tópico. En otras palabras, se requiere que el conocimiento del genero, y del tópico, se encuentren disponibles.

En caso contrario, los desempeños escritos de los estudiantes evaluados serían más pobres, y de menor desempeño, pues dependen del conocimiento previo de cada evaluado.

Se plantea un estudio experimental de asignación aleatoria, en que estudiantes de 5to grado son expuestos a un procedimiento de activación de genero y tópico previo a escribir.

Grupo control



Grupo experimental

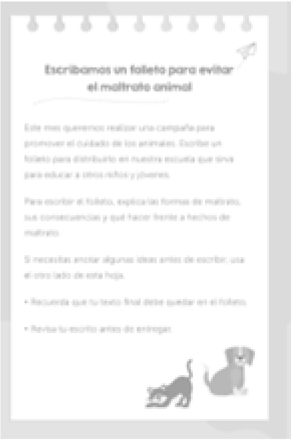
Datos experimentales

Muestra. Se reclutan 344 estudiantes de quinto grado (10 años aproximados), de ocho escuelas diferentes.

Procedimiento. Cada uno de los estudiantes es asignado a una condición experimental al interior de cada escuela. De esta forma, todos los estudiantes que participan de una escuela son aleatorizados entre las dos condiciones, son separados en salas de clases diferentes, y son sometidos al procedimiento experimental asignado.

Los estudiantes del grupo control (n = 175), reciben las instrucciones para escribir sobre el maltrato animal. Mientras que los estudiantes en la condición experimental (n = 169) cuentan con material no verbal que presenta contenidos relevantes del tópico. Además, cuentan con ejemplos de escritura, sobre el género en el que deben escribir.

Grupo control



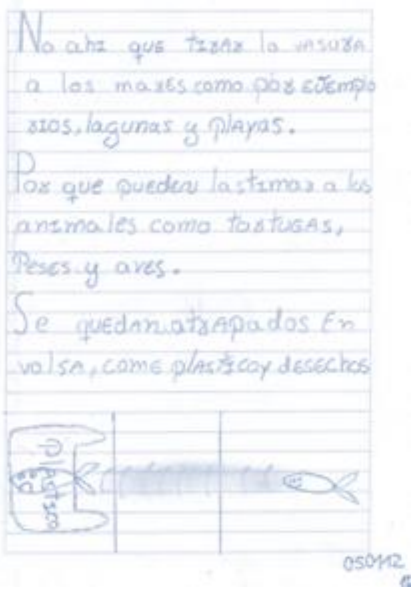
Grupo experimental

No se cumple propósito comunicativo → Propósito comunicativo es satisfecho


Lo efecto negativo
Cuando tiras plastico al mar los peses más pequeños pueden comerlo o enredarse en ella.

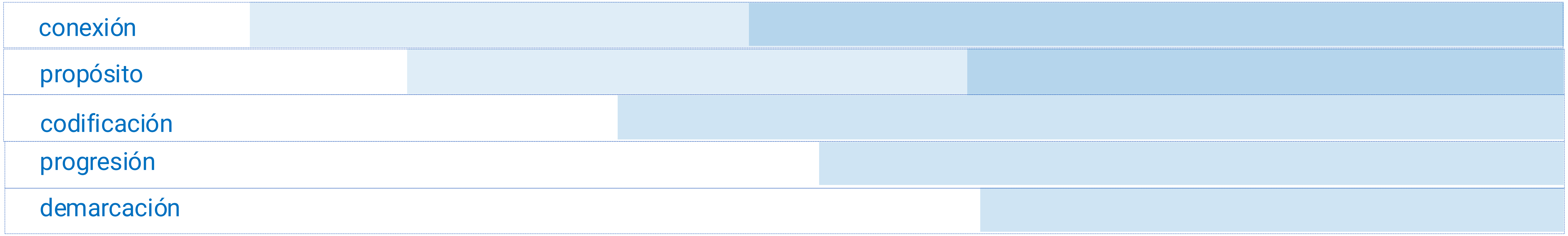
Min Score


los plastico fueron creados por el hombre, contamina el medio ambiente, mata animales marinos,
2) muerte de animales marinos etc
3) reciclar, no contaminar


No ahi que tirar la vasura a los mares como por ejemplo rios, lagunas y playas.
Por que quedan lastimas a los animales como tortugas, peses y aves.
Se quedan atrapados en volsa, como plasticoy desechos


Plastico en el mar:

Max Score



Tarea: escribir un brochure describiendo los efectos de la contaminación en el mar, y como puede ser prevenida.

Nota: diagrama referencial al mapa de ítem-personas.

Paso 1. Hipótesis nula

Expectativas del estudio. La expectativa del estudio es que los estudiantes expuestos al procedimiento de activación tendrán mejores desempeños en la tarea de escritura, que sus pares sin activación.

De no ser este el caso, los resultados entre ambos grupos no serían distinguibles. Para formular **hipótesis nula** sobre este estudio, tenemos que formalizar al escenario en que las expectativas de los investigadores no fueran ciertas.

$$H_0\colon\ \mu_c - \mu_t = 0$$

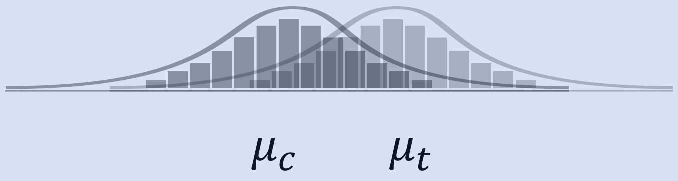
H_0 refiere a la hipótesis nula.

μ_c refiere a la media de puntajes del grupo control.

μ_t refiere a la media de puntajes del grupo tratado o grupo control.

Hipótesis nula

$$H_0\colon\ \mu_c - \mu_t = 0$$



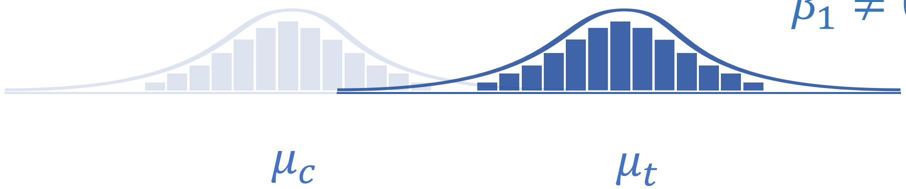
Modelo que genera los datos

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

Dónde,
$\beta_1 = 0$

Hipótesis alternativa

$$H_1\colon\ \mu_c - \mu_t \neq 0$$



Modelo que genera los datos

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

Dónde,
$\beta_1 \neq 0$

Paso 2. Hipótesis alternativa

La hipótesis alternativa contiene a las expectativas de los investigadores. Sin embargo, en términos formales contiene a todos los otros valores que la hipótesis no puede contener. Formalmente, la hipótesis nula, y la hipótesis alternativa son conjuntos disjuntos.

Esta hipótesis podría ser formulada de forma direccional (i.e., indicar que una cifra es mayor que la otra). En este ejemplo, emplearemos la forma no direccional.

$$H_1\colon \mu_c - \mu_t \neq 0$$

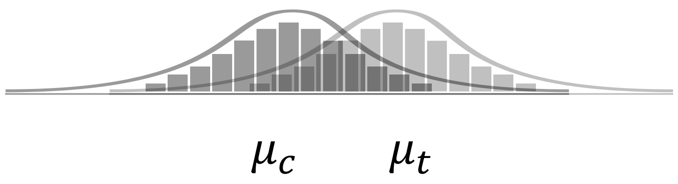
H_1 refiere a la hipótesis alternativa.

μ_c refiere a la media de puntajes del grupo control.

μ_t refiere a la media de puntajes del grupo tratado o grupo control.

Hipótesis nula

$$H_0\colon \mu_c - \mu_t = 0$$



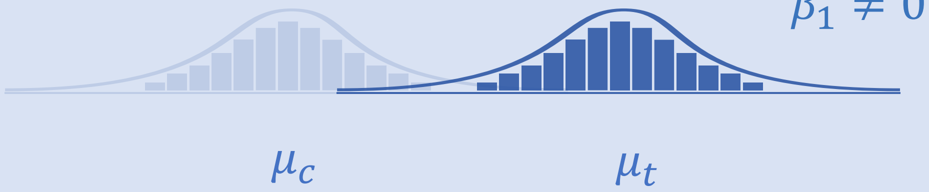
Modelo que genera los datos

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

Dónde,
$\beta_1 = 0$

Hipótesis alternativa

$$H_1\colon \mu_c - \mu_t \neq 0$$



Modelo que genera los datos

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

Dónde,
$\beta_1 \neq 0$

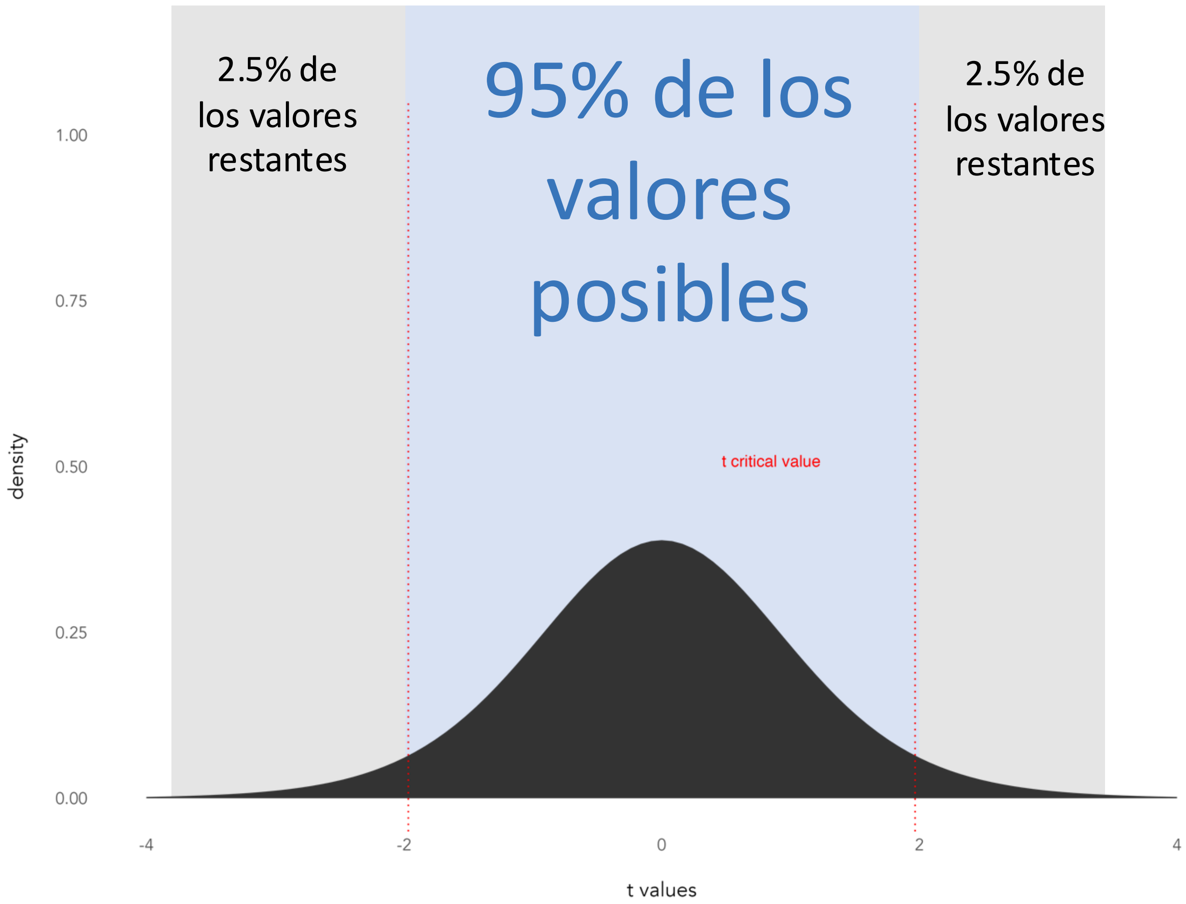
Paso 3. Nivel de significancia

En términos de notación el grado de significancia se le llama Alpha ($\alpha = 0.05$). Por convención este valor es de 0.05, y refiere a la cantidad de valores posibles, respecto a la distribución probabilística que vamos a considerar como no generados por el modelo contenido en la hipótesis nula.

Vamos a replantear lo anterior. Comenzamos con un mecanismo generador de datos, respecto a la variable de respuesta. Este genera todos los valores posibles. En este modelo, no hay diferencias entre el grupo experimental, y el grupo control.

El grado de significancia influye sobre qué tan distante ha de estar el estadígrafo de nuestras observaciones para que indiquemos que es poco plausible que nuestros datos hayan sido generados por el modelo de la hipótesis nula-

El grado de significancia influye sobre el valor critico que empleamos para realizar esta conclusión.



En este caso, si empleamos un $\alpha = 0.05$, el valor critico en la distribución t, para 343 grado de libertad (tamaño muestral, menos uno) es de 1.97. Podemos obtener este valor con el código R:

```
t_critic <- qt(1-0.05/2, df = 344-1)
```

Paso 4. Cálculo de resultados

Ajustamos un modelo de regresión sobre los puntajes asignados sobre las muestras de escrituras de los estudiantes, **score**. Este puntaje varía de 0 a 7, y el cual es la suma de puntajes recibidos en 5 indicadores de una rúbrica. Mayor puntaje, indica muestras de escritura mejor logradas.

La variable **act**, por su parte toma dos valores. Valor 0 que representa al grupo control, y valor 1, que representa el grupo experimental.

Los resultados del modelo de regresión nos entregan términos con los cuales podemos construir las medias de cada grupo. Si el modelo es:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i$$

La media del grupo control es $\hat{\beta}_0$, y la media del grupo experimental es $\hat{\beta}_0 + \hat{\beta}_1$. Por convención nos basta con evaluar sí $\hat{\beta}_1$es diferente de cero, para indicar que las medias de ambos grupos son diferentes.

```
> # load data
> library(dplyr)
> file_url <- url('https://github.com/dacarras/psi2301_20230517/raw/main/data_s2.rds')
> data_s2 <- readRDS(file_url) %>%
+       dplyr::glimpse()
Rows: 344
Columns: 9
$ id_i     <dbl> 110007, 110012, 110013, 1100
$ act      <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
$ score    <dbl> 6, 6, 4, 5, 4, 7, 4, 5, 7, 6
```

```
> # fit model
> lm(score ~ 1 + act, data = data_s2) %>%
+ summary()

Call:
lm(formula = score ~ 1 + act, data = data_s2)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.7219 -1.0743  0.2781  1.2781  2.9257 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4.0743     0.1342  30.359  < 2e-16 ***
act           0.6476     0.1915   3.382 0.000802 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.775 on 342 degrees of freedom
Multiple R-squared:  0.03237,	Adjusted R-squared:  0.02954 
F-statistic: 11.44 on 1 and 342 DF,  p-value: 0.0008022
```

Paso 4. Cálculo de resultados (modelo aplicado)

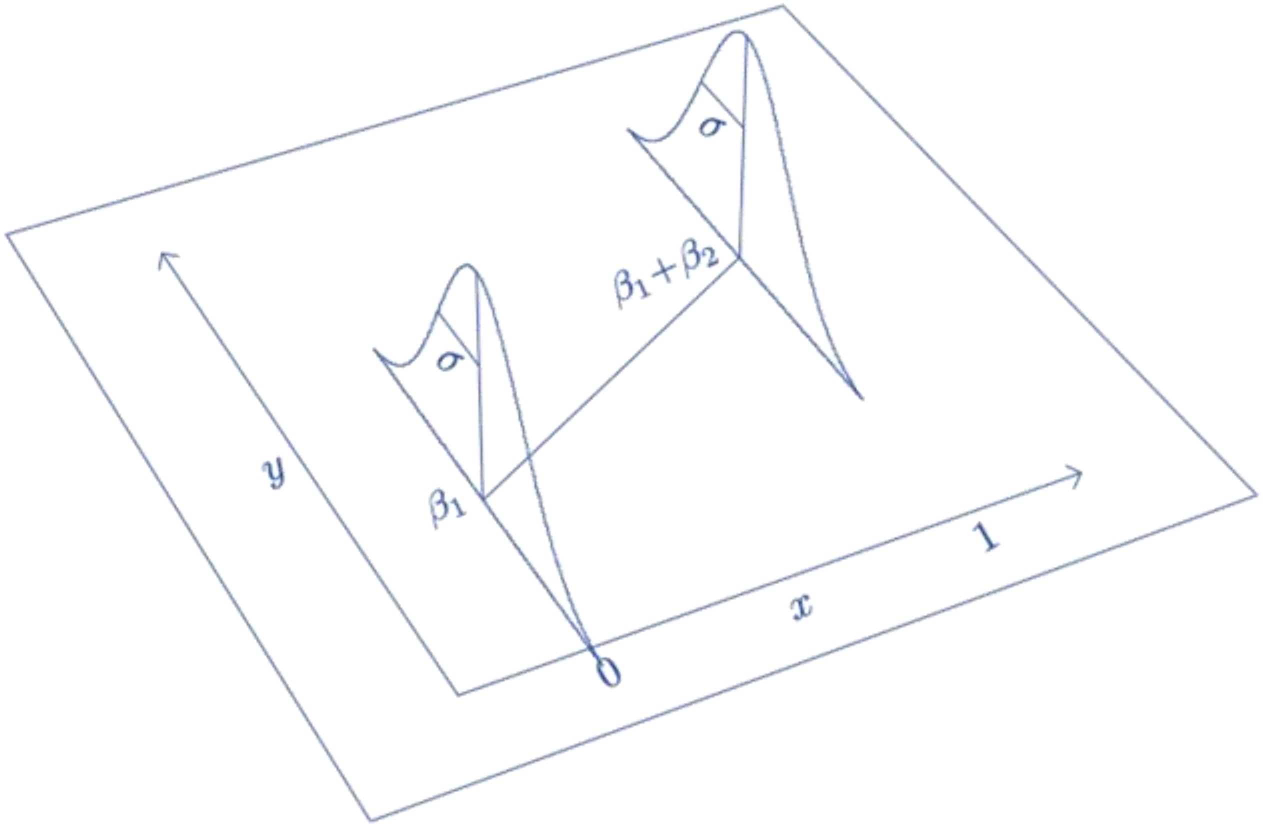


Figure 1.7: Illustration of simple linear regression with a dummy variable

Paso 5. Comparación a criterio

Lo que necesitamos saber, es si el $\hat{\beta}_1 \neq 0$, para tener evidencia a favor de que H_0: $\mu_c - \mu_t = 0$, no sea el caso.

Un modelo de regresión, con una covariable de dos valores, cero y uno, o tipo *dummy*, es un modelo equivalente a un escenario de comparación de medias. De esta manera, podemos descansar en el uso de una distribución probabilística de diferencia de medias. La distribución t, es una distribución probabilística que representa a la variabilidad de todos los valores posibles de diferencias de medias a diferentes tamaños muestrales. Dados los resultados, nuestra diferencia de medias entre el grupo experimental, y el grupo control, $\hat{\beta}_1$ es igual a .64, y equivale a un valor t de 3.38, por encima del valor critico (1.97). Estos resultados indican que nuestra diferencia de medias se encuentra por sobre el azar (i.e., por sobre los valores posibles que genera el modelo nulo).

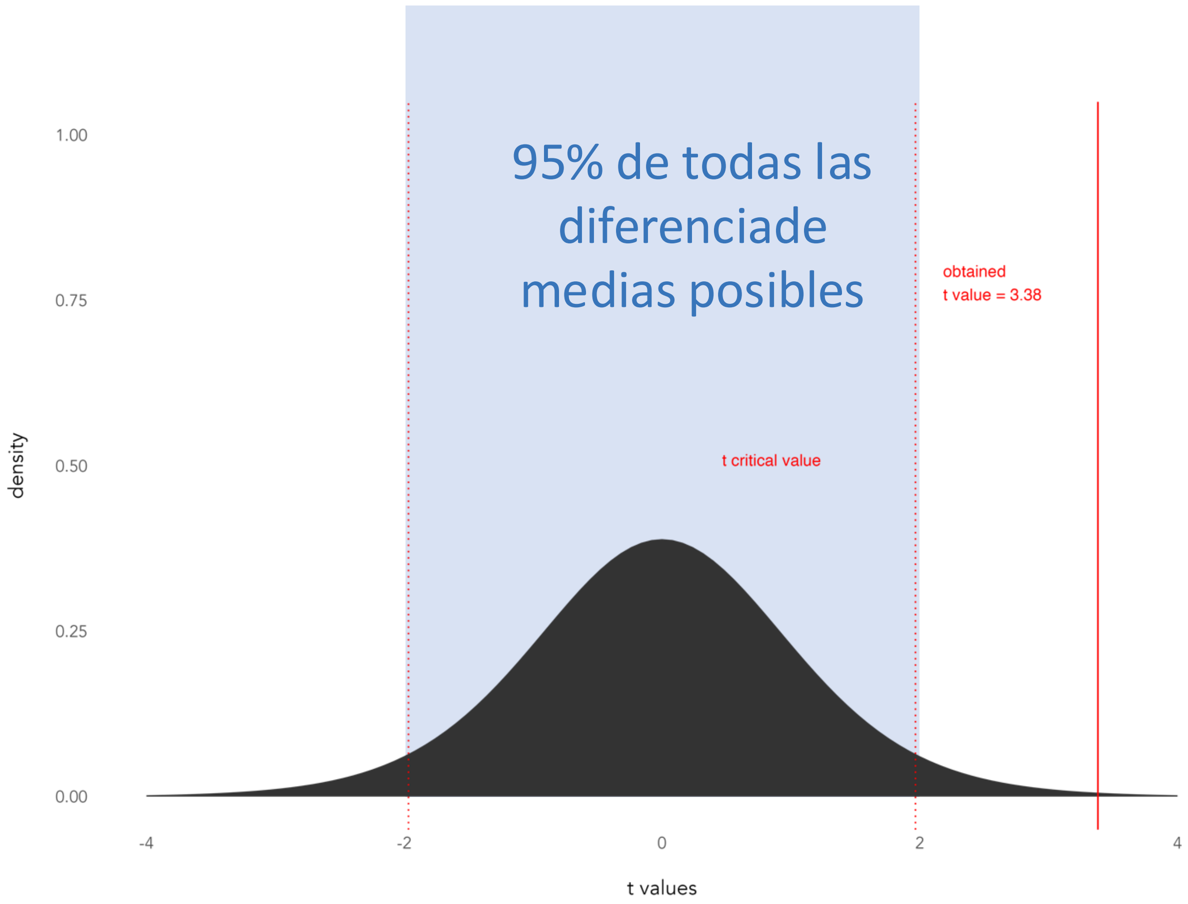
```
> # load data
> library(dplyr)
> file_url <- url('https://github.com/dacarras/psi2301_20230517/raw/main/data_s2.rds')
> data_s2 <- readRDS(file_url) %>%
+         dplyr::glimpse()
Rows: 344
Columns: 9
$ id_i      <dbl> 110007, 110012, 110013, 1100
$ act       <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1
$ score     <dbl> 6, 6, 4, 5, 4, 7, 4, 5, 7, 6
```

```
> # fit model
> lm(score ~ 1 + act, data = data_s2) %>%
+ summary()

Call:
lm(formula = score ~ 1 + act, data = data_s2)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.7219 -1.0743  0.2781  1.2781  2.9257 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4.0743     0.1342  30.359  < 2e-16 ***
act           0.6476     0.1915   3.382 0.000802 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.775 on 342 degrees of freedom
Multiple R-squared:  0.03237,	Adjusted R-squared:  0.02954 
F-statistic: 11.44 on 1 and 342 DF,  p-value: 0.0008022
```

Paso 5. Comparación a criterio

Lo que necesitamos saber, es si el $\hat{\beta}_1 \neq 0$, para tener evidencia a favor de que $H_0: \mu_c - \mu_t = 0$, no sea el caso.

Un modelo de regresión, con una covariable de dos valores, cero y uno, o tipo dummy, es un modelo equivalente a un escenario de comparación de medias. De esta manera, podemos descansar en el uso de una distribución probabilística de diferencia de medias. La distribución t, es una distribución probabilística que representa a la variabilidad de todos los valores posibles de diferencias de medias a diferentes tamaños muestrales. Dados los resultados, nuestra diferencia de medias entre el grupo experimental, y el grupo control, $\hat{\beta}_1$ es igual a .64, y equivale a un valor t de 3.38, por encima del valor critico (1.97). Estos resultados indican que nuestra diferencia de medias se encuentra por sobre el azar (i.e., por sobre los valores posibles que genera el modelo nulo).



Nuestros datos producen una diferencia de medias de t = 3.38, la cual presenta un p valor de “p <.001” en la distribución probabilística de todas las diferencias de medias posibles, para muestras de 344 casos.

Parece poco razonable plantear que el modelo nulo hubiera generado los datos observados, dado estos resultados.

Nuestros datos observados, se encuentran en esta ubicación en los datos generados por el modelo nulo.

dónde
$y_i = 4.07 + .65x_i + e_i$

Y dónde
$\beta_1 = .65, \mathrm{se} = \ .19, \mathrm{t} = 3.38, \mathrm{p} < \ .001$

Paso 6. Realizar una conclusión

Considerando los resultados observados, parece poco razonable que los datos observados puedan ser generados con el modelo nulo. De esta forma, canónicamente lo que se indica, es que **“rechazamos la hipótesis nula”** (Huck, 2012, p135).

No obstante, esto solo significa que de los dos conjuntos disjuntos, que dividen a todos los valores posibles de diferencias de medias, estamos eligiendo al modelo que creemos que sí genero nuestros datos, y estamos descartando un modelo que es poco plausible que haya generado nuestros datos.

Luego de ajustado del modelo, lo que nos queda es la descripción de resultados.

```
> # load data
> library(dplyr)
> file_url <- url('https://github.com/dacarras/psi2301_20230517/raw/main/data_s2.rds')
> data_s2  <- readRDS(file_url) %>%
+           dplyr::glimpse()
Rows: 344
Columns: 9
$ id_i       <dbl> 110007, 110012, 110013, 1100
$ act        <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
$ score      <dbl> 6, 6, 4, 5, 4, 7, 4, 5, 7, 6
```

```
> # fit model
> lm(score ~ 1 + act, data = data_s2) %>%
+ summary()

Call:
lm(formula = score ~ 1 + act, data = data_s2)

Residuals:
    Min      1Q  Median      3Q     Max
-4.7219 -1.0743  0.2781  1.2781  2.9257

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   4.0743     0.1342  30.359  < 2e-16 ***
act           0.6476     0.1915   3.382 0.000802 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.775 on 342 degrees of freedom
Multiple R-squared:  0.03237,	Adjusted R-squared:  0.02954
F-statistic: 11.44 on 1 and 342 DF,  p-value: 0.0008022
```

Descripción de resultados

Para evaluar el efecto de la activación del tópico y el género sobre escritores en formación, se implementó un estudio experimental. Se reclutaron 344 estudiantes de quinto grado, de 8 escuelas diferentes, y fueron asignados al azar a las condiciones experimentales. Los estudiantes del grupo control (n = 175), reciben las instrucciones para escribir sobre el maltrato animal. Mientras que los estudiantes en la condición experimental (n = 169) cuentan con material no verbal que presenta contenidos relevantes del tópico, y cuentan con ejemplos en pizarra sobre el género que deben emplear.

Se aplica un modelo de regresión lineal sobre los puntajes de escritura. Estos puntajes varían de 0 a 7 puntos, donde mayor puntaje indica mayor calidad en lo escrito, en términos de cumplimiento del propósito comunicativo. Estos puntajes los condicionamos de acuerdo a si los estudiantes pertenecen al grupo experimental (valor uno), o al grupo control (valor cero).

El modelo ajustado explica hasta un 3% de la variabilidad de los puntajes ($F(1,342) = 11.44$, $p < .001$). Se observa una diferencia de .64 entre el grupo experimental, y el grupo control, a favor del grupo experimental ($\beta_1 = .65, \mathrm{se} = .19, \mathrm{t} = 3.38, \mathrm{p} < .001$). En términos estandarizados, el tamaño de efecto observado es de tamaño pequeño (Cohen's d = .37).

```
> # load data
> library(dplyr)
> file_url <- url('https://github.com/dacarras/psi2301_20230517/raw/main/data_s2.rds')
> data_s2 <- readRDS(file_url) %>%
+         dplyr::glimpse()
Rows: 344
Columns: 9
$ id_i      <dbl> 110007, 110012, 110013, 1100
$ act       <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
$ score     <dbl> 6, 6, 4, 5, 4, 7, 4, 5, 7, 6
```

```
> # fit model
> lm(score ~ 1 + act, data = data_s2) %>%
+ summary()

Call:
lm(formula = score ~ 1 + act, data = data_s2)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.7219 -1.0743  0.2781  1.2781  2.9257 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4.0743     0.1342  30.359  < 2e-16 ***
act           0.6476     0.1915   3.382 0.000802 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.775 on 342 degrees of freedom
Multiple R-squared:  0.03237,	Adjusted R-squared:  0.02954 
F-statistic: 11.44 on 1 and 342 DF,  p-value: 0.0008022
```

Descripción de resultados

Para evaluar el efecto de la activación del tópico y el género sobre escritores en formación, se implementó un estudio experimental. Se reclutaron 344 estudiantes de quinto grado, de 8 escuelas diferentes, y fueron asignados al azar a las condiciones experimentales. Los estudiantes del grupo control (n = 175), reciben las instrucciones para escribir sobre el maltrato animal. Mientras que los estudiantes en la condición experimental (n = 169) cuentan con material no verbal que presenta contenidos relevantes del tópico, y cuentan con ejemplos en pizarra sobre el género que deben emplear.

Se aplica un modelo de regresión lineal sobre los puntajes de escritura. Estos puntajes varían de 0 a 7 puntos, donde mayor puntaje indica mayor calidad en lo escrito, en términos de cumplimiento del propósito comunicativo. Estos puntajes los condicionamos de acuerdo a si los estudiantes pertenecen al grupo experimental (valor uno), o al grupo control (valor cero).

El modelo ajustado explica hasta un 3% de la variabilidad de los puntajes (F(1,342) = 11.44, p <.001). Se observa una diferencia de .64 entre el grupo experimental, y el grupo control, a favor del grupo experimental ($\beta_1 = .65, se = .19, t = 3.38, p < .001$). En términos estandarizados, el tamaño de efecto observado es de tamaño pequeño (Cohen's d = .37).

```
> # load data
> library(dplyr)
> file_url <- url('https://github.com/dacarras/psi2301_20230517/raw/main/data_s2.rds')
> data_s2 <- readRDS(file_url) %>%
+          dplyr::glimpse()
Rows: 344
Columns: 9
$ id_i       <dbl> 110007, 110012, 110013, 1100
$ act        <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1
$ score      <dbl> 6, 6, 4, 5, 4, 7, 4, 5, 7, 6
```

```
> # fit model
> lm(score ~ 1 + act, data = data_s2) %>%
+ summary()

Call:
lm(formula = score ~ 1 + act, data = data_s2)

Residuals:
    Min      1Q  Median      3Q     Max 
-4.7219 -1.0743  0.2781  1.2781  2.9257 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4.0743     0.1342  30.359  < 2e-16 ***
act           0.6476     0.1915   3.382 0.000802 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.775 on 342 degrees of freedom
Multiple R-squared:  0.03237,	Adjusted R-squared:  0.02954 
F-statistic: 11.44 on 1 and 342 DF,  p-value: 0.0008022
```

Metodología Cuantitativa

Resumen

Constraste de hipótesis realizado

Contraste de hipótesis

Pasos para realizar un contraste de hipótesis:

1. Definir la hipótesis nula
2. Establecer la hipótesis alternativa
3. Seleccionar el nivel de significancia
4. Recoger y resumir los datos de una muestra
5. Definir el criterio para evaluar la evidencia
6. Realizar una decisión respecto a rechazar o retener la hipótesis nula

Pasos del contraste de hipótesis:

- Hipótesis nula

$$H_0:\ \mu_c - \mu_t = 0$$

- Hipótesis alternativa

$$H_1:\ \mu_c - \mu_t \neq 0$$

- Nivel de significancia

$$\alpha = 0.05$$

- Calcular resultados

```
lm(score ~ 1 + act, data = data_s2)
```

- Valor crítico en distribución probabilística

```
t_critic <- qt(0.975,df=n_sample-1)
```

- Decisión

Si t_obs > a t_critic, entonces es poco razonable que H_o haya generado los datos observados.

Muchas gracias!

Carrasco, D., PhD,
Centro de Medición MIDE UC
Pontificia Universidad Católica de Chile
https://dacarras.github.io/

Referencias

Cobb, G. W. (2007). The Introductory Statistics Course: A Ptolemaic Curriculum. Technology Innovations in Statistics Education, 1(1), 1–16. http://www.escholarship.org/uc/item/6hb3k0nz%5Cnpapers3://publication/uuid/426B647B-BAD9-4278-8F02-990BD513D5FA

Ellis, P. D. (2010). The Essential Guide to Effect Sizes. Cambridge University Press.

McCutchen, D. (2011). From novice to expert: Implications of language skills and writing-relevant knowledge for memory during the development of writing skill. Journal of Writing Research, 3(1), 51–68. https://doi.org/10.17239/jowr-2011.03.01.3

Rabe-Hesketh, S., & Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition (3rd ed.). Stata Press.

Sterba, S. K. (2009). Alternative Model-Based and Design-Based Frameworks for Inference From Samples to Populations: From Polarization to Integration. In Multivariate Behavioral Research (Vol. 44, Issue 6, pp. 711–740). https://doi.org/10.1080/00273170903333574