



# Programa de formación MACHINE LEARNING AND DATA SCIENCE MLDS

Facultad de  
**INGENIERÍA**





# Módulo 1

## Análisis y visualización de datos con Python

Unidad 3

Análisis avanzado de datos con Python

Clase sincrónica

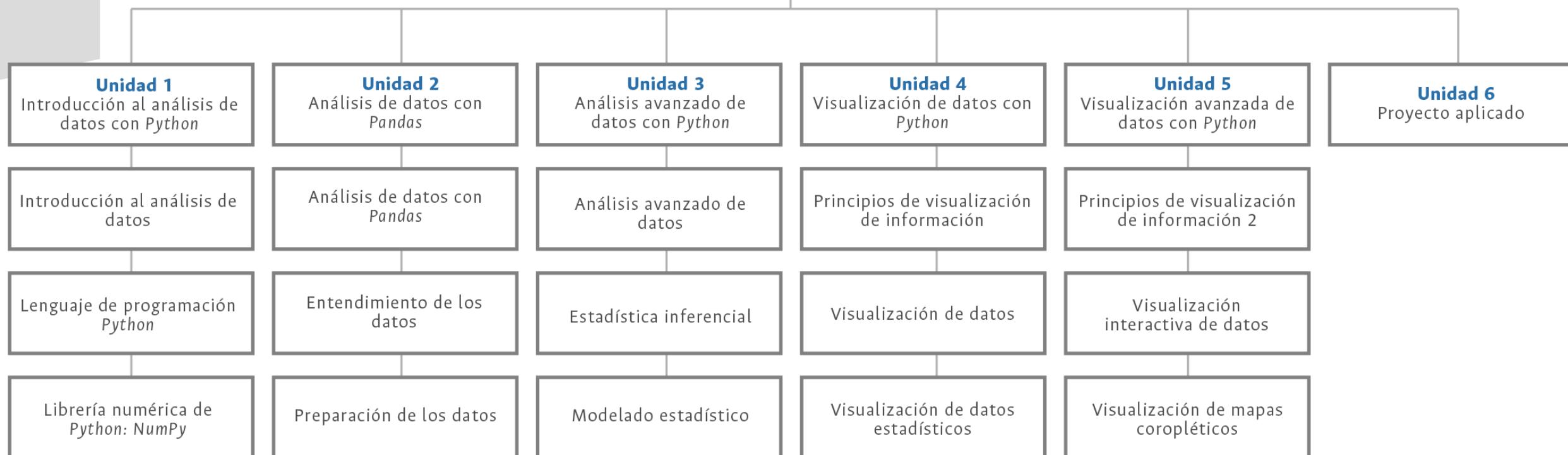
Felipe Restrepo Calle, PhD.

Facultad de  
**INGENIERÍA**





# Mapa de contenidos





## Tabla de contenidos

1

Estadística Inferencial



Intervalos de confianza



Test de hipótesis

2

Modelado estadístico



Análisis de correlaciones

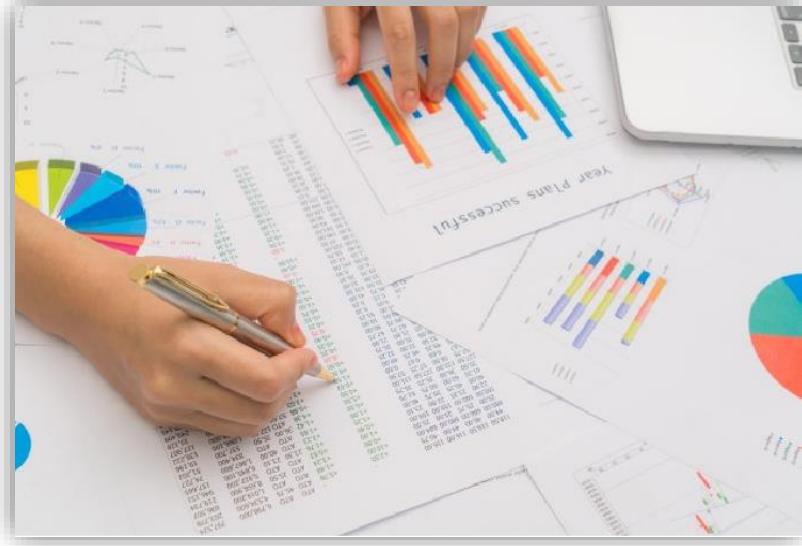


Análisis de regresiones

## Estadística Inferencial



## Estadística descriptiva e inferencial



### Estadística descriptiva

La estadística descriptiva se encarga de organizar, caracterizar y presentar descripciones o resúmenes de los datos con el propósito de apoyar tareas que involucren su entendimiento general. Esto se logra por medio de medidas numéricas y visualizaciones.



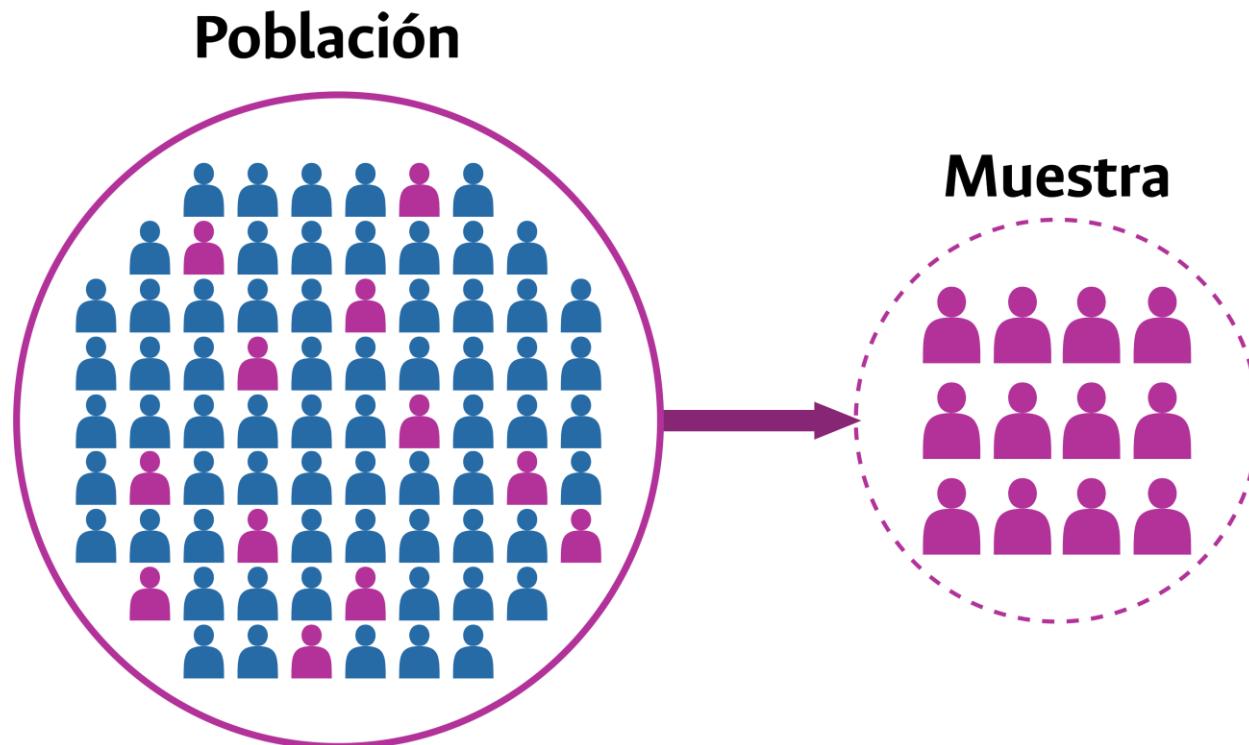
### Estadística inferencial

Es la rama de la estadística que busca inferir información de una población estadística, a partir de la información de una muestra, aplicando conceptos de probabilidad que se ajusten a la distribución de la población. Permite obtener conclusiones o inferencias a partir de los datos hasta llegar, incluso, a realizar predicciones.



## Muestra y población

La **población** es el conjunto de todos los objetos que se tiene la intención de observar. Generalmente, su tamaño es muy grande, por lo que es impráctico e inviable pretender observarla completamente.



La **muestra** es un subconjunto de la población, de la cual se parte para hacer inferencias sobre la población. La forma en que se realiza el muestreo es muy importante para obtener resultados fiables al aplicar métodos inferenciales.



## Parámetros y estimaciones

Un **parámetro poblacional** es una medida que representa a toda la población. Por lo general, es un valor desconocido.

	Parámetros poblacionales	Estadísticos muestrales
Proporción	$\pi$	$p$
Media aritmética	$\mu$	$\bar{x}$
Desviación estándar	$\sigma$	$s$
Varianza	$\sigma^2$	$s^2$

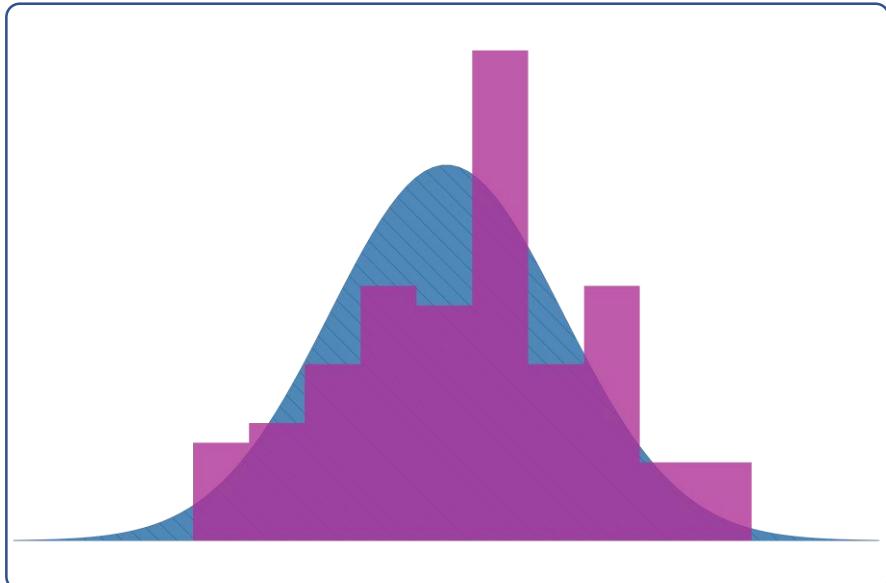
Un **estadístico muestral** es una medida que representa a una muestra de una población.



## Supuestos:

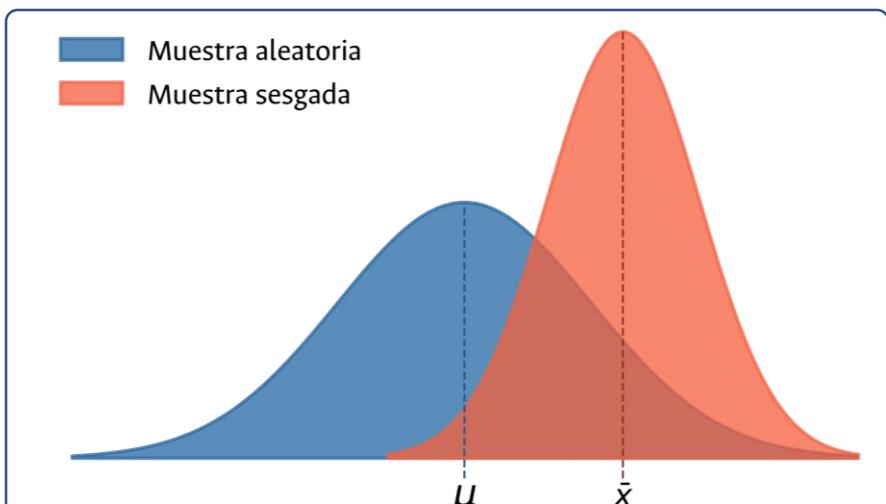
### Distribución idéntica de muestreo

- Al tratar con datos que provienen de una población desconocida, es necesario asumir (o validar) que **en el proceso de muestreo todos los datos provienen de la misma distribución de probabilidad**.
- Se suele asumir que esta distribución es **normal** para aprovechar sus propiedades. La verificación de normalidad puede hacerse visualmente usando histogramas o gráficos p-p o q-q, o mediante una prueba como la de Shapiro-Wilk.



### Muestras aleatorias e independientes

- Para mitigar el riesgo o el sesgo en el proceso de muestreo, se utilizan **técnicas para la selección de muestras aleatorias**.
- Además, es necesario que las muestras sean independientes entre sí, ya que una posible influencia o dependencia entre las mediciones realizadas puede generar un sesgo significativo en los resultados obtenidos.
- El **proceso de muestreo** es de suma importancia en el análisis de datos y es necesario validar que se realizó con las precauciones necesarias.



Estadística Inferencial



Objetivo

## ¿Qué se busca con la estadística inferencial?

Estimar parámetros con confianza



Probar teorías sobre parámetros



## Estadística Inferencial



## Estimar parámetros con confianza

De todos los dueños de perros en la ciudad de Bogotá, ¿qué **proporción** ha utilizado alguna vez un hotel para su mascota?



**Objetivo:**  
Estimar una **proporción** poblacional con confianza

¿Cuál es el tiempo **promedio** de desplazamiento que los empleados gastan para llegar a su trabajo en la ciudad de Lima?



**Objetivo:**  
Estimar una **media** poblacional con confianza

¿Cuál es la **diferencia promedio** en el nivel educativo entre el hijo mayor y el hijo menor en las familias de la ciudad de Quito?

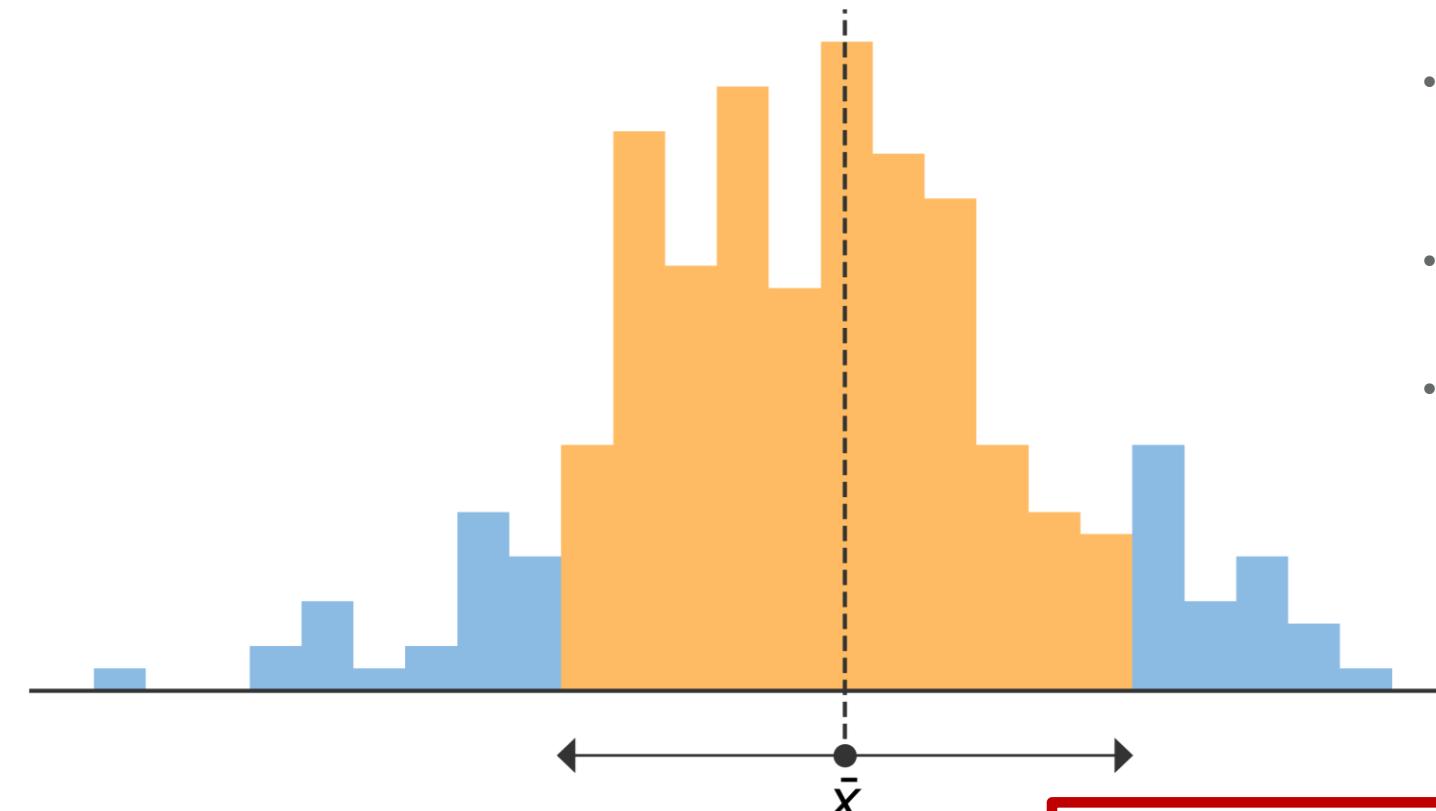


**Objetivo:**  
Estimar una **diferencia de medias** de población con confianza



## Intervalos de confianza

### Estimar parámetros con confianza



- Al realizar estimaciones, solo se puede obtener el que se considera el **mejor estimado** que se alcanza a partir de una muestra.
- El mejor estimado posible proviene del cálculo de valores con la información de una muestra reducida.
- Generalmente, este valor difiere del valor real, por lo que no se puede afirmar que representa a toda la población sin antes intentar identificar qué tanto se puede alejar del valor real.

¿Qué tanto se acerca el valor estimado al valor real?

Estimación del parámetro  $\pm$  margen de error

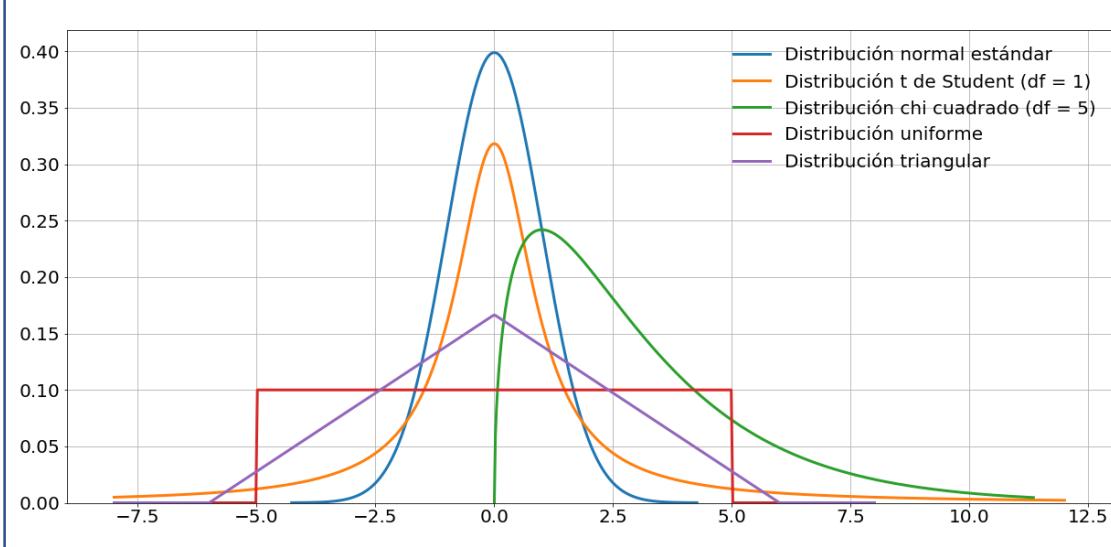


## Margen de error

“unos cuantos” \* errores estándar de distancia

También depende del **nivel de confianza** definido.

Depende del tipo de la distribución de la muestra:



**Error estándar:** desviación estándar de la distribución de muestreo. Por ejemplo, el error estándar de la media:

$$SE = \frac{s}{\sqrt{n}}$$

Desviación estándar  
de la muestra

Tamaño de la muestra



## Nivel de confianza y significancia

Para calcular el **margin de error** es necesario definir el **nivel de confianza** que se desea tener en la estimación.

### $\alpha$ = Nivel de significancia

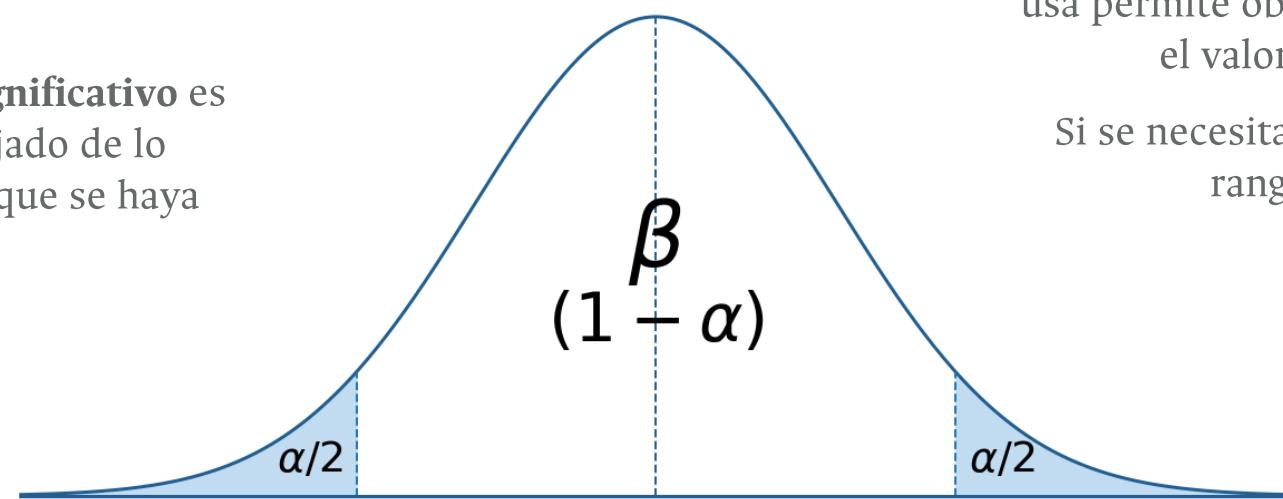
La significancia es una **probabilidad baja** en la que se falla al obtener un intervalo que contenga el valor real.

En estadística, un resultado **significativo** es algo sorprendente que está alejado de lo esperado y que es improbable que se haya obtenido por azar.

### $\beta$ = Nivel de confianza

El nivel de **confianza** es una probabilidad relativamente alta en la que la muestra que se usa permite obtener un intervalo que contenga el valor real del parámetro poblacional.

Si se necesita un nivel de confianza mayor, el rango del intervalo será más amplio.



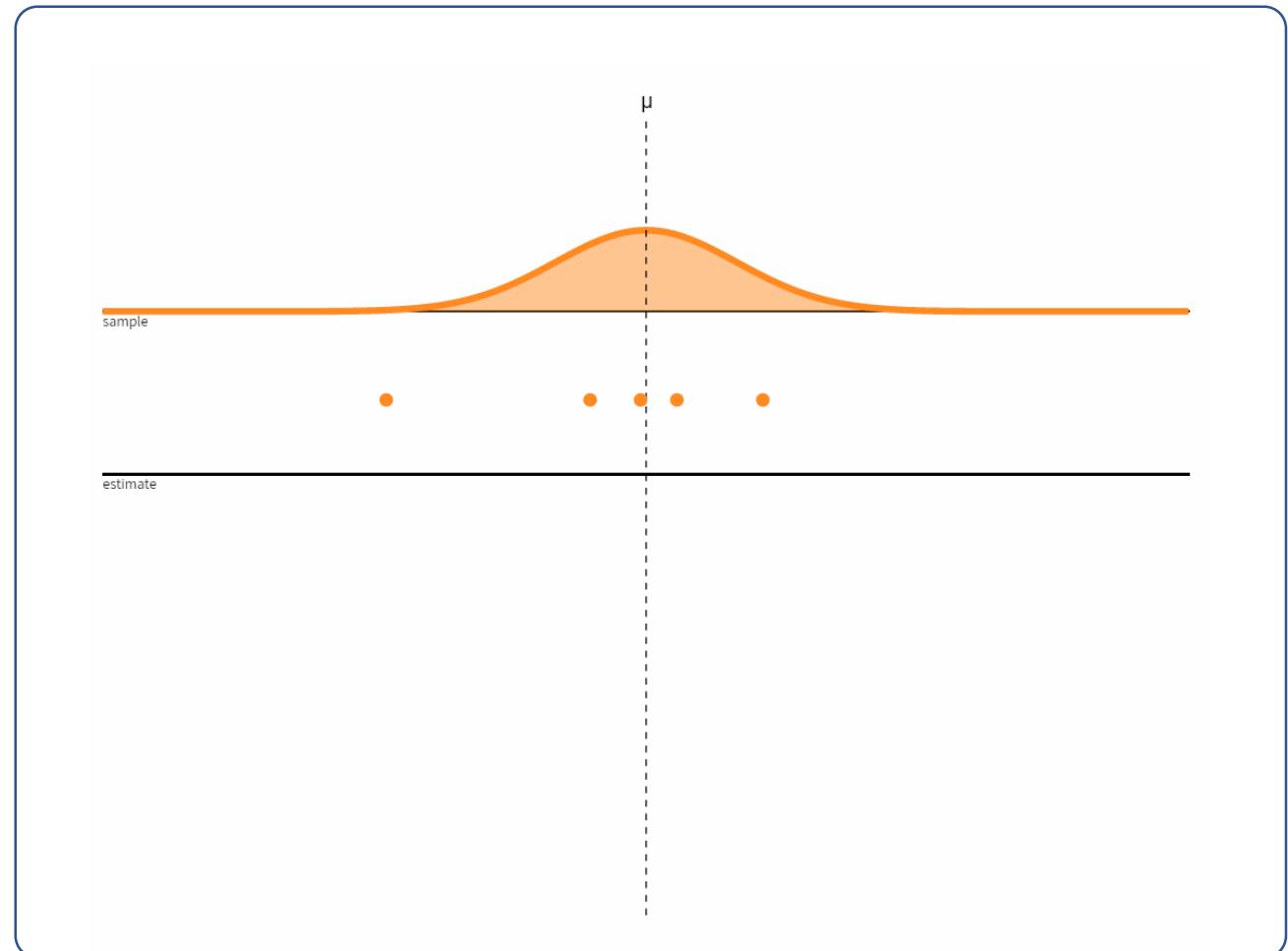


## Interpretación

La creación de un intervalo de confianza con un **nivel de confianza** específico, por ejemplo del 95%, implica que, si se repite el muestreo muchas veces (en las mismas condiciones), ese porcentaje de las muestras obtenidas incluirán dentro de su intervalo el parámetro real.

Conforme crece el **nivel de confianza**, se amplía también el rango del intervalo para incluir más valores.

Así mismo, a medida que aumenta el **tamaño de la muestra**, aumenta la precisión del estimado, puesto que incluye más información y disminuye el tamaño de los intervalos.



## Estadística Inferencial



## Intervalos de confianza



Estadística Inferencial



Objetivo

## ¿Qué se busca con la estadística inferencial?

Estimar parámetros con confianza



Probar teorías sobre parámetros



## Estadística Inferencial



## Probar teorías sobre parámetros

¿Ha **aumentado** la **proporción** de padres colombianos que creen que la falta de sueño de su hijo adolescente se debe al uso de dispositivos electrónicos?



**Objetivo:**  
Probar una teoría sobre una **proporción** de la población

¿Los estudiantes de último semestre de la UNAL tienen un **mayor** Índice de Masa Corporal (IMC) **medio en comparación** con los recién admitidos?



**Objetivo:**  
Probar una teoría que compara **dos medias** poblacionales

¿Ha **aumentado** la **temperatura media** de la superficie terrestre de una de las regiones del Amazonas en los últimos 20 años?



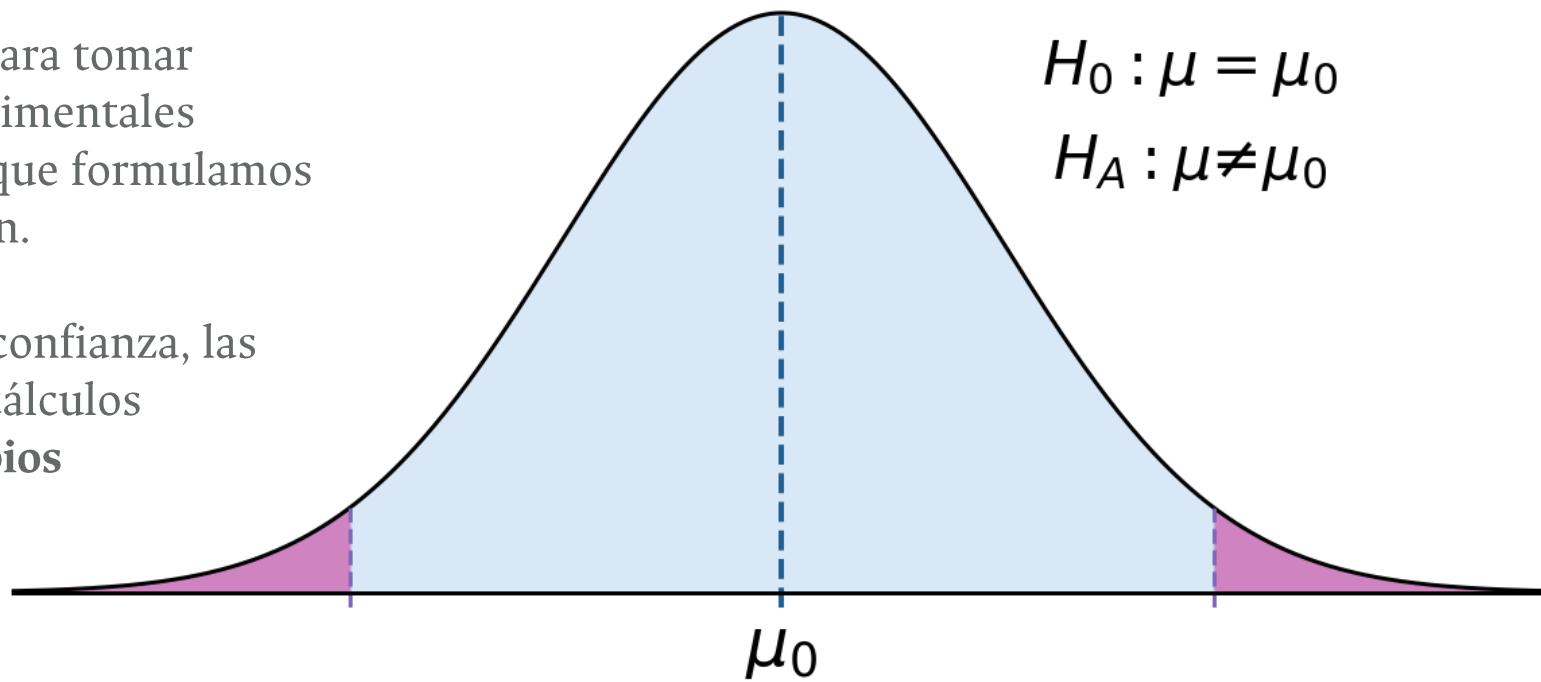
**Objetivo:**  
Probar una teoría sobre la **diferencia de medias** de población con confianza



## Test o prueba de hipótesis

Probar teorías sobre parámetros

- Método estadístico que puede ser usado para tomar decisiones sobre conjuntos de datos experimentales basándose en una **suposición o hipótesis** que formulamos sobre los parámetros reales de la población.
- A diferencia del enfoque de intervalos de confianza, las pruebas de hipótesis no pretenden hacer cálculos estimados, sino **validar resultados o cambios significativos** en los datos.





## Hipótesis nula y alternativa

- Es necesario formular **dos hipótesis iniciales mutuamente excluyentes** sobre la población.
- Se busca determinar qué hipótesis está mejor soportada por los datos.

### Hipótesis Nula

 $H_0$ 

- La **hipótesis nula** implica que no existe una diferencia significativa en el parámetro que se decide evaluar y que la observación no es suficientemente inusual.
- **No se puede probar ni desmentir**, solo se puede hallar (o no) evidencia que sugiera su rechazo en favor de la **hipótesis alternativa**.

### Hipótesis Alternativa

 $H_A$ 

- La **hipótesis alternativa** es el efecto que se está intentando verificar.
- Su aprobación se produce cuando hay evidencia significativa que sugiere el rechazo de la **hipótesis nula**.
- **No se puede garantizar su veracidad**, pero se puede minimizar el error con niveles de confianza definidos.

## Prueba de dos colas

## Prueba de cola inferior

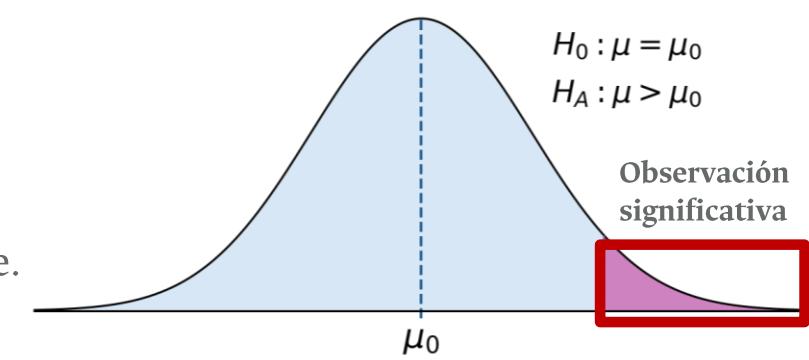
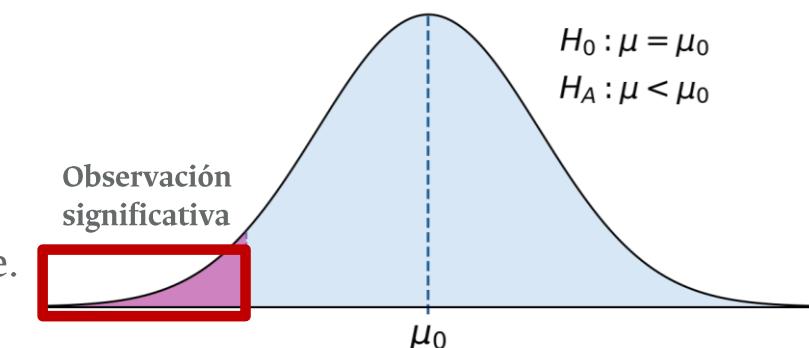
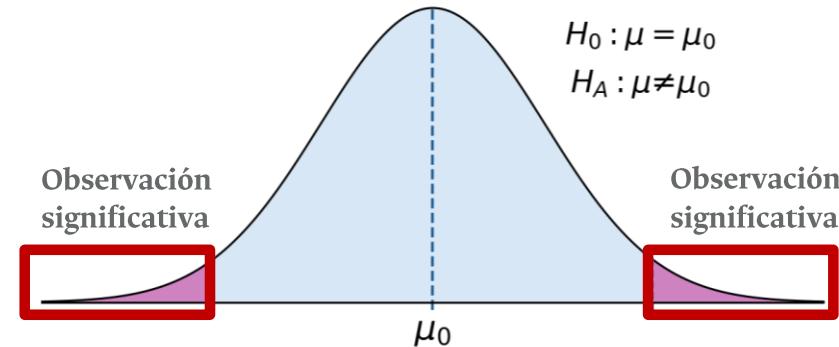
## Prueba de cola superior

### Estadística Inferencial



### Test de hipótesis

- Se emplea cuando se quiere verificar si un parámetro **ha tenido algún cambio** significativo en cualquier dirección, es decir, si es diferente.
- La **hipótesis nula** representa el escenario en el que el valor permanece igual a lo observado inicialmente.



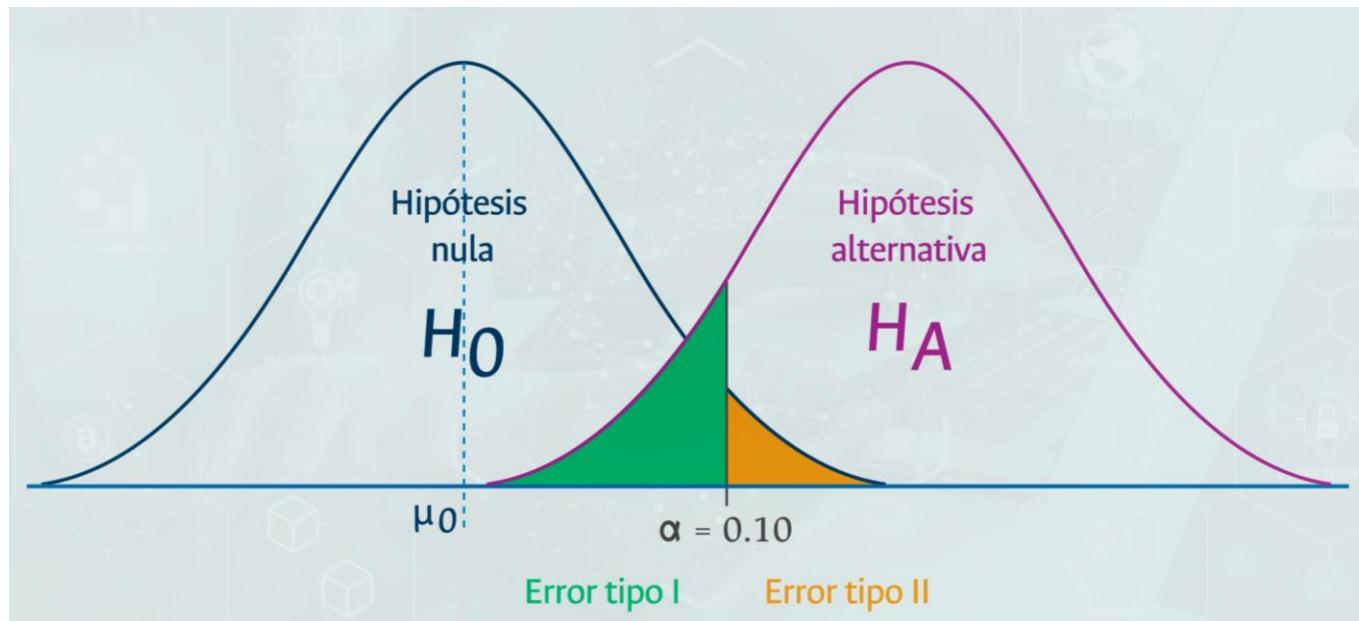
- Cuando se quiere verificar si un parámetro ha **disminuido** significativamente.
- La **hipótesis nula** representa el escenario en el que el valor permanece igual o incluso mayor a lo observado inicialmente.

- Cuando queremos verificar si un parámetro ha **aumentado** significativamente.
- La **hipótesis nula** representa el escenario en el que el valor permanece igual o incluso menor a lo observado inicialmente.



## Tipos de error estadístico

A pesar de tener un nivel de confianza alto, es posible cometer errores al aceptar o rechazar una hipótesis. Por lo tanto, es necesario considerar los posibles errores que se pueden cometer e incluso orientar el análisis para minimizar su impacto.

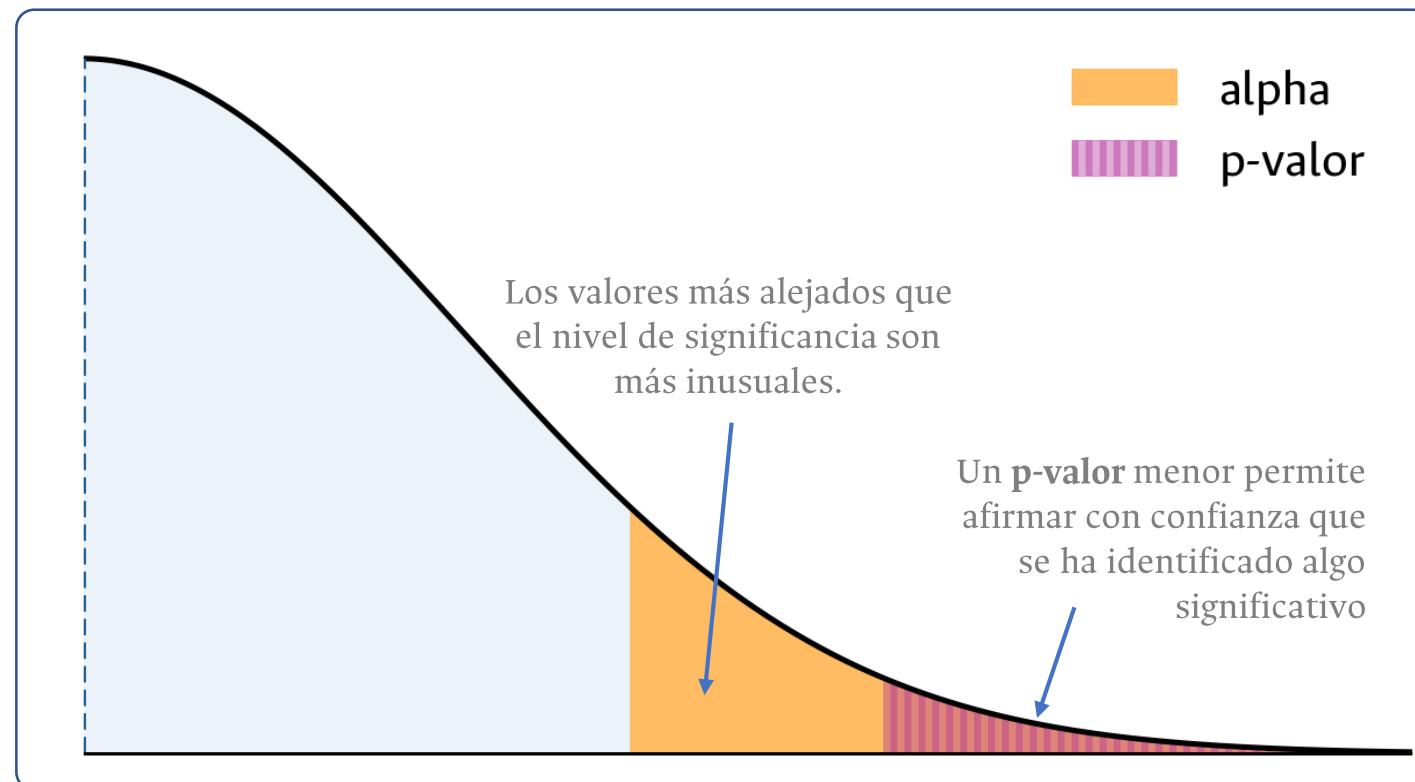


	Evidencia real que apoya $H_0$	Evidencia real que apoya $H_A$
No rechazar $H_0$	Decisión correcta (No rechazar $H_0$ )	Error Tipo II Falso Negativo
Rechazar $H_0$	Error Tipo I Falso Positivo	Decisión correcta (Rechazar $H_0$ )



## p-valor

- Al realizar estimaciones, se obtiene el **p-valor** de la observación (o estimación).
- Es la **probabilidad** de encontrar los resultados observados si la hipótesis nula fuera cierta.
- Se usa el p-valor **para rechazar o no** la hipótesis nula y optar por la hipótesis alternativa.
- Debido al error de muestreo, no se puede garantizar que el resultado obtenido sea correcto, ya que se desconocen los detalles de la población.
- Sin embargo, si se asume que los datos siguen una distribución, se puede determinar la probabilidad de que un hallazgo se produzca por causa del azar.
- Cuando el p-valor de una observación es menor que el **nivel de significancia ( $\alpha$ )**, se puede optar por **rechazar la hipótesis nula con confianza**.



## Estadística Inferencial



## Test de hipótesis





## Tabla de contenidos

1

Estadística Inferencial



Intervalos de confianza



Test de hipótesis

2

Modelado estadístico

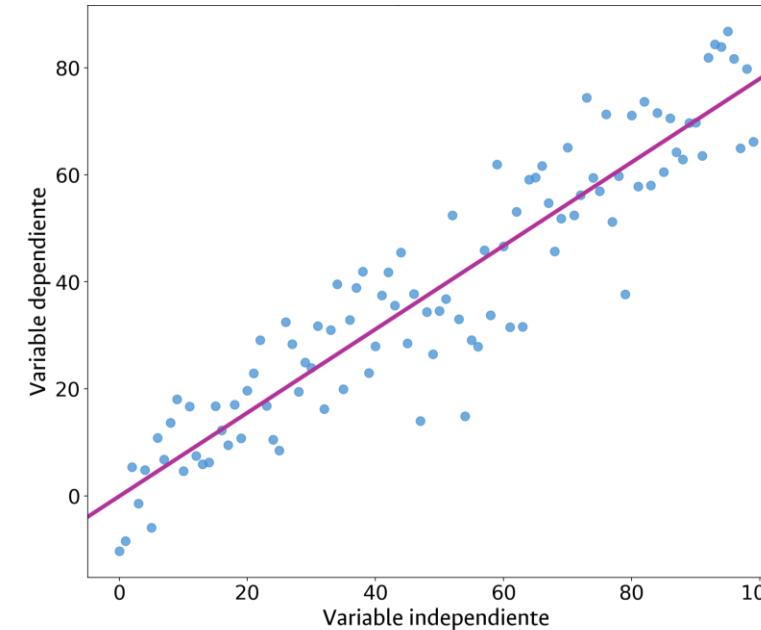
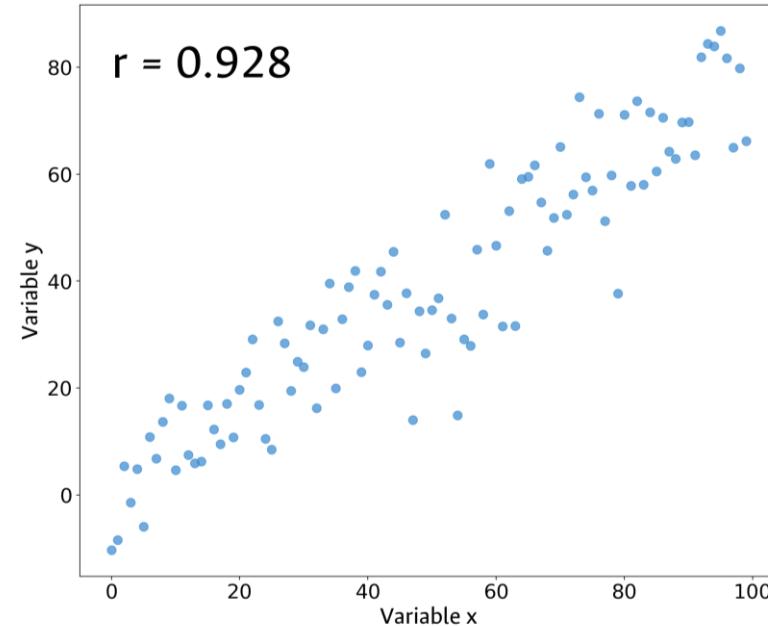


Análisis de correlaciones



Análisis de regresiones

## Modelado estadístico



## Análisis de correlaciones

Busca **identificar relaciones o dependencias** entre dos o más variables. Estas relaciones se cuantifican con **coeficientes de correlación**, que representan numéricamente el grado de relación entre las variables analizadas.



## Análisis de regresiones

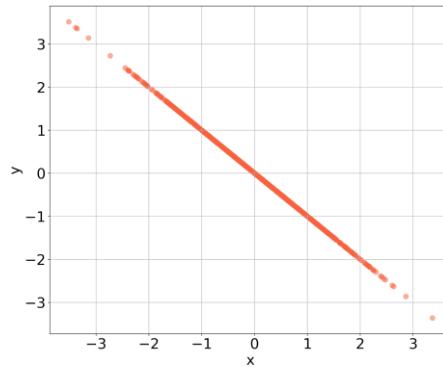
Busca **modelar relaciones y asociaciones** específicas entre variables. A diferencia del análisis de correlación, con la regresión se definen **funciones matemáticas** que permiten predecir el valor de una variable dependiente a partir de una (o varias) variable(s) independiente(s).

## Análisis de correlación

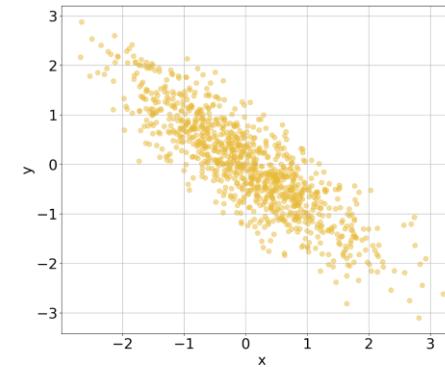


## Coeficientes de correlación

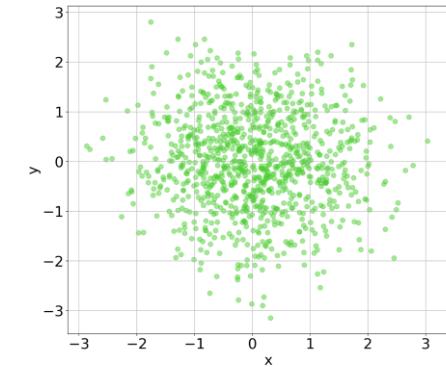
- Los **coeficientes de correlación** permiten cuantificar la relación entre variables, apoyados por la estadística inferencial para establecer los niveles de significancia de estas relaciones.
- Los coeficientes de correlación pueden tener valores entre **-1.0 y 1.0**.
- Una **correlación positiva** significa que las variables tienen una **relación directa**, mientras que una **correlación negativa** indica que las variables tienen una **relación inversa** entre sí.
- Mientras más cerca está de los extremos -1.0 y 1.0, mayor es la fuerza de la correlación. Los valores cercanos a 0 manifiestan una correlación lineal nula entre las variables.



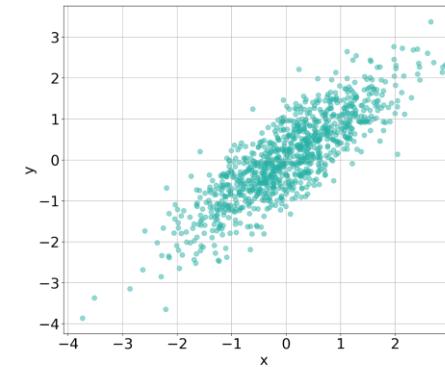
$$r = -1.0$$



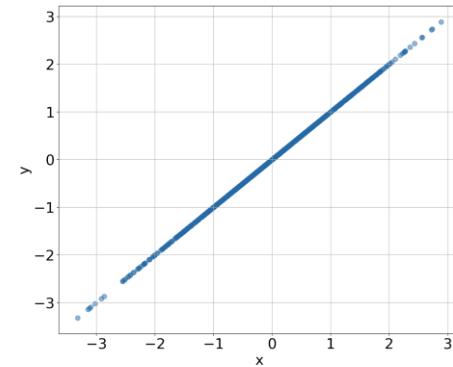
$$r = -0.8$$



$$r = 0.0$$



$$r = 0.8$$

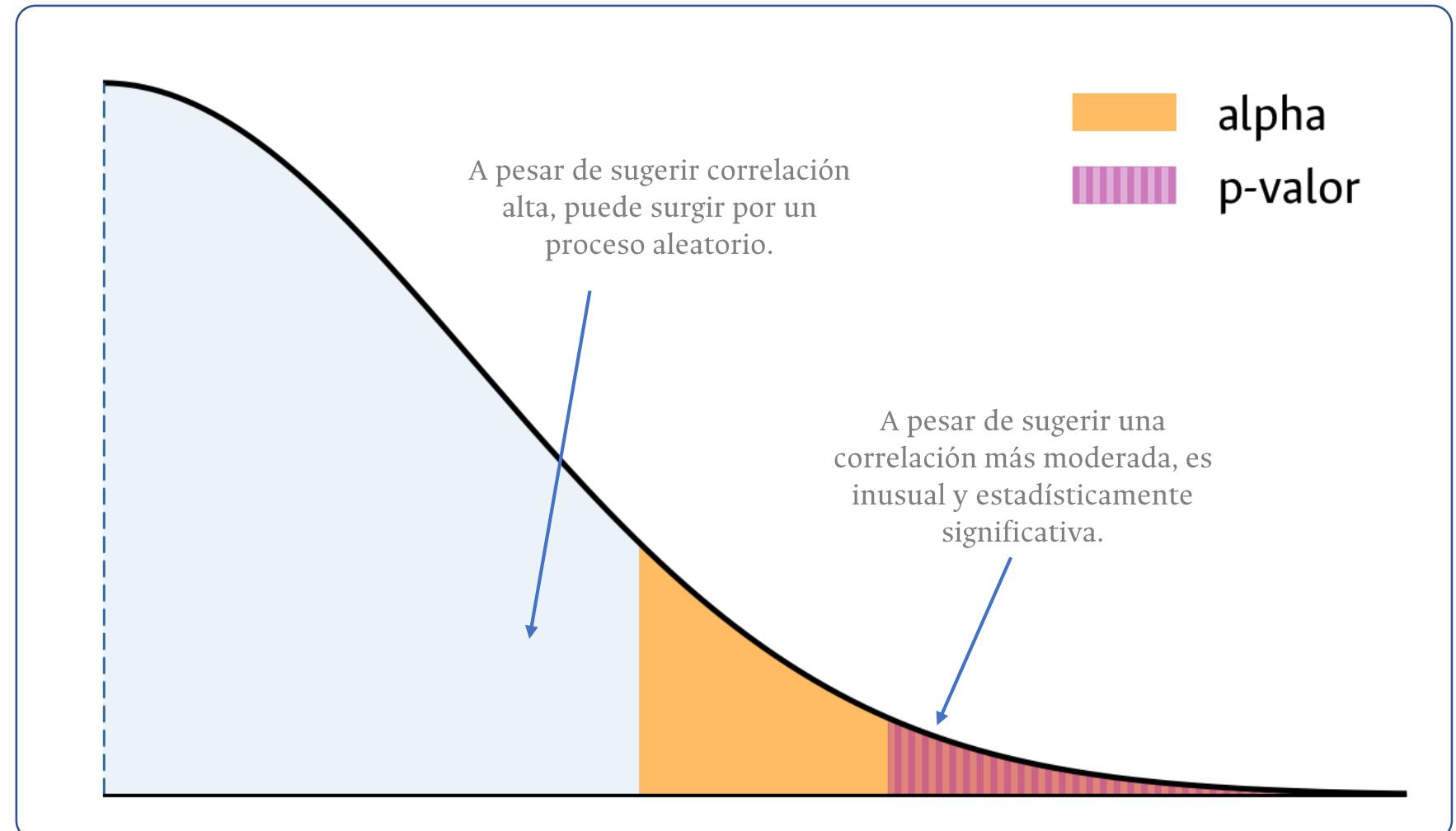


$$r = 1.0$$

## Análisis de correlación

 Coeficientes de correlación

- Al igual que con otras estimaciones, el cálculo de correlaciones puede apoyarse de un **nivel de significancia**.
- El **p-valor** de la correlación es muy importante para evitar señalar relaciones que se puedan haber producido por la aleatoriedad de los datos.

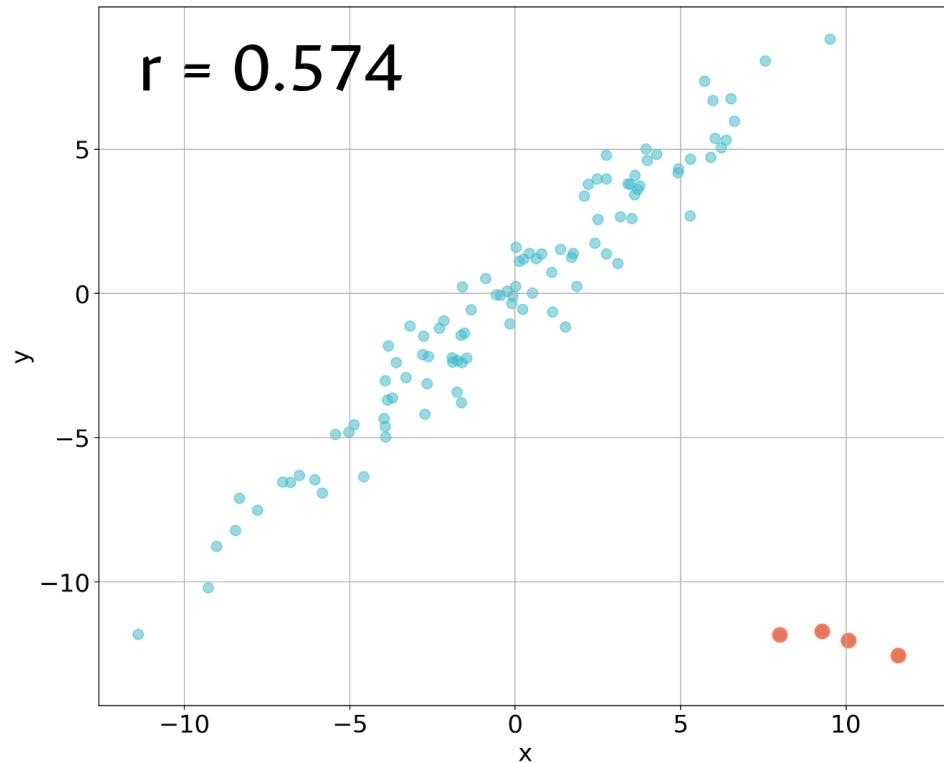
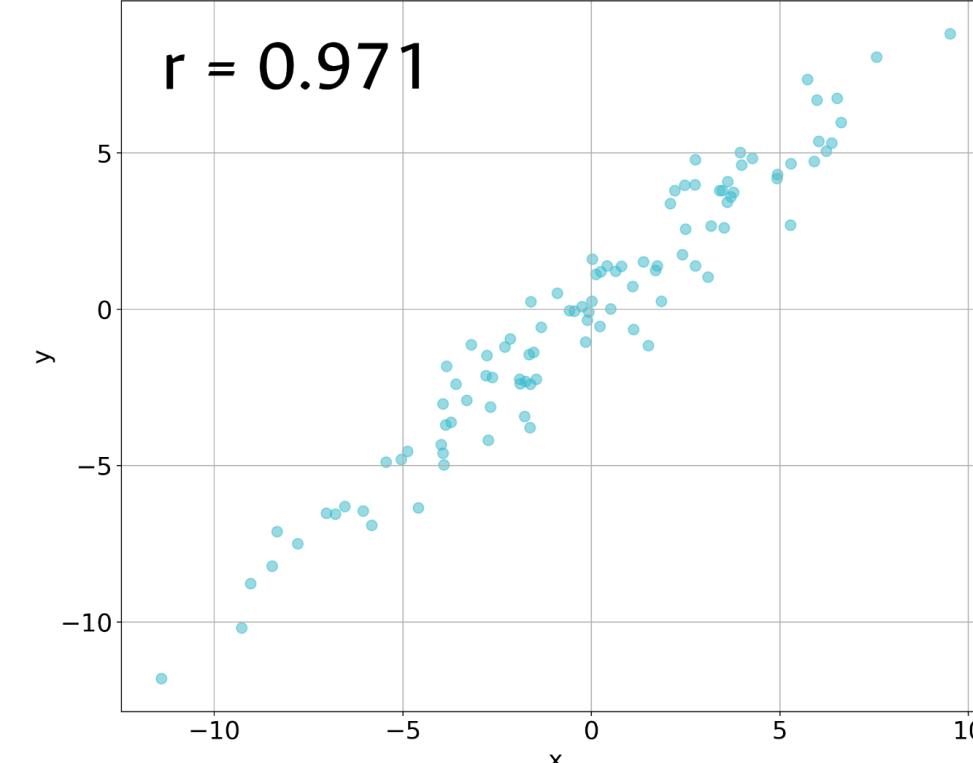


## Análisis de correlación



## Coeficientes de correlación

- El coeficiente de correlación es muy sensible a valores de datos atípicos o outliers.
- Realice una correcta limpieza de los datos antes de hacer este tipo de análisis.

**Con outliers****Sin outliers**

## Análisis de correlación



## Coeficientes de correlación

Un coeficiente de correlación bajo no implica que no exista relación entre las variables, ya que estas pueden tener una relación no lineal que no sea detectada con los coeficientes de correlación típicos. Siempre verifique por medio de gráficos de dispersión.

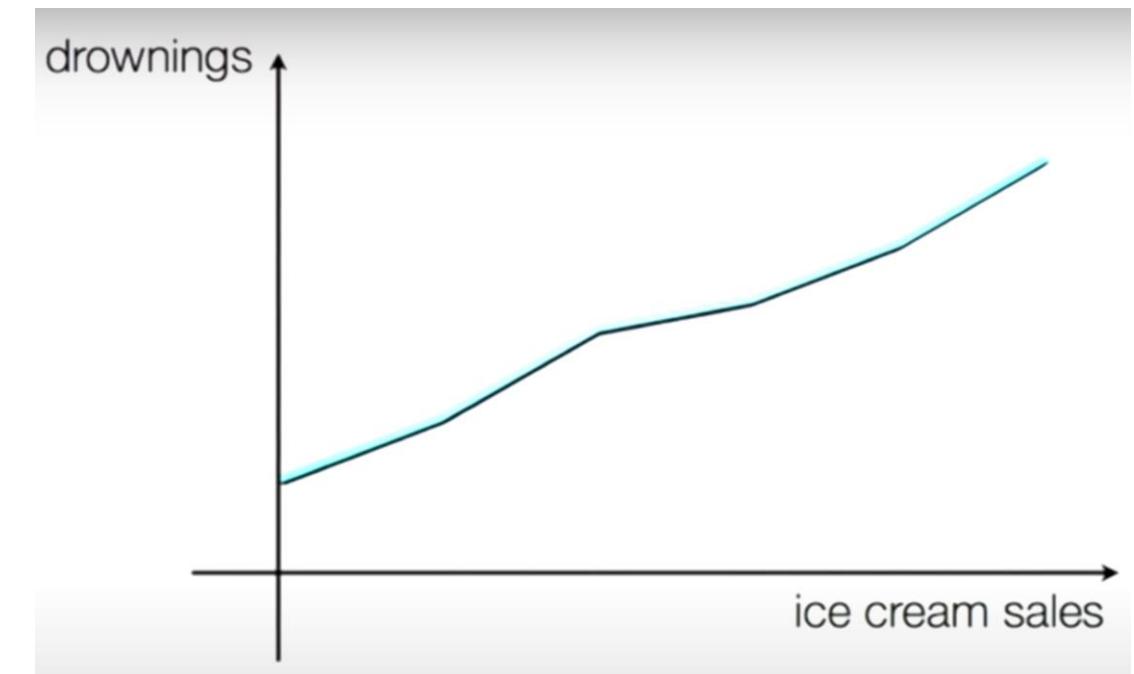
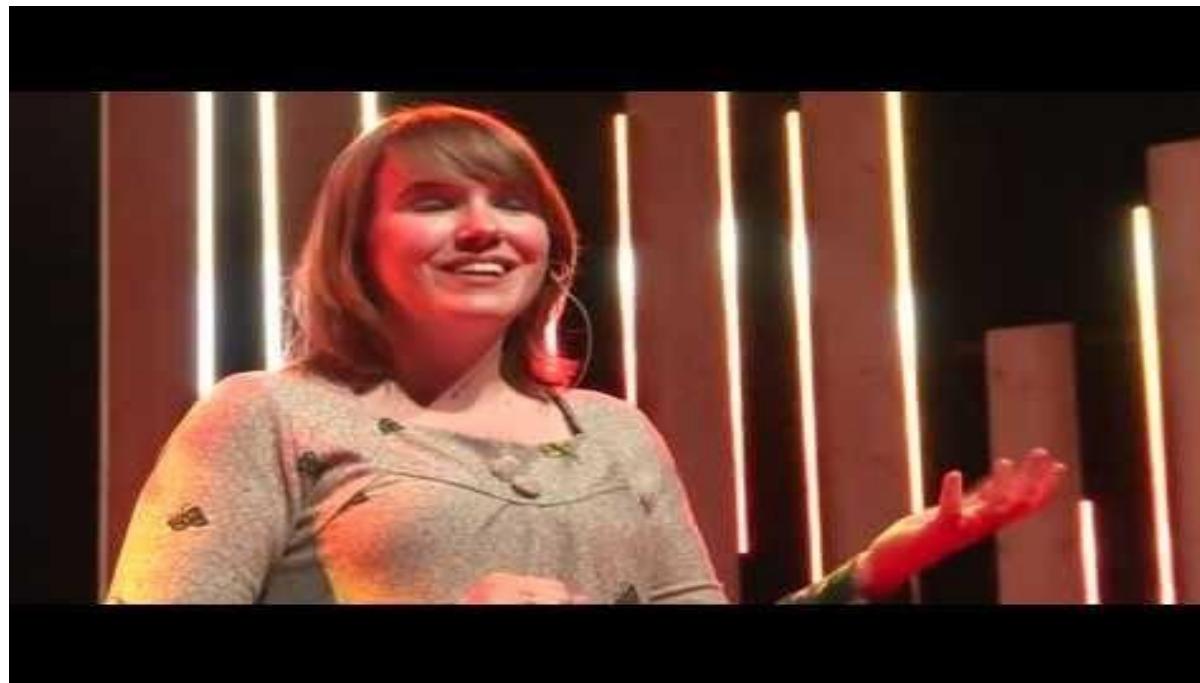


## Análisis de correlación



## Coeficientes de correlación

**Correlación no implica causalidad:** a partir de una correlación significativa **no** se puede concluir que los cambios en una variable **causan** cambios en otra. Sólo los experimentos controlados adecuadamente permiten determinar si una relación es causal.



Fuente: TEDx. The danger of mixing up causality and correlation: Ionica Smeets. Online:  
<https://tedx.tumblr.com/post/35843089997/the-danger-of-mixing-up-causality-and-correlation>  
Nov, 2012

## Análisis de regresión

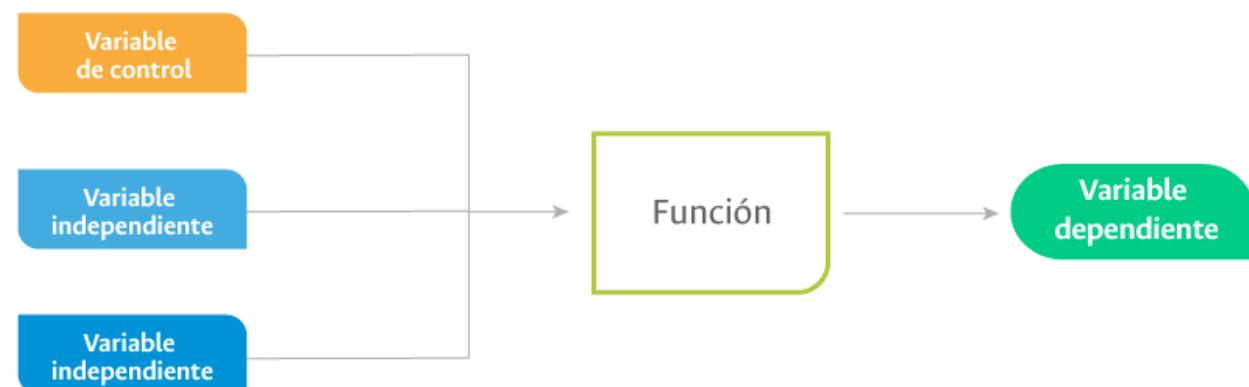
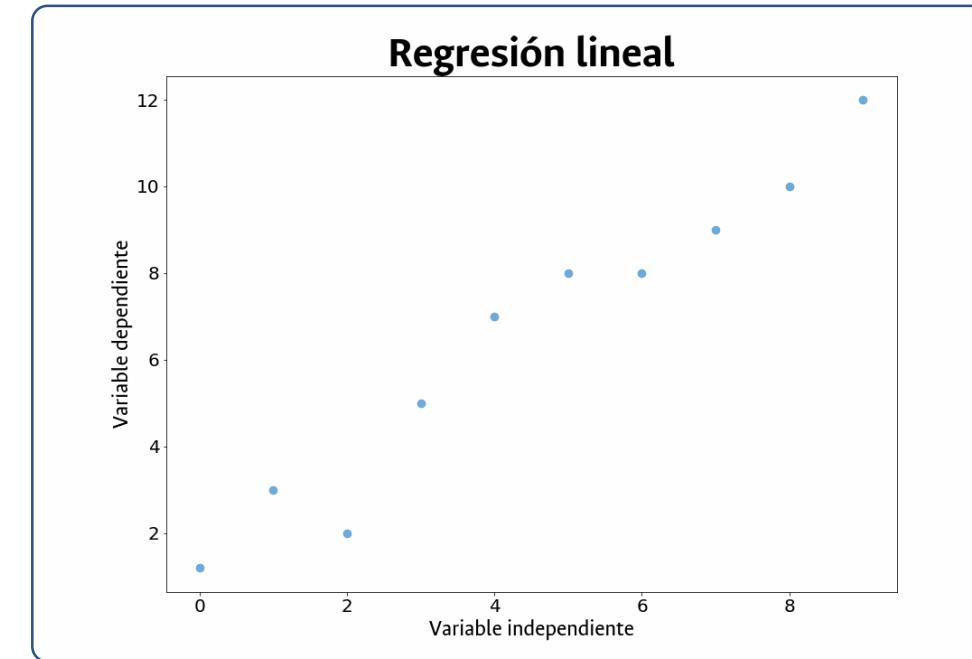


## Objetivo de la regresión

La función principal de la regresión es **examinar y predecir el valor de una variable dependiente** a partir de una o más **variables independientes**.

Permite establecer y **representar relaciones y asociaciones entre variables** por medio de un modelo matemático.

Con la regresión es posible crear **funciones matemáticas** que se ajusten a los datos.



## Análisis de regresión

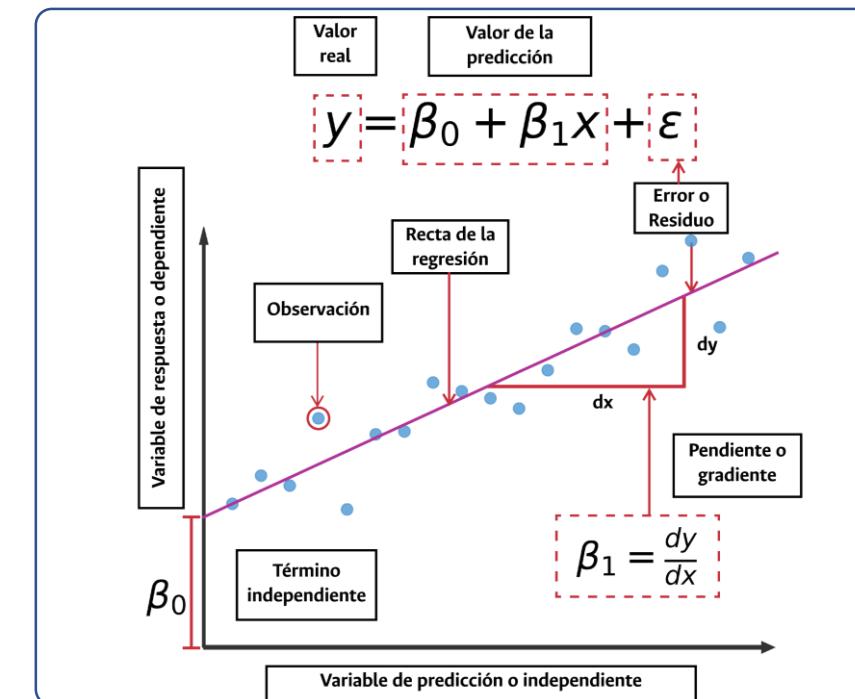
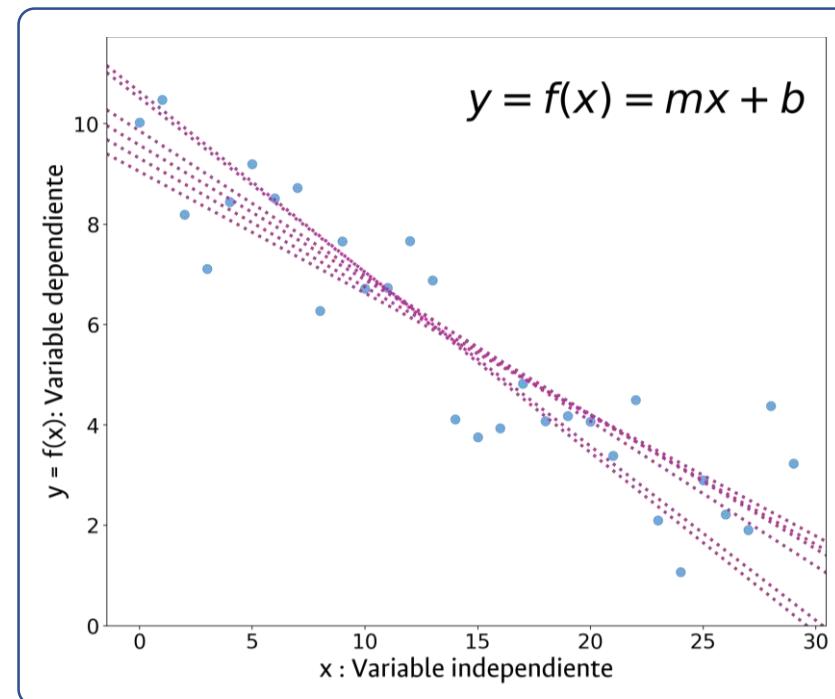


## Regresión lineal

La **regresión lineal** es un modelo estadístico que examina la relación **lineal** entre **dos variables** y plantea una relación lineal proporcional, en la que si una variable aumenta o disminuye, la otra también varía.

El objetivo es encontrar la **función matemática** que más se ajuste a los datos y minimice el error entre las observaciones iniciales y las predicciones.

Uno de los métodos más utilizados en la regresión lineal es el **método de mínimos cuadrados ordinarios** (OLS - Ordinary Least Squares).



## Análisis de regresión



## Regresión multilineal

La **regresión lineal múltiple** es un modelo estadístico que examina la relación **lineal** entre **más de dos variables**, donde una es la variable dependiente o de respuesta, y las otras son variables independientes de predicción continuas o categóricas.

La función del modelo de regresión lineal múltiple es muy similar a la del modelo de regresión lineal de dos variables, con un coeficiente y una variable en la fórmula por cada variable predictoría adicional.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_i x_i + \dots + u$$

Diagrama que ilustra la ecuación de regresión multilineal:

- Variables predictoras  $x_i$** : Se representan por flechas que apuntan hacia la variable  $x_i$  en la ecuación.
- Coeficientes de regresión  $\beta_i$** : Se representan por flechas que apuntan hacia los términos  $\beta_i x_i$  en la ecuación.
- Variable de respuesta**: Se indica con una flecha apuntando hacia la variable  $Y$ .
- Error**: Se indica con una flecha apuntando hacia la variable  $u$ .

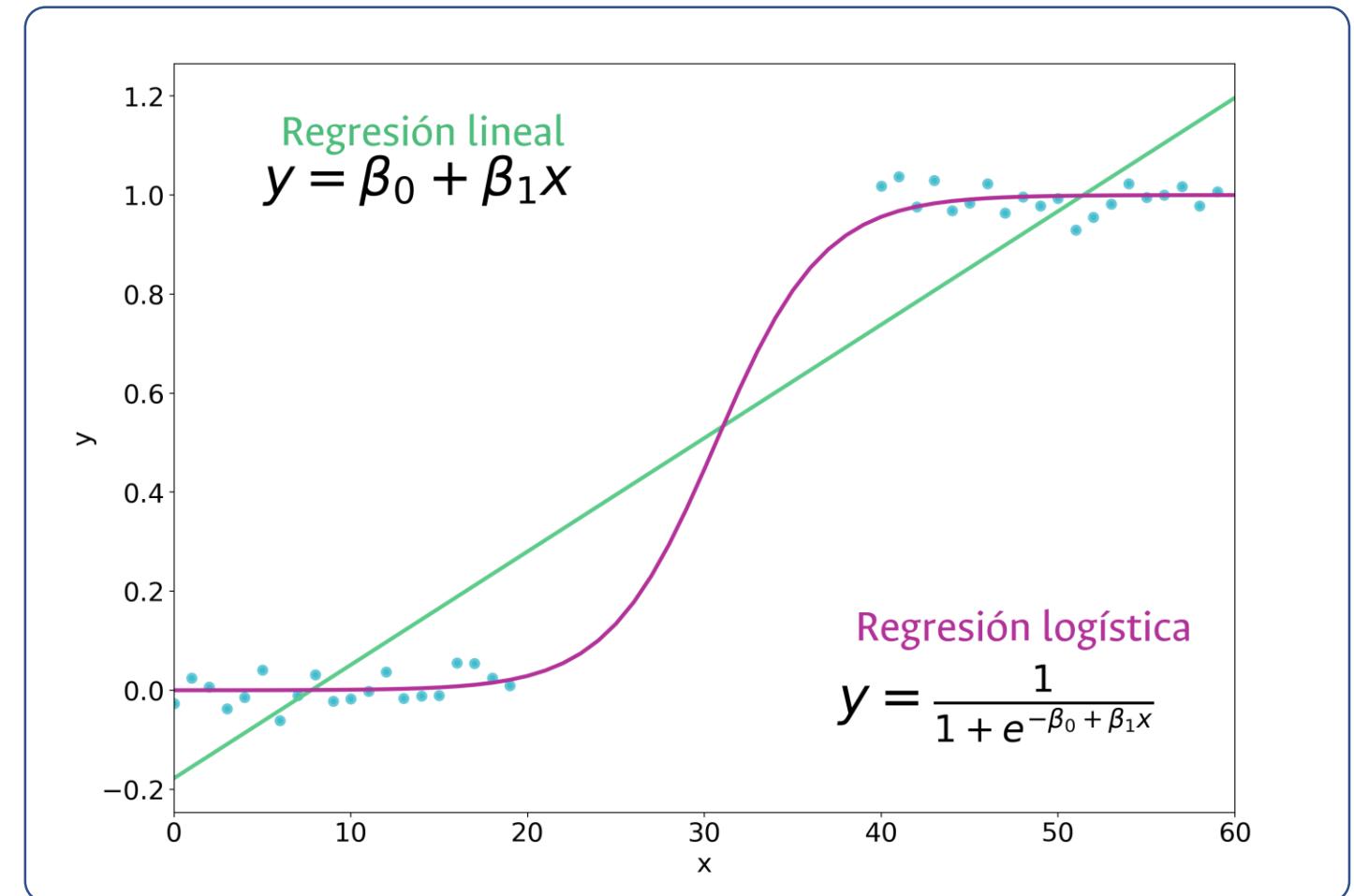
## Análisis de regresión



## Regresión logística

La **regresión logística** es un modelo estadístico que examina la relación entre **dos variables**, que simula un modelo logístico binario.

Se emplea para modelar la **probabilidad de ocurrencia de un evento** en función de variables independientes.



## Análisis de regresión



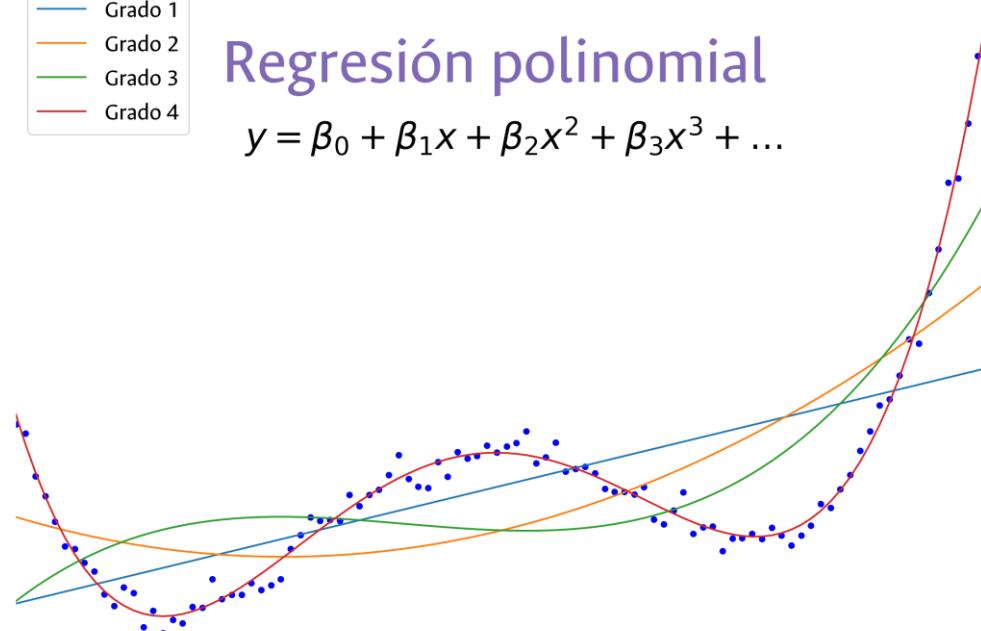
## Regresión no lineal

Existen otros tipos de modelos de regresión que corresponden a distintas estructuras de modelado matemático.

- Grado 1
- Grado 2
- Grado 3
- Grado 4

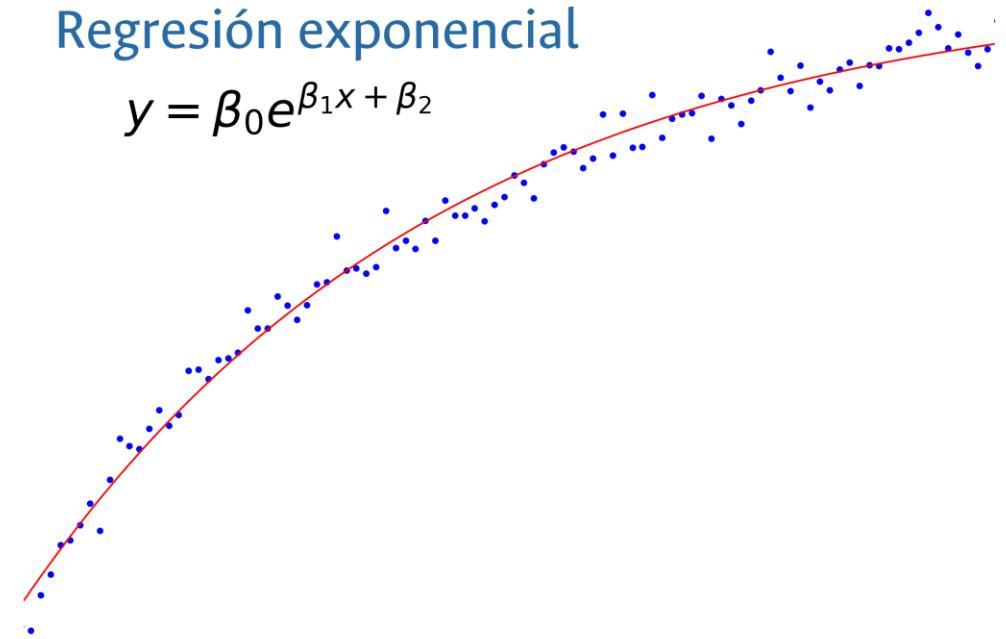
## Regresión polinomial

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \dots$$



## Regresión exponencial

$$y = \beta_0 e^{\beta_1 x + \beta_2}$$



3

## Actividades Unidad 3



## Logística

**Actividades Unidad 3**

1

**Objeto Virtual de Aprendizaje (OVA)**

- Análisis avanzado de datos con Python



4

**Documentos (pdf)**

- Selección test estadístico
- Guía de referencia Python



2

**Estadística inferencial:**

- Taller guiado (notebook)
- Quiz 5 (notebook)



5

**Tarea 3 – Análisis avanzado de datos con Python**

- Notebook



3

**Modelado estadístico:**

- Taller guiado (notebook)
- Quiz 6 (notebook)



6

**Campuswire:**

- Dudas en los foros
- Colaboración entre estudiantes



## Derechos de imágenes

Estadística inferencial <https://pixabay.com/es/photos/an%C3%A1lisis-negocio-mano-gr%C3%A1ficos-3782319/>

Estadística descriptiva <https://www.freepik.com/free-photos-vectors/business>“ Business photo created by janno028

Estadística inferencial <https://www.freepik.com/free-photos-vectors/school> School photo created by rawpixel.com

Estimar parámetros con confianza <https://www.freepik.com/free-photos-vectors/business> Business photo created by jcomp

Probar teorías sobre parámetros <https://www.freepik.com/free-photos-vectors/business> Business photo created by pressfoto

Estimar parámetros con confianza [https://www.freepik.com/free-photo/front-view-dogs-costumes\\_6877142.htm](https://www.freepik.com/free-photo/front-view-dogs-costumes_6877142.htm)

Estimar parámetros con confianza [https://www.freepik.com/free-photo/businessman-holding-disposable-coffee-cup-using-mobile-phone\\_8402321.htm](https://www.freepik.com/free-photo/businessman-holding-disposable-coffee-cup-using-mobile-phone_8402321.htm)

Estimar parámetros con confianza [https://www.freepik.com/free-photo/family-spending-time-together\\_856019.htm](https://www.freepik.com/free-photo/family-spending-time-together_856019.htm)

Probar teorías sobre parámetros [https://www.freepik.com/free-photo/young-student-sitting-table-sleeping\\_5913141.htm](https://www.freepik.com/free-photo/young-student-sitting-table-sleeping_5913141.htm)

Probar teorías sobre parámetros [https://www.freepik.com/free-photo/workaholic-office\\_5401107.htm](https://www.freepik.com/free-photo/workaholic-office_5401107.htm)

Probar teorías sobre parámetros [https://www.freepik.com/free-photo/close-shot-broken-camel-thorn-tree-desert-with-sand-dunes-clear-sky\\_7747898.htm](https://www.freepik.com/free-photo/close-shot-broken-camel-thorn-tree-desert-with-sand-dunes-clear-sky_7747898.htm)

Fundamentos <https://seeing-theory.brown.edu/probability-distributions/index.html#section3>

Intervalo de confianza <https://seeing-theory.brown.edu/frequentist-inference/index.html>

Intervalo de confianza <https://www.freepik.com/photos/background> Background photo created by jcomp

Test de hipótesis [https://www.freepik.com/free-photo/student-using-augmented-reality-studying\\_3954533.htm](https://www.freepik.com/free-photo/student-using-augmented-reality-studying_3954533.htm)

Análisis de correlación [https://www.freepik.com/free-photo/magnifying-glass-red-stock-market-graph-paper\\_3095565.htm](https://www.freepik.com/free-photo/magnifying-glass-red-stock-market-graph-paper_3095565.htm)

Análisis de regresión [https://www.freepik.es/vector-gratis/conexion-lineas-puntos\\_1894934.htm](https://www.freepik.es/vector-gratis/conexion-lineas-puntos_1894934.htm)

Actividades Unidad 3 <https://www.freepik.es/fotos-vectores-gratis/negocios> Designed by katemangostar



## Referencias

- Agrawal, Y. (Enero 21 del 2019). Hypothesis testing in Machine learning using Python [Prueba de hipótesis en aprendizaje automático usando Python]. Towards data science. <https://towardsdatascience.com/hypothesis-testing-in-machine-learning-using-python-a0dc89e169ce>
- Ahmed, S. (Septiembre 19 del 2018). An Intuitive Perspective to Linear Regression [Una perspectiva intuitiva a la regresión lineal]. Hackernoon. <https://hackernoon.com/an-intuitive-perspective-to-linear-regression-7dc566b2c14c>
- Basic Statistical Test Flow Chart. (s.f.). Shippensburg University. <https://webspace.ship.edu/pgmarr/Geo441/Statistical%20Test%20Flow%20Chart.pdf>
- Bronshtein, A. (Mayo 8 del 2017). Simple and Multiple Linear Regression in Python [Regresión lineal simple y multiple en Python]. Towards data science. <https://towardsdatascience.com/simple-and-multiple-linear-regression-in-python-c928425168f9>
- Brownlee, J. (Agosto 15 del 2018). A Gentle Introduction to Statistical Hypothesis Testing [Una breve introducción a las pruebas de hipótesis estadísticas].
- Correlation Coefficients. (13 de enero de 2018). Data Vedas. <https://www.datavedas.com/correlation-coefficients/>
- Dougherty, C. (2020). Introduction to Econometrics (second edition). Oxford University Press, Oxford. <https://home.ubalt.edu/ntsbarsh/Business-stat/StatsticalTables.pdf>
- Forsyth, D. (2018). Probability and Statistics for Computer Science. Muestra y población. Doi: 10.1007/978-3-319-64410-3
- Freeman, M. (2019). Binary Prediction Metrics [Metricas de predicción binaria]. University of Washington. <http://mfviz.com/binary-predictions/>
- Freeman, M. (2017). Hierarchical models [Una introducción al modelo jerárquico]. University of Washington. <http://mfviz.com/hierarchical-models/>
- Gow, M. (Mayo 27 del 2019). Test Statistic Cheat Sheet: Z, T, F, and Chi-Squared [Hoja de trucos de estadística de prueba: Z, T, F y Chi cuadrado]. Medium. <https://medium.com/@marin892/test-statistic-cheat-sheet-z-t-f-and-chi-squared-58f4d9db5b99>



## Referencias

Gomila, J. G. (2015). A Simple but Efective Logistic Regression Derivation [Una derivación de regresión logística simple pero efectiva].  
<http://juangabrielgomila.com/en/logistic-regression-derivation/>

Gunderson, B., West, B.T. & Shedden, K. ( s.f.). Understanding and Visualizing Data with Python. Universidad de Michigan (Coursera).  
<https://www.coursera.org/learn/understanding-visualization-data>

Gunderson, B., West, B.T. & Shedden, K.( s.f.). Inferential Statistical Analysis with Python. Universidad de Michigan (Coursera).  
<https://www.coursera.org/learn/inferential-statistical-analysis-python>

Hypothesis testing. (Enero 12 del 2018). Data Vedas. <https://www.datavedas.com/hypothesis-testing/>

Inferential Statistics in Python. (Enero 26 del 2018). Data Vedas. <https://www.datavedas.com/inferential-statistics-in-python/>

Marshall, E. (s.f.). The Statistics Tutor's Quick Guide to Commonly Used Statistical Tests [La Guía rápida del tutor de estadísticas para las pruebas estadísticas comúnmente utilizadas]. University of Sheffield. <http://www.statstutor.ac.uk/resources/uploaded/tutorsquickguidetostatistics.pdf>

Statistics cheat sheet. Massachusetts Institute of Technology. [https://web.mit.edu/~csvoss/Public/usabo/stats\\_handout.pdf](https://web.mit.edu/~csvoss/Public/usabo/stats_handout.pdf)

Pérez, A. (2019). Trabajando con Python: Machine Learning y la Regresión. Medium. <https://medium.com/qu4nt/el-machine-learning-y-la-regresi%C3%B3n-python-c94db2968a4e>

Rodrigo, J.A. (Junio, 2016). Correlación lineal y Regresión lineal simple. [https://www.cienciadedatos.net/documentos/24\\_correlacion\\_y\\_regresion\\_lineal](https://www.cienciadedatos.net/documentos/24_correlacion_y_regresion_lineal)

Rooney (Diciembre 31 de 2019). Correlation (Pearson, Spearman, and Kendall) [Correlación (Pearson, Spearman, and Kendall)]. Kaggle.  
<https://www.kaggle.com/kiyoung1027/correlation-pearson-spearman-and-kendall/report>



## Referencias

Seabold, Skipper & Perktold, J.(2010). Modelado econométrico y estadístico con Python. Statsmodels [Actas de la 9<sup>a</sup> Conferencia de Python en la Ciencia]. <https://www.statsmodels.org>

Standard Error, Standard Error of Mean and Central Limit Theorem. (6 de enero de 2018). Data Vedas. <https://www.datavedas.com/standard-error-standard-error-of-mean-and-central-limit-theorem/>

Stjernholm M. B. (2016). Statistics for Non-Statisticians. Second Edition.

Virtanen, P., et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods. <https://scipy.org/>

Jones, J. (2020). Stats: Introduction to Estimation [Estadística: introducción a la estimación]. <https://people.richland.edu/james/lecture/m170/ch08-int.html>

Kormanik, K., Laraway, S. & Rogers, R. (s.f.). Statistics. San Jose State University. <https://www.udacity.com/course/statistics--st095>

Lioi, J. (s.f.). Statistical Inference Reference Sheet. University of Arizona.  
<https://www.math.arizona.edu/~jlioi/Statistical%20Inference%20Reference%20Sheet.pdf>

Machine Learning Mastery. <https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/>

McLeod, S. A. (2019, May 30). What does the Z table tell you? [¿Qué te dice la table de puntuaciónZ?]. Simply psychology.  
<https://www.simplypsychology.org/z-table.html>



# Créditos

Facultad de  
**INGENIERÍA**

**Autores**

Felipe Restrepo Calle, PhD

**Asistente docente**

Alberto Nicolai Romero Martínez

**Diseño instruccional**

Claudia Patricia Rodríguez Sánchez

**Diseño gráfico**

Clara Valeria Suárez Caballero  
Milton R. Pachón Pinzón

**Diagramadora PPT**

Daniela Duque García

2022

