

Libro de Estadística - Daniel Christello

Temario o contenido del libro de estadística



Imagen de contenido de estadística

[Previous](#)

[Next](#)

## Descripción del contenido del libro

El contenido del libro se organiza en torno a **estadística descriptiva** y **estadística inferencial**, abordando temas esenciales y métodos aplicados, con una estructura detallada que facilita el aprendizaje paso a paso. Así mismo los ejemplos y sus cálculos se hacen en python y los archivos correspondientes [se comparten en github](#).

A continuación, se proporciona una descripción general de las secciones y capítulos.

### Capítulo 1: Introducción

Lectura altamente recomendable antes de leer este libro completo o a modo de consulta.

- **Prefacio e Introducción**

- Motivaciones y razones para el estudio de estadística.

- Importancia de los métodos estadísticos.

## Capítulo 2: Fundamentos

Esta sección se enfoca en conceptos fundamentales, organización de datos y visualización.

- **Análisis Exploratorio de Datos (EDA)**

- Observación de datos: tipos de estadística y procesos exploratorios.
- Diseño experimental y normas a seguir.

- **Organización y Visualización de Datos**

- Ordenamiento de datos: límites, histogramas y fronteras.
- Visualización: gráficos y diagramas estadísticos.

## Capítulo 3: Estadística Descriptiva

- **Medidas de Tendencia Central y Dispersión**

- Media, mediana y moda.
- Desviación estándar, sesgo, simetría, cuartiles y percentiles.

- **Valores Atípicos (Outliers) y Regla Empírica**

- Identificación de valores atípicos.
- Uso de puntuación Z-score y gráficos de control.

## Capítulo 4: Probabilidad

Explora los fundamentos teóricos necesarios para abordar la inferencia estadística.

- **Conceptos Básicos de Probabilidad**

- Definición y sucesos simples.
- Reglas de probabilidad: suma, multiplicación y teorema de Bayes.

## Capítulo 5: Distribuciones

- **Distribuciones de Probabilidad**

- Distribuciones discretas y continuas.
- Distribuciones importantes: binomial, geométrica, Poisson y otras

- **Modelado y Ajuste de Distribuciones**

- Clasificación y comparación entre distribuciones teóricas y experimentales.

## Capítulo 6: Estadística Inferencial

Esta sección introduce el análisis de inferencias a partir de muestras y la validación de hipótesis.

- **Estimación de Parámetros**

- Intervalos de confianza: media, varianza y proporciones.
- Error de estimación y tamaño de muestra.

## Capítulo 7: Pruebas de Hipótesis

- **Hipótesis**

- Hipótesis nula y alternativa: definición y enfoques.
- Valor crítico, errores tipo I y II, y potencia de prueba.

- **Comparación de Medias y Varianzas**

- Pruebas t de Student.
- Pruebas de hipótesis para varianzas.

## Capítulo 8: Análisis Multivariante

- **Correlación y Regresión Lineal**

- Correlación: coeficientes de Pearson y Spearman.
- Regresión lineal simple y múltiple: ajustes,  $R^2$  ajustado y optimización.

## Capítulo 9: Pruebas No Paramétricas

Dirigida a métodos alternativos para datos que no cumplen supuestos de normalidad.

- **Pruebas de Bondad de Ajuste y Chi-cuadrado**

- Análisis de independencia y pruebas de ajuste.

- **Pruebas No Paramétricas**

- Prueba de la mediana, prueba de Wilcoxon y prueba Kruskal-Wallis.

- **Análisis ANOVA**

- Diseño de factores: ANOVA de un factor y multifactorial.
  - Comparación de medias entre grupos.

## Capítulo 10: Otros Métodos

- **Supervivencia**

- Análisis de supervivencia y pruebas de correlación no lineal.

## Capítulo 11: Introducción a la Estadística Bayesiana

Una introducción breve a conceptos bayesianos y aplicaciones prácticas.

- Interpretación probabilística bayesiana.
- Ejemplo de uso práctico con problemas de inferencia.

### Resumen de características clave:

- **Explicaciones teóricas y prácticas** con ejemplos aplicados.
- **Distribuciones de probabilidad fundamentales** y sus aplicaciones.
- **Metodologías de pruebas de hipótesis** detalladas con discusión de errores y potencia.
- **Pruebas no paramétricas** y análisis ANOVA para datos complejos.
- **Estadística descriptiva** como base para comprender la inferencia estadística.

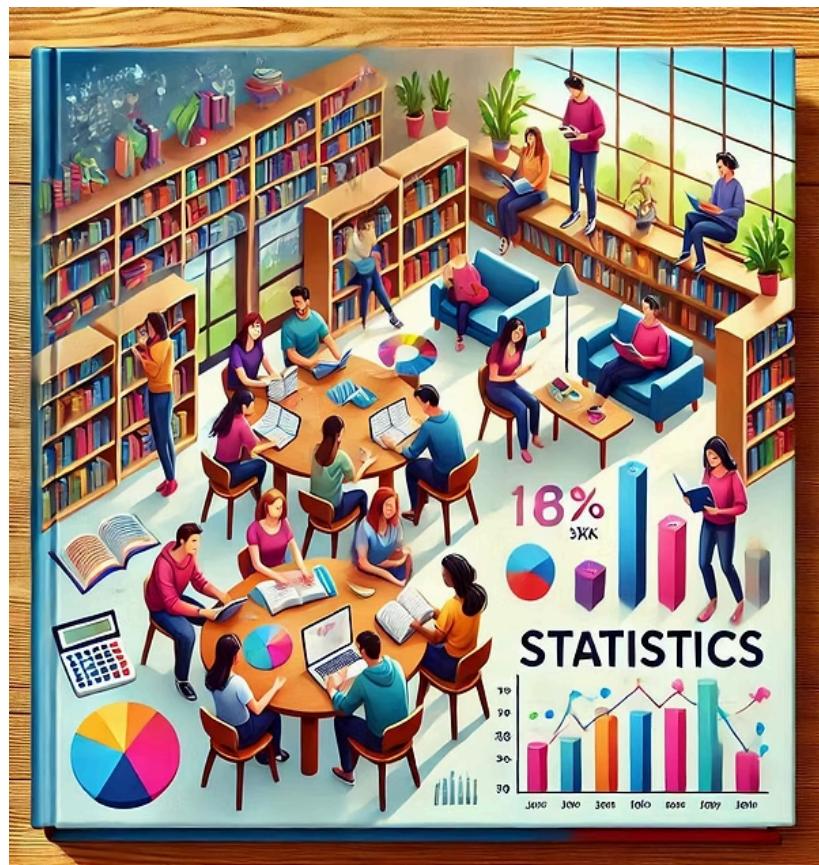
Este libro es una guía integral para estudiantes, profesionales y cualquier persona que busque comprender los fundamentos de la estadística con aplicaciones prácticas y herramientas sólidas.





Libro de Estadística - Daniel Christello

### Mis razones para escribir sobre estadística



Aprendizaje de estadística

[Previous](#)

[Next](#)

Además de las razones generales para aprender estadística, explicadas en [la introducción](#), quiero mencionar aquí algunas razones particulares tanto para entender estadística como para, en mi caso, dedicar tiempo a escribir sobre esta materia.

- Me encontré con la estadística en la década del 80. En mis primeros años de ingeniería tuve que estudiarla y rendirla por primera vez. Era una época donde la estadística era una herramienta muy útil para las ciencias en general pero no era tan popular como ahora en la era de la inteligencia artificial. Recuerdo que en aquella oportunidad mi hermano Rubén (estudiante avanzado de ingeniería) me regalo el libro "Probabilidad y Estadística para Ingenieros de Irwin Miller - John E. Freund". Esta acción, sumado a la buena didáctica del profesor que me tocó en su momento, hicieron que pegara buena onda con la estadística.
- Durante mis años de carrera, no me desarrolle en la ingeniería, sino más bien en gestión. En estos años utilicé principios básicos de estadística que me ayudaron tanto para desarrollar e interpretar encuestas y estudios de mercado como para desarrollar productos, servicios y hacer forecasting de los mismos. Y aunque esta historia

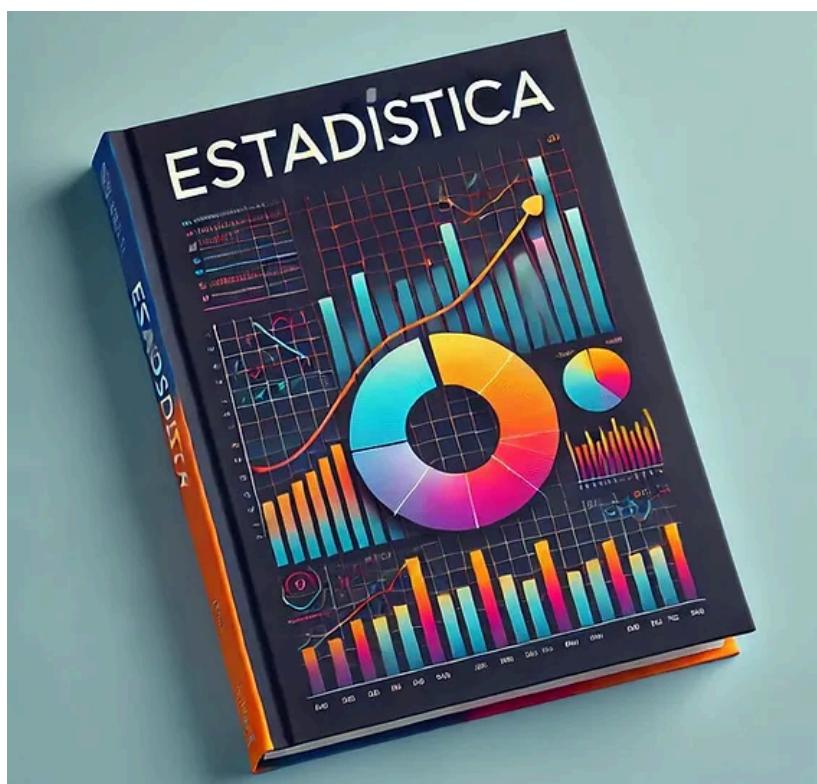
es muy larga, pude comprobar la utilidad de la estadística para entender mejor la probabilidad de los sucesos y tomar así mejores decisiones.

- En la última década me puse a estudiar y certificarme en ciencia de datos. Allí tuve un reencuentro con la estadística al notar que esto del machine learning esta basado en el procesamiento de datos bajo algoritmos eminentemente estadísticos.
- Es importante resaltar aquí, que al estudiar estas cuestiones estadísticas, pude observar una escasa cantidad de material actualizado escrito en español sobre la estadística actual, que si bien es la misma de siempre requiere de una actualización importante debido a las herramientas tecnológicas con que en la actualidad se pueden procesar los datos.
- En paralelo en esta ultima década tuve el regalo de compartir mi vida con un hijo soñado y hecho realidad a quién por un lado se le ha hecho más simpática la matemática que otras disciplinas y por otro lado decidió comenzar su secundaria en una escuela tecnológica. En la primaria pude ver como le ensañaron visualización entendimiento de datos con histogramas y otros conceptos. En la secundaria es **muy probable** que profundice en conceptos estadísticos. Y ni que hablar en la universidad. *Es mi deseo y mi inspiración, que este material le resulte útil cuando llegue ese momento.*

## Dedicatoria

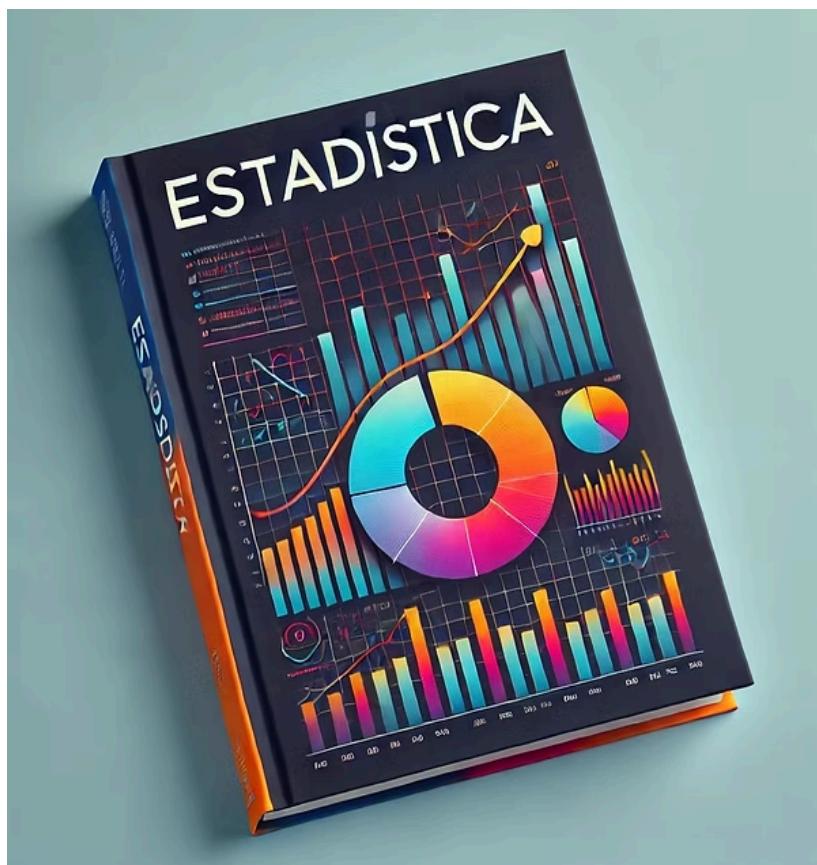
A mi hijo Galo Christello:

- Tengo pocas cosas para dejarte, que perduren en el tiempo, esté o no contigo, en cualquier momento y en cualquier lugar. Esta es una de ellas. Quizás te sirva o quizás no. En cualquier caso, espero que sepas que me hizo muy bien escribir pensando en que te lo puedo dedicar a vos. Es de esas cosas que se llaman amor.





La importancia de ser idóneo en estadística



Estadística en portada de un libro

Previous

Next

Todos los días nos enfrentamos a situaciones con resultados inciertos y debemos tomar decisiones basadas en datos, que muchas veces están... incompletos. Por ejemplo:

- Se espera que los estudiantes universitarios de primer año seleccionen una carrera al ser admitidos, aunque muchos de ellos aún no tengan una meta profesional clara.
- Los pacientes con cáncer pueden ser invitados a participar en estudios clínicos para probar medicamentos experimentales, aun cuando se desconoce información crítica sobre efectos secundarios, tasas de supervivencia y tasas de recurrencia.

Las decisiones empresariales también se toman frecuentemente en entornos de incertidumbre, donde los directivos no pueden predecir con certeza la futura conducta de los factores que influirán en los resultados de las diferentes opciones consideradas. Por ejemplo:

- He tenido la oportunidad de desarrollar (como product manager) internet en Argentina. En ese contexto ser preciso para estimar la demanda y con ello la capacidad de la empresa para satisfacer a la misma se diseño en

un entorno de alta incertidumbre por tratarse de un nuevo mercado que aunque sabíamos que iba a crecer, no era sencillo estimar cuánto. Otro tanto me sucedió con otros nuevos productos y servicios como IPTTransit y CDN.

- En la transformación digital de un periódico los ingresos por publicidad se hacen difíciles de estimar en tanto la publicidad se va volcando del papel a internet, en tanto que los precios de la publicidad digital son muchos menores que los de la edición impresa.
- De manera similar, un inversor no puede predecir con seguridad si los mercados financieros estarán boyantes, estables o deprimidos, pero aún así debe tomar decisiones sobre acciones, bonos e instrumentos de mercado para equilibrar su cartera.

Con esto, se entiende que, las afirmaciones que hacemos a menudo contienen un lenguaje que sugiere un **nivel de certeza** que no siempre es justificado. Consideremos las siguientes declaraciones:

- "La demanda de ancho de banda de internet, será el doble en un plazo de un año"
- "En la transformación digital de un periódico los ingresos por publicidad se crecerán un 15% aún cuando la tirada en papel tienda a cero."
- "El precio de las acciones de tal o cual será más alto dentro de seis meses que ahora."
- "Si el déficit presupuestario público es cero, los tipos de interés irán bajando el resto del año."
- "La renta anual de un titulado universitario en economía será mayor que la renta anual de una persona con estudios tecnológicos"

En el momento en que se hacen estas afirmaciones, es imposible estar completamente seguro de que son ciertas. Incluso un analista que crea que el precio de las acciones de Tesla aumentará en los próximos meses no puede garantizarlo. Por lo tanto, estas afirmaciones deben formularse de una manera que reconozca la incertidumbre inherente, por ejemplo:

- "El precio de las acciones de Tesla **probablemente** será más alto dentro de seis meses que ahora."
- "Si el déficit presupuestario público es tan elevado como se prevé, es **probable** que los tipos de interés se mantengan altos el resto del año."
- "La renta anual de un titulado universitario **probablemente** será mayor que la renta anual de una persona sin estudios universitarios."

Es esencial pensar cuidadosamente en cómo se expresan las ideas, especialmente cuando se refieren a probabilidades o a situaciones de incertidumbre. No es adecuado sustituir afirmaciones demasiado precisas por otras innecesariamente vagas. **¿Qué significa "probablemente" o "es probable que"?** Se debe tener especial cuidado al expresar las ideas de forma clara y precisa para reflejar con exactitud las probabilidades involucradas.

### ¿Por qué estadística?

La estadística es una herramienta valiosa para abordar este tipo de incertidumbre y tomar decisiones informadas. Puede ayudar a extraer el máximo conocimiento de la información disponible y a definir con claridad lo que sabemos y lo que no sabemos. Por ejemplo, transforma declaraciones vagas como "Este medicamento puede causar náuseas" en afirmaciones más precisas como "Tres de cada mil pacientes experimentan náuseas al tomar este medicamento" o "Si no tomas este medicamento, hay un 95 % de probabilidad de que mueras". Sin el apoyo de la estadística, la interpretación de los datos puede ser defectuosa, como se vio en el "Problema de los Tanques Alemanes" durante la Segunda Guerra Mundial, donde el análisis estadístico proporcionó estimaciones mucho más precisas que los métodos de inteligencia estándar.

A lo largo del tiempo se han visto una cantidad considerable de decisiones críticas informadas por la estadística. Por ejemplo, la reciente implementación de modelos estadísticos para gestionar los impactos del cambio climático ha sido fundamental. Gobiernos y organizaciones en todo el mundo han utilizado datos estadísticos para proyectar los efectos de las emisiones de carbono en eventos climáticos extremos, como las olas de calor y los huracanes más frecuentes y devastadores. La capacidad de interpretar y aplicar estos datos correctamente permite a los

líderes políticos tomar decisiones basadas en evidencia, como el diseño de estrategias para mitigar los efectos del clima extremo y planificar infraestructuras resilientes.

Para un científico de datos, la estadística es indispensable. Desde la limpieza y exploración de datos hasta la creación de modelos predictivos, la comprensión de conceptos como la inferencia, la probabilidad y la validación de modelos permite a los científicos de datos sacar conclusiones fiables. En 2024, por ejemplo, los científicos de datos han estado utilizando métodos estadísticos avanzados para optimizar algoritmos de inteligencia artificial en aplicaciones como la detección temprana de enfermedades y la predicción de patrones de consumo energético. Sin estas bases estadísticas, el riesgo de sobreajuste y malas interpretaciones sería enorme, comprometiendo la integridad de los modelos y sus aplicaciones.

Para un economista, la estadística es la base sobre la que se construyen las teorías y los modelos que explican el comportamiento económico. Analizar las fluctuaciones del mercado, estudiar el impacto de las políticas gubernamentales o proyectar el crecimiento económico requiere una comprensión profunda de métodos estadísticos. Con la economía global enfrentando desafíos significativos en 2024, como la inflación y las disrupciones en las cadenas de suministro, los economistas han usado modelos econométricos para asesorar sobre medidas de estabilización y prever los efectos a largo plazo de las decisiones financieras.

Para un profesional de las finanzas, la estadística es una herramienta crítica para gestionar riesgos, evaluar inversiones y construir portafolios diversificados. El análisis estadístico permite comprender la volatilidad de los activos, calcular la probabilidad de eventos extremos y optimizar la asignación de recursos. Este año, la capacidad de manejar modelos de riesgo estadístico ha sido crucial, especialmente en la evaluación de criptomonedas y activos digitales, que siguen mostrando comportamientos impredecibles. Sin una base estadística sólida, los financieros estarían navegando a ciegas en mercados cada vez más volátiles y globales.

Además, la estadística ha sido clave en el sector de la salud, donde los estudios recientes sobre la efectividad de las vacunas actualizadas contra variantes emergentes del COVID-19 y otros virus respiratorios han guiado las campañas de salud pública. Comprender las tasas de eficacia y los intervalos de confianza de los estudios clínicos ayuda a la población a tomar decisiones informadas sobre su salud y bienestar.

En un mundo donde la cantidad de información y la complejidad de los problemas solo aumenta, ser idóneo en estadística no es solo una habilidad técnica, sino una herramienta esencial para comprender el entorno en el que vivimos y **tomar decisiones fundamentadas**. En términos generales, la estadística ayuda a:

1. Clarificar las preguntas y definir las variables clave.
2. Determinar el tamaño muestral necesario.
3. Describir la variación y hacer afirmaciones cuantitativas sobre parámetros estimados.
4. Realizar estimaciones, inferencias y predicciones basadas en los datos.

¿Por qué estadística? En un mundo cada vez más impulsado por datos, ser idóneo en estadística no es solo una ventaja, sino una necesidad. La idoneidad estadística permite interpretar correctamente los datos y evitar caer en conclusiones erróneas o decisiones mal fundamentadas. Comprender las técnicas estadísticas y sus aplicaciones prácticas nos da el poder de analizar fenómenos complejos y comunicar hallazgos de manera precisa. Además, en contextos críticos como la medicina, la economía, las finanzas o la investigación científica, una comprensión deficiente de la estadística puede tener consecuencias desastrosas. Por eso, aprender y dominar la estadística no es solo un ejercicio académico, sino una habilidad esencial para cualquier profesional que aspire a manejar la incertidumbre de manera efectiva y ética.



# Estadística

Una oportunidad para compartir conceptos, anécdotas y ejemplos relacionados al tema.

Más sobre el contenido

## 01 Introducción

Prefacio. Introducción. Contenido. Referencias.

Ver más

## 02 Fundamentos

Métodos Estadísticos. Observar Experimentar. Diseño de experimentos. Normas Experimentales. Poblaciones y muestras.

Ver más

## 03 Descriptiva

Ordenamiento. Clasificación Niveles. Medición. Limites. Fronteras. Anchuras. Huecos. Visualización. Frecuencias. Muestras. Grados. Libertad. Aspectos. Ejemplos. Confusión. Media. Mediana. Moda. Mitad. Desvío., Rango. Intervalo.

Ver más

## 04

### Probabilidad

Sucesos. Probabilidad.  
Reglas de Probabilidad.  
Teorema de Bayes. Probable. Posible.  
Conteo y espacio muestral.  
Permutaciones.  
Conteo. P.  
Variables. Aleatorias.  
Distribución. Probabilidad.

[Ver más](#)

---

## 05

### Distribuciones

Clasificación. Agrupamiento. Fitting.  
Distribuciones Discretas. Distribuciones  
Varias. Bernoulli. Binomial. Multinomial.  
Geométrica. Negativa. Hipergeométrica.  
Poisson. Uniforme. Normal. t-Student.  
Estadísticos. Distribución de Proporciones.  
Distribución de Medias. Distribución chi-  
cuadrado. TCL. Análisis de Normalidad.

[Ver más](#)

---

## 06

### Inferencias

Estadística inferencial. Suceso Infrecuente.  
Intervalo de Confianza. IC. Población  
Finita. Error de  $\bar{x}$  y  $p^{\wedge}$ . Estimadores y  
requisitos. Tamaño de la muestra.

[Ver más](#)

---

## 07

### Pruebas de Hipótesis

t-Student. Normal. Tipos. Falsos ++ y --.  
Contraste.  
Prueba. Hipótesis. Nula. Alternativa. Nivel.  
Significancia. Errores. Tipo 1. Tipo 2. Valor  
Crítico. Región Crítica. Estadístico.  
Valor. p o P. Alfa. Beta. Potencia de una  
prueba. Hipótesis sobre Proporción.  
Hipótesis sobre Media. Hipótesis sobre  
Varianza. Hipótesis de 2 Proporciones.  
Hipótesis de 2 Medias. Varianzas de dos  
muestras. Rachas.

[Ver más](#)

---

## 08

### Análisis Multivariable

Correlaciones Lineales. Pearson.  
Spearman. Kendall. Regresiones.  
Costo. Desvíos. r2. IC e IP.

Valor P. Regresión múltiple. R2 Ajustado.  
Ridge. Lasso. Polinómica. Bondad. Ajuste.  
Tablas. Contingencia. Chi-Cuadrado.  
Independencia.  
Experimentos. Multinomiales. Fisher.  
McNemar. Análisis. Varianza. Factores.  
ANOVA. Requisitos. Tamaños.

[Ver más](#)

---

**09**

## Pruebas No Paramétricas

Pruebas No Paramétricas. Signo.  
Wilcoxon. Kruskal-Wallis. Correlación.  
Spearman. Kendall.

[Ver más](#)

---

**10**

## Supervivencia

Análisis. Tiempos. Supervivencia. Kaplan.  
Meier. Cox.

[Ver más](#)

---

**11**

## Métodos Bayesianos

Estadística Bayesiana. Conceptos.  
Ejemplos.

[Ver más](#)

---

chreinvent@gmail.com

©2021 por Chr. Creada con Wix.com

DANIEL CHRISTELLO