

# Estadística Matemática

Agosto-Diciembre 2021

**Profesor:** Emilien Joly (ext. 531, of. H105, emilien.joly@cimat.mx)

**Ayudante:** Santiago Arenas (santiago.arenas@cimat.mx)

**Horarios de clase:** martes y viernes, 11:00-12:20

**Horarios de prácticas:** miércoles, 11:00-12:20

## Objetivos

1. Familiarizarse con conceptos de teoría de probabilidad de relevancia directa para la formulación de modelos estadísticos y el manejo de propiedades elementales de métodos estadísticos.
2. Estudiar los llamados métodos de estadística descriptiva (numéricos y gráficos) a la luz de dichos conceptos probabilísticos.
3. Adiestrarse en el entendimiento y manipulación de herramientas matemáticas para inferencia.
4. Estudiar propiedades matemáticas primordiales de estimadores clásicos.
5. Conocer algunas clases de modelos estadísticos de uso general.

## Temario

### 1. Preliminares

#### 1.1. Repasos de probabilidad

##### 1.1.1. Espacios de probabilidad

-Medidas, variables aleatorias, función de distribución, densidad, def de  $P_X$ , cuantiles.

##### 1.1.2. Calculo con medidas de probabilidad

-Calculo de valores esperados, varianzas, desviación estándar, independencia, desigualdad de Markov, desigualdad de Chebyshev, desigualdades clásicas de analysis, *i.i.d.*,  $\geq \varepsilon$  o  $> \varepsilon$

##### 1.1.3. Herramientas de probabilidad para caracterización de distribuciones en $\mathbb{R}$

-Casos particulares de  $\mathbb{E}[g(X)]$ , función  $F_X$ , momentos, momentos centrados, función generatriz de momentos, función característica,  $F_X$  caracteriza  $P_X$ , caracterización con funciones de prueba, localización, dispersión, simetría

#### 1.2. Vectores aleatorios

##### 1.2.1. Espacios productos

- $\sigma$ -algebras productos, medida producto, marginales, cálculos de marginales (caso densidad y general), espacios  $(\mathbb{R}^n, \mathcal{B}^n, P_X)$

##### 1.2.2. Cálculos multivariados

- $\mathbb{E}[g(X)]$  con  $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $g : \mathbb{R}^n \rightarrow \mathbb{C}$ ,  $g : \mathbb{R}^n \rightarrow \mathcal{M}_{n \times m}(\mathbb{R})$ , ejemplos de  $e^{itX}$  y de la matriz de covarianza.

##### 1.2.3. Ley de vectores aleatorios

-Ley del vector  $\neq$  ley de cada coordenada, función generatriz, caracterización con funciones de prueba, comentario sobre los cuantiles

#### 1.3. Modelos de probabilidad, modelos de estadística

##### 1.3.1. Introducción sobre la diferencia entre probabilidad y estadística

-Modelación  $\leftrightarrow$  Estudio de características de una distribución

##### 1.3.2. Modelos estadísticos

-espacio  $(\Omega, \mathcal{B}, P_\theta, \theta \in \Theta)$ , Modelos parametricos y no-parametricos, comentarios sobre lo conocido y desconocido en un modelo, observación, muestra, espacio  $(\Omega^n, \mathcal{B}^n, P_\theta, \theta \in \Theta)$ , modelos identificables, error de especificación

- 1.4. Cantidades empíricas
    - 1.4.1. Distribución empírica
      - def de  $P_n$ , notación  $P_n f$ , función de distribución empírica, teorema de Glivenko-Cantelli, momentos empíricos
    - 1.4.2. Estimadores
      - funciones medibles de  $X_1, \dots, X_n$ , notación  $\hat{\cdot}$ ; consistencia, consistencia fuerte, sesgo, riesgo, descomposición sesgo-varianza, información de Fisher, cota de Cramer-Rao
  - 1.5. Convergencia en probabilidad/en distribución/casi-segura
    - 1.5.1. Definiciones de convergencias estocásticas
      - Convergencia en probabilidad, Conv. en distribución, conv casi-segura
    - 1.5.2. Implicaciones entre convergencias
      - Lema de Slutsky,  $\mathbb{P}$  implica cv en distribución, cv c.s. implica cv en  $\mathbb{P}$
    - 1.5.3. Herramientas principales de caracterización de convergencia en distribución
      - Cramer-Wold device, teorema de Levy
  - 1.6. Ley de grandes números, teorema del límite central, método delta.
    - 1.6.1. Ley de grandes números
      - Forma débil, forma fuerte
    - 1.6.2. Teorema del límite central
      - Teorema del límite central, teorema de Lindeberg-Feller
2. Fundamentos de estimación
    - 2.1. Principios de estimación
      - 2.1.1. Método de momentos
        - Momentos empíricos, convergencia
      - 2.1.2. Estimador de cuantiles
        - Notación  $X_{(i)}$ , cuantile  $\hat{q}_{n,\beta}$ , consistencia
      - 2.1.3. Método del máximo de verosimilitud
        - Notación  $f(\cdot|\theta)$ , log-verosimilitud, consistencia, normalidad asintótica
    - 2.2. Regiones de confianza
      - 2.2.1. Definiciones
        - Region de confianza de nivel  $\alpha$ , region de confianza asintótica.
      - 2.2.2. Intervalos de confianza
        - Método, de cuantiles, desigualdades exactas, varianza uniformemente acotada, estimación consistente de  $\sigma^2$ , estabilización con el método delta, intervalos simultáneos
    - 2.3. Pruebas de hipótesis
      - 2.3.1. Formalismo
        - Hipótesis  $H_0, H_1$ , riesgo de primer especie, de segundo especie, estadística de prueba, p-valor, dualidad prueba de hipótesis-region de confianza
      - 2.3.2. Ejemplos
        - Prueba de ajuste de la media, prueba de Kolmogorov-Smirnov, Shapiro-Wilk, prueba de  $\chi^2$ , prueba de Neyman-Pearson
  3. Modelo gaussiano lineal
    - 3.1. Vectores gaussianos
      - gaussianas que no forman un vector gaussiano, función característica, linealidad, independencia, densidad, TCL, definición-teorema  $\chi^2$ , ley de Student, teorema de Cochran, independencia de los estimadores  $\bar{X}_n$  y  $\hat{\sigma}_n^2$
    - 3.2. Condicionamiento
      - Soporte discreto finito, definición de Kolmogorov, proyección  $L^2$ , densidad condicional
    - 3.3. Modelo gaussiano lineal
      - 2 definiciones, ejemplos, estimadores de  $m$  y  $\sigma^2$ , optimalidad

#### 4. Regresión lineal

##### 4.1. Modelo y estimación

-errores *i.i.d.*, regresión sencilla, estimador *MV*, generalizaciones, casos heterocedásticos y homocedásticos

##### 4.2. Regresión logística

-regresión de parámetro, verosimilitud, algoritmos de minimización

##### 4.3. Validación del modelo

-prueba de student, prueba de Fisher, distribución de Fisher

##### 4.4. Predicción

-Notación  $Y^*$ , descomposición error de predicción-error de estimación, intervalos de predicción, estudio de varianza

#### 5. Temas opcionales

##### 5.1. $M$ -estimadores

##### 5.2. $U$ -estadísticas

##### 5.3. Bootstrap

##### 5.4. Simulación de variables aleatorias

##### 5.5. Transforma de Box-Cox

## References

- [1] Gut, A. (2009) *An intermediate course in probability*. Springer Science & Business Media. [link]
- [2] Christensen, R. (2011). *Plane answers to complex questions: the theory of linear models*. Springer Science & Business Media. [link]
- [3] Serfling, R. J. (2009). *Approximation theorems of mathematical statistics*. John Wiley & Sons. [link]
- [4] Van der Vaart, A. W. (2000). *Asymptotic statistics*. Cambridge university press. [link]
- [5] Wasserman, L. (2013). *All of statistics: a concise course in statistical inference*. Springer Science & Business Media. [link]
- [6] Williams, D. (1991). *Probability with martingales*. Cambridge university press. [link]
- [7] Feller, W. (1971) *An introduction to Probability Theory and its Applications*. John Wiley & Sons. [link]
- [8] Knuth, D. (1978) *The art of computer programming*. Vol. 1: Fundamental algorithms. Atmospheric Chemistry & Physics. [link]