# Métodos Estadísticos para IA

## Resumen ejecutivo

Esta monografía desarrolla y mejora los contenidos de la Clase 1 — Parte 1: ‘Métodos Estadísticos para IA’. Cubre definiciones fundamentales de estadística y probabilidad, distribuciones discretas y continuas, cuantiles, generación de números aleatorios, y una introducción a la inferencia estadística — todo orientado a aplicaciones prácticas en Inteligencia Artificial. Cada tema incluye explicaciones ampliadas, fórmulas clave, ejemplos aplicados a la vida real y recomendaciones para ejercicios y notebooks prácticos.

## Objetivos de la monografía

1. Aclarar y ampliar las nociones básicas de probabilidad y estadística necesarias para IA.  
2. Explicar las principales distribuciones (discretas y continuas) con ejemplos reales y fórmulas.  
3. Proveer ejemplos prácticos (Python / SciPy / NumPy) y ejercicios para asentar conceptos.  
4. Conectar los métodos estadísticos con problemas típicos de IA (regresión, clasificación, conteos, tests A/B).

## Introducción a la Estadística y su rol en IA

La estadística proporciona el lenguaje y las herramientas para tratar la incertidumbre y los datos en IA. Desde la estimación de parámetros hasta la validación experimental (A/B tests) y el modelado probabilístico (Bayesiano / frecuentista), los conceptos estadísticos son fundamentales para construir modelos fiables.

## Definiciones básicas

Población: conjunto total de observaciones de interés.

Muestra: subconjunto seleccionado de la población para análisis.

Muestra aleatoria: cada individuo de la población tiene probabilidad de ser seleccionado.

Variable aleatoria (VA): función que asigna un valor numérico a cada resultado de un experimento aleatorio.

## Probabilidades y distribuciones

Probabilidad (informal): cuantifica la incertidumbre de eventos.

Dada una VA X:

- Si X es discreta: se describe mediante la función de masa de probabilidad (PMF, f\_X(x)).  
- Si X es continua: se describe mediante la función de densidad de probabilidad (PDF, f\_X(x)).

Función de distribución acumulada (CDF): F\_X(x)=P(X ≤ x). Sirve tanto para variables discretas como continuas.

## Esperanza y varianza

Esperanza (media): E[X] = suma/integral de x·P(X=x) o ∫ x f\_X(x) dx. Interpretable como el centro de la distribución.

Varianza: Var(X)=E[(X−E[X])²]. Descripción de la dispersión. Desviación estándar = sqrt(Var(X)).  
Propiedades: E[aX+b]=aE[X]+b, Var(aX+b)=a²Var(X). Linealidad de la esperanza.

## Distribuciones de probabilidad discretas (teoría + ejemplos)

### Distribución Binomial

Definición: modela el número de éxitos en n ensayos independientes de Bernoulli con

probabilidad p de éxito.

Notación: X ~ Binomial(n, p).

PMF: P(X=k) = C(n,k) p^k (1−p)^{n−k} para k=0,...,n.

Esperanza: E[X] = n p. Var(X)=n p (1−p).

Ejemplo aplicado (control de calidad): en una línea de producción se inspeccionan 50 piezas (n=50) y la probabilidad histórica de defecto es p=0.02. X representa la cantidad de piezas defectuosas. Con Binomial se calculan probabilidades de tolerancia y se diseña el muestreo.

Código (ejecutar en Jupyter):

```python

import numpy as np

from scipy.stats import binom

n, p = 50, 0.02

k = np.arange(0, n+1)

pmf = binom.pmf(k, n, p)

```

### Distribución Geométrica

Definición: modela el número de ensayos hasta el primer éxito (incluyendo el ensayo exitoso).

Notación: X ~ Geometric(p).

PMF: P(X=k) = (1−p)^{k−1} p para k=1,2,....

Esperanza: E[X]=1/p. Var(X)=(1−p)/p^2.

Ejemplo aplicado (marketing / ventas): modelar cuántos clientes visita un vendedor hasta cerrar la primera venta. Si p es la probabilidad de cerrar por visita, la geométrica ayuda a estimar recursos y pipeline.

Código (ejecutar en Jupyter):

```python

from scipy.stats import geom

p = 0.1

rvs = geom.rvs(p, size=1000)

```

### Distribución de Poisson

Definición: modela el número de eventos que ocurren en un intervalo de tiempo o espacio, cuando los eventos ocurren con tasa promedio λ y son raros/independientes.

Notación: X ~ Poisson(λ).

PMF: P(X=k)=e^{−λ} λ^k / k! para k=0,1,2,....

Esperanza y varianza: E[X]=Var(X)=λ.

Ejemplo aplicado (servicios): número de llamadas a un centro de soporte por minuto. Si la tasa esperada es λ=3, Poisson da la probabilidad de recibir 0,1,2,... llamadas en un minuto y se usa para dimensionar turnos.

Código (ejecutar en Jupyter):

```python

from scipy.stats import poisson

lam = 3

pmf = poisson.pmf(k, lam)

```

## Distribuciones de probabilidad continuas (teoría + ejemplos)

### Distribución Uniforme Continua

Definición: todos los valores en un intervalo [a,b] son igualmente probables.  
PDF: f(x)=1/(b−a) para a ≤ x ≤ b. Esperanza (a+b)/2. Var=((b−a)^2)/12.

Ejemplo aplicado (simulación / inicialización): la inicialización uniforme de pesos en redes neuronales o simulaciones cuando no hay preferencia por un subintervalo concreto.

### Distribución Normal (Gaussiana)

Definición: variable continua con la clásica «curva de campana». Notación: X ~ N(μ, σ²).  
PDF: f(x) = (1/(σ sqrt(2π))) exp(−(x−μ)^2/(2σ^2)).

Propiedades: simétrica respecto a μ; la mayor parte de la masa está dentro de μ ± 2σ.  
Teorema central del límite (TCL): la suma (o media) de muchas variables independientes y con varianza finita tiende a una normal (bajo condiciones), explicación clave de la ubiquidad de la normal.

Ejemplo aplicado: error de medición (sensores), estaturas poblacionales, ruido en modelos de regresión lineal.

Tipificación: Z = (X−μ)/σ transforma N(μ,σ²) a N(0,1). Uso práctico: tablas de probabilidad y cálculos con scipy.stats.norm.

Código (ejemplo):

```Python

from scipy.stats import norm

mu, sigma = 0, 1

p = norm.cdf(1.96, loc=mu, scale=sigma) # prob P(Z ≤ 1.96)

```

### Distribución Chi-cuadrado (χ²)

Definición: suma de los cuadrados de k variables normales estándar independientes. Notación: χ²\_k.

Uso: pruebas de bondad de ajuste, estimación de varianzas y construcción de intervalos de confianza para varianza.

Ejemplo aplicado: contrastar si la varianza de un proceso industrial cumple una especificación (test sobre σ²).

### Distribución t de Student

Definición: aparece al estimar la media poblacional cuando la varianza poblacional es desconocida y el tamaño muestral es pequeño.

Notación: t\_{ν} con ν grados de libertad. Uso frecuente en tests t para comparar medias.  
Ejemplo aplicado: comparar rendimiento medio de dos versiones de un algoritmo en muestras pequeñas (p.ej. A/B test con pocos usuarios).

## Cuantiles

Definición: el cuantil c (0 ≤ c ≤ 1) es el valor x\_c tal que P(X ≤ x\_c)=c. Ejemplos: mediana (c=0.5), cuartiles (0.25, 0.75).

Aplicación práctica: seleccionar umbrales, detectar outliers (p.ej. percentiles 1% y 99%), y construir boxplots para EDA.

Funciones: scipy.stats.distribution.ppf (percent point function) devuelve cuantiles; la función .rvs genera aleatorios.

## Números pseudo‑aleatorios y semillas

Las librerías modernas (NumPy / SciPy) generan números pseudoaleatorios reproducibles si se setea una semilla (seed).

Ejemplo: np.random.seed(42) para reproducir experimentos. Importante para reproducibilidad en experiments ML y pruebas.

## Introducción a la inferencia estadística

Definición: conjunto de métodos para deducir propiedades de una población a partir de una muestra, con un riesgo de error cuantificado.

Métodos paramétricos importantes:

- Estimación de parámetros (estimadores puntuales y por intervalos, ejemplo: estimador de máxima verosimilitud).

- Contraste de hipótesis (tests t, chi-cuadrado, pruebas sobre proporciones, etc.).

Conceptos: error tipo I (α), error tipo II (β), potencia (1−β), p‑valor, intervalos de confianza.

Ejemplo práctico (A/B testing): Se desea comprobar si la versión B mejora la conversión respecto a A. Se calcula la diferencia de proporciones, se formula H0: p\_B = p\_A vs H1: p\_B > p\_A, se calcula la estadística y el p‑valor y se decide con un umbral α (p.ej. 0.05).

## Aplicaciones directas en Inteligencia Artificial

1) Regresión lineal: asume ruido gaussiano en la variable dependiente; estimadores MLE ≈ mínimos cuadrados.

2) Clasificación binaria: modelo Bernoulli para la variable objetivo; la función de pérdida log-loss se deriva de la verosimilitud Bernoulli.

3) Conteos: usar modelos Poisson o regresión Poisson (o negative binomial si hay sobredispersión).

4) Bayesianismo: uso de priors conjugados (ej. Beta para Bernoulli/Binomial) para actualizar creencias a posteriori.

Mini-ejemplo (Beta-Binomial): con una prior Beta(α,β) y datos con k éxitos en n ensayos, la posterior es Beta(α+k, β+n−k). Útil para estimar tasas de conversión.

## Sugerencias de mejora para las diapositivas

1. Añadir esquemas visuales: PMF/PDF y CDF juntos para cada distribución (discreta vs continua).  
2. Incluir ejemplos numéricos completos (con números concretos) y resultados de código en notebooks.

3. Mostrar simulaciones que ilustren el Teorema Central del Límite.

4. Añadir una sección práctica: mini-proyecto (p.ej. análisis de datos de un centro de llamadas y modelado con Poisson).

5. Incorporar una página con resumen de funciones útiles en scipy.stats y snippets listos para copiar.

## Conclusión

Esta monografía ofrece una versión ampliada, aplicada y práctica de la Clase 1. El objetivo fue combinar teoría con ejemplos concretos y recetas prácticas reproducibles en Python, para que el lector no solo comprenda los conceptos, sino que también pueda aplicarlos en problemas reales de IA.

Nota: los fragmentos de código incluidos funcionan en un entorno con NumPy y SciPy instalados.

Generado automáticamente por ChatGPT — Puede pedirse ampliación (v.2) con gráficos embebidos y notebooks Jupyter.