

Clase 4: Teoremas fundamentales y paradigmas estadísticos

FM849 - Programación Científica para Proyectos de Inteligencia Artificial (IA)

6 de enero de 2026

Estadística frecuentista

La estadística frecuentista es un enfoque que busca describir las probabilidades asociadas a un evento como la frecuencia relativa de ocurrencia de dicho evento en un conjunto grande de experimentos o ensayos repetidos bajo las mismas condiciones.

Ejemplo

Desde el enfoque frecuentista, ¿cuál es la probabilidad de lanzar un dado y obtener un número par?

Para responder esta pregunta, imaginemos que podemos lanzar infinitamente el dado. Lo que se esperaría es que aproximadamente la mitad de los lanzamientos resulten en un número par (2, 4, o 6), si es que es un dado normal. Así, $\mathbb{P}(X \in \{2, 4, 6\}) = \frac{3}{6} = 0.5$.

Ley de los grandes números

En el ejemplo anterior, “esperamos” que la frecuencia relativa de obtener un número par se acerque a $\frac{1}{2}$ a medida que el número de lanzamientos del dado aumenta, pero ¿es esto realmente así?

Rigurosidad matemática

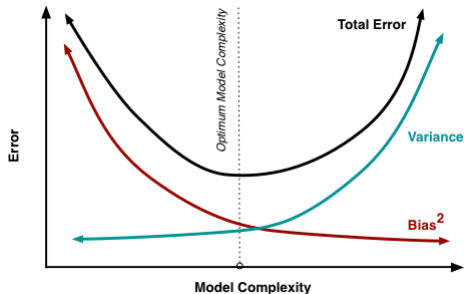
En las matemáticas, cuando se tiene una creencia que no está probada ni refutada, se dice que es una *conjetura*.

La apuesta que realizamos es una conjetura. Pueden probarlo ustedes mismos. Lancen un dado muchas veces y calculen la frecuencia relativa de obtener un número par. ¿Se acerca a $\frac{1}{2}$? Ahora... ¿qué herramienta matemática me asegura que ese es el resultado correcto para la probabilidad?

Ley de los grandes números

Bajo la necesidad de probar este resultado, se enuncia la *ley de los grandes números*. Esta ley establece que, a medida que el número de ensayos independientes de un experimento aumenta, la frecuencia relativa de un evento específico converge hacia la probabilidad teórica de ese evento.

La manera correcta de interpretar esta ley es con el dilema sesgo-varianza, el cual iremos profundizando más adelante en el curso. No caigan en la [falacia del apostador](#).



Aplicación: “la casa siempre gana”

Supongamos que mañana iremos a un casino como curso a jugar la ruleta, y decidimos apostar todo al rojo con los ingresos de la EdV (nadie fue obligado).

Consideremos que en una ruleta común y corriente, existen 18 números rojos, 18 números negros, y 2 números verdes (el 0 y el 00).

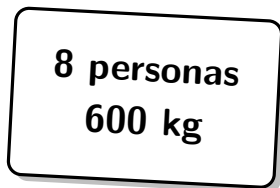
Preguntas abiertas

- ▶ ¿Por qué podemos estar ultraseguros de que la casa siempre gana a largo plazo?
- ▶ ¿Entonces nos han estafado mostrándonos a esas personas que se hacen millonarias?

De nuevo, es importante el pensamiento crítico.

Otro cuestionamiento...

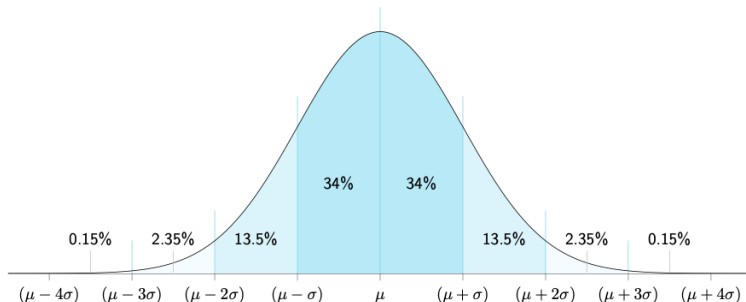
¿Alguna vez se han cuestionado por qué los ascensores aseguran con tanta precisión una masa límite para cierta cantidad de personas?



Teorema central del límite

Aquí entra uno de los resultados más importantes de la teoría de la probabilidad, conocido como el *teorema central del límite*.

Imaginen que entran grupos de 8 personas al ascensor, y miden el promedio de sus masas. Esto lo repiten muchas veces con distintos grupos, y van graficando los resultados obtenidos en un histograma. La distribución tenderá a verse de la siguiente forma, con $\mu = 85$ kg:



Teorema central del límite

En simples palabras, este teorema nos permite afirmar que, independientemente de las distribuciones originales de las variables aleatorias, si generamos una muestra con los promedios muestrales, la distribución resultante tenderá a ser una normal a medida que el tamaño de la muestra de promedios aumenta.

¿Y cuál es la importancia?

La distribución normal es ampliamente estudiada y comprendida, lo que facilita el análisis estadístico y la inferencia en una variedad de contextos prácticos.

Probabilidad condicional

La probabilidad condicionada de un evento A por otro B se anota $\mathbb{P}(A \mid B)$, y representa qué tan probable es que ocurra A , dado que ya ocurrió B .

Ejemplo: Monty Hall adaptado

Elijo al azar a uno de ustedes para participar por 30 décimas. Al frente suyo, dicha persona tendrá 3 cajas. Sólo una de ellas tiene un papel dentro que dice “ganaste las décimas”, y yo sé cuál es. Las otras dos están vacías.

La persona elige una caja al azar. Después, yo abro una de las cajas que no eligió y sé que está vacía. Ahora sólo quedan 2 cajas posibles con el premio. Le doy la oportunidad de arrepentirse y cambiar de decisión. **¿Qué harían ustedes si fueran esa persona?**

Probabilidad condicional

Definamos el evento W como “ganar las décimas” y el evento C como “cambiar de caja después de que yo abro una vacía”. Inicialmente, ninguna caja tiene preferencia, entonces $\mathbb{P}(W) = \frac{1}{3}$.



Si la persona decide cambiar de caja, la probabilidad de ganar es $\mathbb{P}(W \mid C) = \frac{2}{3} > \frac{1}{3} = \mathbb{P}(W)$.

Acá nace la paradoja, porque naturalmente pensaríamos que la nueva probabilidad de ganar es $\mathbb{P}(W) = \frac{1}{2}$, que es igual a la probabilidad de perder. Esto sólo ocurriría si yo no supiese dónde está el premio.

Con este experimento, se comprobó que el cerebro engaña al humano, haciendo que se quede con su opción inicial, lo que genera que pierda el concurso.

Independencia de eventos

Dos eventos A y B se dicen “independientes” (denotado $A \perp B$) si la ocurrencia de uno no afecta la probabilidad de ocurrencia del otro.

Con las definiciones que hemos visto, podemos deducir que si A y B son independientes, entonces deben cumplir que $\mathbb{P}(A \mid B) = \mathbb{P}(A)$ y $\mathbb{P}(B \mid A) = \mathbb{P}(B)$.

Ejemplo: lanzamiento de una moneda

Supongan que lanzan una moneda al aire y quieren adivinar si saldrá cara o sello. Definen A como el resultado del primer lanzamiento, y B como el resultado del segundo. ¿Son A y B independientes? ¿Por qué?

Spoiler: Sí, porque el resultado del primer lanzamiento no afecta al del segundo.

Teorema de Bayes

Con la definición matemática de la probabilidad condicional, es decir, para A y B eventos tales que $\mathbb{P}(B) > 0$,

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

se puede deducir el teorema de Bayes. El teorema enuncia que las probabilidades condicionales dependen de la condicionalidad inversa, o sea:

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(B \mid A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

Teorema de Bayes

Este teorema es muy relevante en estadística bayesiana, ya que a cada uno de los elementos de esta ecuación los dota de un nombre:

- ▶ $\mathbb{P}(A)$: probabilidad *a priori* de A (creencia previa).
- ▶ $\mathbb{P}(B \mid A)$: verosimilitud de B dado A (datos observados).
- ▶ $\mathbb{P}(B)$: probabilidad marginal de B (normalización por motivos matemáticos).
- ▶ $\mathbb{P}(A \mid B)$: probabilidad *a posteriori* de A dado B (creencia actualizada).

Con esto en mente, tenemos una herramienta útil que no teníamos en el enfoque frecuentista: la capacidad de poder incorporar conocimiento previo en nuestros modelos estadísticos.

Estadística bayesiana vs. frecuentista

Ya tienen suficiente información para elegir su camino...

Disagreements now:

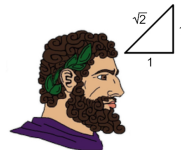


The observed frequency is accurate enough.



Noooo! You need to factor in your priors!

Disagreements then:



I've accidentally found an irrational number.



THROW HIM OVERBOARD.