Apuntes - Probabilidad y variables aleatorias

Curso: Universidad 1

La probabilidad cuantifica la incertidumbre. En este tema se establece el lenguaje de sucesos, la probabilidad condicionada y el teorema de Bayes, y se introducen las variables aleatorias y las distribuciones más usadas: binomial, Poisson, uniforme, exponencial y normal.

Actividad de introducción

Un test rápido de una enfermedad acierta en el 95 % de los enfermos y en el 90 % de los sanos. Si sale positivo, ¿qué probabilidad hay de estar enfermo? La mayoría responde que alrededor del 95 %, pero si la enfermedad es rara la respuesta correcta puede ser menor del 10 %. La intuición falla a menudo con el azar, y por eso necesitamos un cálculo de probabilidades riguroso. En este tema repasamos los sucesos y la probabilidad condicionada, estudiamos las variables aleatorias que asignan números a los resultados de un experimento y conocemos los modelos que más se usan en ciencia, ingeniería y economía: binomial, Poisson, exponencial y normal.

Probabilidad y sucesos

Definición: probabilidad y sucesos

Un **experimento aleatorio** tiene un espacio muestral Ω de resultados posibles; un **suceso** es un subconjunto de Ω. La probabilidad cumple los **axiomas de Kolmogorov**: $P(A) \geq 0$, $P(\Omega) = 1$ y aditividad para sucesos incompatibles. De ahí se deduce $P(A^c) = 1 - P(A)$ y $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Con casos equiprobables, $P(A) = \frac{\text{casos favorables}}{\text{casos posibles}}$, y se cuenta con combinatoria: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.

Probabilidad condicionada e independencia

Definición: probabilidad condicionada e independencia

La **probabilidad condicionada** es $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ (con $P(B) > 0$). Dos sucesos son **independientes** si $P(A \cap B) = P(A)P(B)$, equivalentemente $P(A \mid B) = P(A)$. **Teorema de la probabilidad total:** si $A_1, \ldots, A_n$ es una partición de Ω, $P(B) = \sum P(A_i)P(B \mid A_i)$. **Teorema de Bayes:** $P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)}$.

Variables aleatorias

Propiedad

Una **variable aleatoria** X asigna un número a cada resultado. Si es discreta, se describe con la función de probabilidad $p(x) = P(X = x)$; si es continua, con una **densidad** f ($f \geq 0$, $\int f = 1$, $P(a)$. La función de distribución es $F(x) = P(X \leq x)$. La **esperanza** es $E[X] = \sum xp(x)$ (o $\int x f(x)\,dx$) y la **varianza** $\operatorname{Var} X = E[X^2] - E[X]^2$. Linealidad: $E[aX + b] = aE[X] + b$, $\operatorname{Var}(aX + b) = a^2 \operatorname{Var} X$; para independientes, $\operatorname{Var}(X + Y) = \operatorname{Var} X + \operatorname{Var} Y$.

Distribuciones importantes

Propiedad

Binomial $B(n, p)$: número de éxitos en n pruebas independientes; $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, $E = np$, $\operatorname{Var} = np(1-p)$. **Poisson** $P(\lambda)$: $P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}$, $E = \operatorname{Var} = \lambda$. **Uniforme** en $[a, b]$: densidad constante, $E = \frac{a+b}{2}$, $\operatorname{Var} = \frac{(b-a)^2}{12}$. **Exponencial** λ: $f(x) = \lambda e^{-\lambda x}$, $P(X > x) = e^{-\lambda x}$, $E = 1/\lambda$. **Normal** $N(\mu, \sigma)$: campana simétrica; se tipifica con $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$ y se usan tablas de Φ. Regla 68-95-99,7: $P(|Z|$. La binomial se aproxima por una normal cuando np y $n(1-p)$ son grandes (con corrección de continuidad).

Método: calcular una probabilidad con una distribución

1. Identifica la distribución y sus parámetros.
2. Escribe el suceso con la variable aleatoria.
3. Tipifica o usa la función de probabilidad.
4. Calcula y comprueba que el resultado está entre 0 y 1.

Ejemplo 1

Un test detecta una enfermedad con prevalencia 1 %, sensibilidad 95 % y 5 % de falsos positivos.

$P(E \mid +) = \frac{0{,}01 \cdot 0{,}95}{0{,}01 \cdot 0{,}95 + 0{,}99 \cdot 0{,}05} = \frac{0{,}0095}{0{,}059} \approx 0{,}161$: aunque el test dé positivo, la probabilidad de estar enfermo es solo del 16 %.

Ejemplo 2

En $B(10; 0{,}3)$**:** $P(X = 3) = \binom{10}{3} 0{,}3^3\, 0{,}7^7 \approx 0{,}2668$**;** $E = 3$**;** $\operatorname{Var} = 2{,}1$**.**

Ejemplo 3

Si $X \sim N(100, 15)$: $P(X$ y $P(85$.

¡Atención!

Confundir $P(A \mid B)$ con $P(B \mid A)$; creer que incompatibles implica independientes (si $P(A), P(B) > 0$, son dependientes); sumar varianzas sin independencia o restarlas al restar variables; olvidar tipificar con la desviación típica y no con la varianza; usar la normal sin corrección de continuidad para una binomial.

Herramientas y modelos

Método: resolver un problema de probabilidad

1. Define con precisión los sucesos y escribe los datos como probabilidades, distinguiendo las condicionadas.
2. Si hay etapas o causas posibles, dibuja un diagrama de árbol o una tabla de contingencia.
3. Usa el complementario cuando el suceso contenga «al menos uno».
4. Aplica la probabilidad total para el suceso final y Bayes para invertir la condición.
5. Comprueba que los resultados están entre 0 y 1 y que tienen sentido en el contexto.

Definición: modelos discretos

Una variable binomial $B(n, p)$ cuenta los éxitos en n pruebas independientes con probabilidad de éxito p: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$, con media np y varianza $np(1-p)$. Una variable de Poisson $P(\lambda)$ cuenta sucesos raros que ocurren al azar en un intervalo con media λ: $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, con media y varianza iguales a λ. La hipergeométrica sustituye a la binomial cuando se extrae sin reemplazamiento de una población pequeña.

Definición: modelos continuos

La distribución exponencial de parámetro λ, con densidad $\lambda e^{-\lambda t}$, describe tiempos de espera sin memoria y tiene media $\frac{1}{\lambda}$. La normal $N(\mu, \sigma)$ tiene forma de campana simétrica; se calcula tipificando, $Z = \frac{X-\mu}{\sigma}$, y usando la tabla de $N(0, 1)$. Entre $\mu \pm \sigma$ hay un 68 % de la probabilidad, entre $\mu \pm 2\sigma$ un 95 % y entre $\mu \pm 3\sigma$ un 99,7 %.

Definición: función de distribución

Para cualquier variable aleatoria X, la función de distribución es $F(x) = P(X \leq x)$. Es creciente, vale 0 en $-\infty$ y 1 en $+\infty$, y permite calcular la probabilidad de cualquier

intervalo: $P(a$. En una variable discreta tiene forma de escalera, con saltos en los valores posibles cuya altura es la probabilidad de cada valor. En una variable continua es una función continua cuya derivada es la densidad, de modo que $F(x) = \int_{-\infty}^{x} f(t)\,dt$. Las tablas de la normal que se usan en los exámenes son precisamente valores de la función de distribución de $N(0,1)$, y los percentiles se obtienen resolviendo $F(x) = p$.

Propiedad: esperanza y varianza de sumas

La esperanza es lineal: $E(X+Y) = E(X) + E(Y)$ para cualesquiera variables, sean o no independientes, y $E(aX+b) = aE(X) + b$. La varianza, en cambio, cumple $\operatorname{Var}(aX+b) = a^2 \operatorname{Var}(X)$, y $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$ solo si X e Y son independientes. Como consecuencia, la media de n variables independientes con la misma varianza σ^2 tiene varianza $\frac{\sigma^2}{n}$: promediar reduce la dispersión, y esa es la razón por la que las muestras grandes dan estimaciones más precisas. Esta idea es el punto de partida de la inferencia estadística.

Teorema central del límite

La suma o la media de muchas variables independientes con la misma distribución se comporta aproximadamente como una normal, sea cual sea la distribución de partida. Por eso la binomial con np y $n(1-p)$ mayores que 5 se aproxima por $N\left(np, \sqrt{np(1-p)}\right)$, usando la corrección por continuidad de medio punto, y por eso la normal aparece tan a menudo en la naturaleza.

Ejemplo 4

En una urna hay 3 bolas rojas y 2 azules. Se sacan dos sin devolución. ¿Cuál es la probabilidad de que la segunda sea roja?

Por la probabilidad total: $P(R_2) = \frac{3}{5}\cdot\frac{2}{4} + \frac{2}{5}\cdot\frac{3}{4} = \frac{6+6}{20} = \frac{3}{5}$. Sorprende a muchos, pero sin información sobre la primera bola la segunda tiene la misma probabilidad que la primera.

Ejemplo 5

Un examen tipo test tiene 10 preguntas con 4 opciones. ¿Qué probabilidad hay de aprobar (5 o más aciertos) contestando al azar?

$X \sim B(10;\,0{,}25)$. Sumando $P(X=k)$ para $k = 5, \ldots, 10$ se obtiene aproximadamente $0{,}078$: menos de un 8 %. La media de aciertos al azar es solo $2{,}5$.

Ejemplo 6

Las bombillas de una fábrica de Vitoria duran según una normal de media 800 h y desviación típica 40 h. ¿Qué proporción dura menos de 720 h?

$P(X$: alrededor del 2,3 % de las bombillas.

Probabilidad e intuición

La probabilidad es una de las ramas de las matemáticas en las que la intuición engaña con más frecuencia. La falacia del jugador lleva a creer que, tras varias rojas seguidas en una ruleta, el negro es más probable, cuando cada tirada es independiente. La falacia del fiscal confunde la probabilidad de una prueba sabiendo que alguien es inocente con la probabilidad de que sea inocente sabiendo la prueba, es decir, confunde $P(A \mid B)$ con $P(B \mid A)$. La paradoja del cumpleaños o el problema de las tres puertas muestran que contar bien los casos y escribir con cuidado las condicionadas es imprescindible. Los modelos de probabilidad, por su parte, son idealizaciones: suponen independencia, probabilidades constantes o formas concretas de distribución que conviene contrastar con los datos. Las compañías de seguros, los bancos, los hospitales o los servicios de emergencias toman decisiones diarias con estos modelos, y entender sus hipótesis ayuda a interpretar correctamente los riesgos que leemos en las noticias.

¡Atención!

Sucesos incompatibles e independientes no son lo mismo: dos sucesos incompatibles con probabilidad positiva nunca son independientes. En una variable continua la probabilidad de un valor concreto es 0, así que $P(X \leq a) = P(X$. Y la varianza de una suma solo es la suma de varianzas si las variables son independientes.

¿Sabías que?

La teoría de la probabilidad nació en 1654 en las cartas entre Blaise Pascal y Pierre de Fermat sobre cómo repartir las apuestas de una partida interrumpida. En 1933 Andréi Kolmogórov la fundamentó con tres axiomas, y hoy es la base de la estadística, de la física cuántica y del aprendizaje automático. En España, el matemático Juan Caramuel publicó ya en 1670 un estudio sobre juegos de azar, uno de los primeros tratados de probabilidad escritos fuera de Francia y los Países Bajos, y en la actualidad los sorteos de la lotería se usan en clase como ejemplo clásico de esperanza negativa para el jugador.

El consejo de Otto

dibuja un diagrama de árbol. Las probabilidades de cada rama se multiplican y las de las ramas finales con el mismo resultado se suman: es la probabilidad total sin fórmulas.