Apuntes - Distribuciones binomial y normal

Curso: 2 Bachillerato Ciencias

Actividad de introducción: la campana de las estaturas

Si mides a todos los chicos y chicas de tu instituto y haces un histograma de sus estaturas, obtendrás una figura con forma de campana: muchos valores cerca de la media y cada vez menos al alejarse de ella, de forma simétrica. Lo mismo ocurre con el peso de las barras de pan de una panadería, con las notas de un examen numeroso o con la duración de las baterías de un modelo de teléfono. Esa campana es la distribución normal, la más importante de la estadística. Piensa ahora en otro tipo de situación: lanzas ocho tiros libres y cuentas las canastas, o revisas diez envases y cuentas los defectuosos. Aquí la variable solo toma valores enteros y sigue la distribución binomial. En este tema aprenderás a trabajar con las dos, a usar la tabla de la normal y a aproximar una binomial con muchas pruebas por una normal, una técnica muy frecuente en los problemas de la PAU.

Variable aleatoria discreta

Definición: variable aleatoria discreta

Una **variable aleatoria** X asigna un número a cada resultado de un experimento aleatorio; es **discreta** si toma valores aislados x_i con probabilidades $p_i = P(X = x_i)$, con $\sum p_i = 1$. Su **media** es $\mu = E(X) = \sum x_i p_i$, su **varianza** $\sigma^2 = \sum x_i^2 p_i - \mu^2$ y su **desviación típica** $\sigma = \sqrt{\sigma^2}$.

Distribución binomial

Definición: distribución binomial $B(n, p)$

X es el número de éxitos en n pruebas independientes de Bernoulli (dos resultados: éxito con probabilidad p, fracaso con $q = 1 - p$, constantes en cada prueba). Se escribe $X \sim B(n, p)$.

Fórmulas de la binomial

$P(X = k) = \binom{n}{k} p^k q^{n-k}$, $k = 0, 1, \ldots, n$. Media $\mu = np$; varianza $\sigma^2 = npq$; desviación típica $\sigma = \sqrt{npq}$.

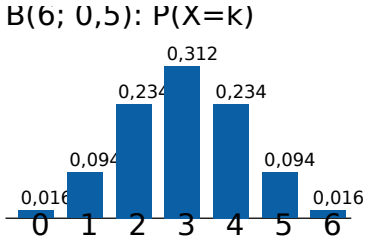


Figura: distribución binomial $B(6;0{,}5)$, simétrica porque $p = 0{,}5$.

Ejemplo 1

Se lanza 5 veces una moneda trucada con $P(\text{cara}) = 0{,}3$.
$P(X = 2) = \binom{5}{2} \cdot 0{,}3^2 \cdot 0{,}7^3 = 10 \cdot 0{,}09 \cdot 0{,}343 = 0{,}3087$.

Ejemplo 2

En $B(8;0{,}25)$, $P(X \geq 1) = 1 - P(X = 0) = 1 - 0{,}75^8 \approx 1 - 0{,}1001 = 0{,}8999$.

Distribución normal

Definición: variable continua y distribución normal

Una variable aleatoria **continua** tiene una función de densidad $f \geq 0$ con área total 1; $P(a \leq X \leq b)$ es el área bajo f entre a y b, y $P(X = a) = 0$. $X \sim N(\mu, \sigma)$ tiene densidad en forma de campana, simétrica respecto de μ, con σ midiendo la dispersión.

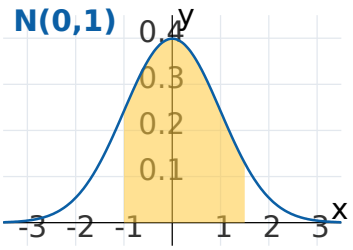


Figura: la probabilidad $P(-1 \leq Z \leq 1{,}5)$ es el área sombreada bajo la normal estándar.

Tipificación y uso de la tabla

Si $X \sim N(\mu, \sigma)$, entonces $Z = \dfrac{X - \mu}{\sigma} \sim N(0,1)$ y $P(X \leq x) = \Phi\left(\dfrac{x - \mu}{\sigma}\right)$. La tabla da $\Phi(z) = P(Z \leq z)$ para $z \geq 0$. Para negativos: $\Phi(-z) = 1 - \Phi(z)$. Además $P(a$ y $P(Z > z) = 1 - \Phi(z)$.

Regla empírica 68-95-99,7

En $N(\mu, \sigma)$: $P(\mu - \sigma$; $P(\mu - 2\sigma$; $P(\mu - 3\sigma$.

Método: probabilidades con la normal

1. Identifica μ y σ y escribe lo pedido como $P(X \leq x)$, $P(X > x)$ o $P(a$.
2. Tipifica restando μ y dividiendo entre σ.
3. Reduce a Φ con la simetría $\Phi(-z) = 1 - \Phi(z)$.
4. Lee la tabla (o calcula) y da la respuesta en el contexto.

Ejemplo 3

$X \sim N(100, 15)$. $P(X > 130) = P\left(Z > \frac{30}{15}\right) = 1 - \Phi(2) = 1 - 0{,}9772 = 0{,}0228$.

Ejemplo 4

$X \sim N(20, \sigma)$ con $P(X > 26) = 0{,}0668$. Entonces $\Phi\left(\frac{6}{\sigma}\right) = 0{,}9332 \Rightarrow \frac{6}{\sigma} = 1{,}5 \Rightarrow \sigma = 4$.

Aproximación de la binomial por la normal

Teorema de De Moivre-Laplace

Si $X \sim B(n, p)$ con $np \geq 5$ y $nq \geq 5$, entonces X se aproxima por $Y \sim N\left(np, \sqrt{npq}\right)$.
Como se pasa de discreta a continua, se aplica la **corrección de continuidad**:
$P(X \leq k) \approx P(Y \leq k + 0{,}5)$, $P(X \geq k) \approx P(Y \geq k - 0{,}5)$,
$P(X = k) \approx P(k - 0{,}5 \leq Y \leq k + 0{,}5)$.

Ejemplo 5

$X \sim B(100; 0{,}4)$: $np = 40$, $\sigma = \sqrt{24} \approx 4{,}899$.
$P(X \leq 45) \approx P(Y \leq 45{,}5) = \Phi\left(\frac{5{,}5}{4{,}899}\right) = \Phi(1{,}12) = 0{,}8686$.

¡Atención!

- Usar la binomial cuando las pruebas no son independientes o p cambia.
- Olvidar la corrección de continuidad en la aproximación.
- Confundir σ^2 con σ al tipificar.
- No usar la simetría para valores negativos de z.

Cómo se plantean estos problemas en la PAU

Observación

Los ejercicios de distribuciones suelen tener dos apartados: uno de binomial (calcular $P(X = k)$, $P(X \geq 1)$ o la media) y otro de normal (tipificar, usar la tabla y a veces

hallar un parámetro o un percentil). Se facilita la tabla de la normal estándar. Indica siempre la distribución («$X \sim B(10;0{,}3)$», «$X \sim N(70,5)$»), la tipificación y la lectura de la tabla. Para hallar un valor x a partir de una probabilidad se hace el camino inverso: se busca z en la tabla y se despeja $x = \mu + z\sigma$.

Ejemplo 6

Notas $X \sim N(6;1{,}5)$. ¿Qué nota deja por encima al 10 % de los alumnos? Se busca z con $\Phi(z) = 0{,}90$: $z \approx 1{,}28$. $x = 6 + 1{,}28 \cdot 1{,}5 = 7{,}92$.

Método: probabilidades con la normal

1. Identifica μ y σ y tipifica $Z = \frac{X-\mu}{\sigma}$.
2. Expresa lo pedido con Φ y la simetría $\Phi(-z) = 1 - \Phi(z)$.
3. Busca el valor en la tabla.
4. Para la binomial con $np \geq 5$ y $nq \geq 5$, aproxima por $N(np, \sqrt{npq})$.

Cómo resolver un problema de distribuciones

Método: binomial o normal

1. Pregúntate qué mide la variable. Si cuenta éxitos en un número fijo de pruebas independientes con la misma probabilidad, es binomial; si es una medida continua (peso, tiempo, longitud) con forma de campana, es normal.
2. En la binomial, escribe n y p, y usa la fórmula para valores concretos; para «al menos uno», usa el complementario.
3. En la normal, tipifica siempre: resta la media y divide entre la desviación típica.
4. Dibuja la campana y sombrea la zona pedida: así sabrás si hay que restar de 1 o usar la simetría.
5. Si la binomial tiene muchas pruebas y se cumplen $np > 5$ y $n(1-p) > 5$, aproxímala por una normal con la misma media y desviación típica, aplicando la corrección de continuidad.

Ejemplo 7

Una aerolínea vende 210 billetes para 200 plazas y el 8 % de los pasajeros no se presenta. ¿Qué probabilidad hay de que falten asientos?

El número de pasajeros que se presentan es una binomial de 210 pruebas con probabilidad 0,92. Como hay muchas pruebas, se aproxima por una normal de media 193,2 y desviación típica de unos 3,93. Faltan asientos si se presentan 201 o más, que con la corrección de continuidad corresponde a valores mayores que 200,5. Tipificando se obtiene un valor de z de unos 1,86 y una probabilidad de aproximadamente 0,031: en unos 3 vuelos de cada 100 habrá que indemnizar a algún pasajero.

¡Atención!

La tabla de la normal solo da probabilidades de la forma «menor o igual que» para valores positivos de z. Para valores negativos o para «mayor que», hay que usar la simetría de la campana y el complementario. En una distribución continua la probabilidad de un valor exacto es cero, así que da igual escribir menor o menor o igual. En la binomial, en cambio, sí importa: «más de 25» significa 26 o más. Por eso, al aproximar una binomial por una normal, hay que traducir con cuidado cada desigualdad y sumar o restar 0,5 según corresponda.

Interpretar los parámetros

Propiedad: qué significan la media y la desviación típica

La media indica dónde está el centro de la campana y la desviación típica indica lo ancha que es. Si dos poblaciones tienen la misma media pero distinta desviación típica, la de mayor desviación tiene más valores extremos. Para comparar datos de poblaciones distintas, como las notas de dos institutos con exámenes diferentes, se comparan los valores tipificados: indican a cuántas desviaciones típicas de la media está cada dato. En control de calidad se vigilan los dos parámetros: si la media se desplaza, aumentan los rechazos por un lado; si la desviación crece, aumentan por los dos.

Ejemplo 8

Lucía saca un 8 en un examen de media 6 y desviación 1,2; Pablo saca un 8,5 en otro de media 5 y desviación 2. ¿Quién ha quedado relativamente mejor?

Lucía está a $\frac{8-6}{1{,}2} \approx 1{,}67$ desviaciones típicas por encima de su media y Pablo a $\frac{8{,}5-5}{2} = 1{,}75$. Aunque las notas parecen parecidas, Pablo ha quedado en una posición relativa algo mejor: lo supera un 4 % de su grupo, mientras que a Lucía la supera un 4,75 % del suyo.

La binomial en situaciones reales

Propiedad: cuándo es razonable el modelo binomial

El modelo binomial exige que las pruebas sean independientes y que la probabilidad de éxito no cambie de una prueba a otra. En la práctica esto se cumple bien al lanzar monedas o dados, o al extraer con reemplazamiento. Si se extrae sin reemplazamiento de un grupo pequeño, la probabilidad cambia en cada extracción y la binomial no es adecuada; pero si el grupo es muy grande comparado con la muestra, como los clientes de una ciudad o las piezas de una gran fábrica, el cambio es tan pequeño que la binomial da resultados excelentes. Antes de aplicar la fórmula conviene escribir en una frase qué es una prueba, qué es un éxito y por qué las

pruebas son independientes: en la PAU esa justificación suele puntuar.

Observación: la campana de Gauss

La distribución normal se llama también campana de Gauss, en honor de Carl Friedrich Gauss, que la usó a principios del siglo XIX para estudiar los errores de las medidas astronómicas. Antes, Abraham de Moivre había descubierto que la binomial con muchas pruebas se parece a esa campana, y Pierre-Simon Laplace generalizó el resultado. El teorema central del límite explica por qué aparece tanto: cuando una magnitud es la suma de muchos efectos pequeños e independientes, como la estatura, que depende de muchos genes y de la alimentación, su distribución se acerca a una normal.

El consejo de Otto

En una normal, dibuja la campana y sombrea lo que te piden. Si el área sombreada incluye el centro, será más de la mitad; si no, menos. Te avisa de resultados absurdos.