

5. Modelos de Distribución

1. Distribución Binomial
2. Distribución de Poisson
3. Distribución Normal
4. Combinación lineal de normales
5. Teorema Central del Límite
6. Aproximaciones por la normal

Distribución de probabilidad

- Una distribución de probabilidad define el comportamiento de una variable aleatoria.
 - Ejemplo: en el caso discreto especifica todos los valores posibles de la v.a. y la prob. de que ocurra cada uno de ellos.
- Toda v.a. tiene una distribución de probabilidad (conocida o desconocida).
- Distribuciones teóricas (se dan en realidad):
 - Discretas: Binomial, Poisson, Hipergeométrica, (Uniforme), ...
 - Continuas: Normal, Exponencial, ...
- Distribuciones en el muestreo:
 - t de Student, χ^2 , F de Snedecor, ...

Distribución de Bernoulli

- Una v.a. X tiene una distribución de Bernoulli de parámetro p si su función de cuantía es

X	0	1
f	q	p

- Se trata de una prueba básica, donde la probabilidad de 1 (éxito) es p y la de 0 (fracaso) es q .
- $p + q = 1$
- Ejemplo: en el lanzamiento de un dado precisamos un 6 para ganar.

X	0	1
f	$5/6$	$1/6$

- Parámetros:

- $E(X) = p$; $\text{Var}(X) = pq$ (comp.)
- $\psi(t) = E(e^{tX}) = p \cdot e^t + q, -\infty < t < \infty$

Distribución Binomial

- Una v.a. tiene una distribución Binomial con parámetros (n, p) si está formada por la suma de n pruebas de Bernoulli de parámetro p .
 - Si tenemos una muestra $\{X_i\}_1^n$ consistente en n pruebas de Bernoulli, entonces la v.a. $X = X_1 + X_2 + \dots + X_n$ tiene una distribución Binomial $X \sim B(n, p)$
- Ejemplo: en un colegio la probabilidad de que un niño contraiga la gripe en un año es de 0'3. Considerar un aula con 20 niños.
 - El nº X de niños con gripe viene dado por: $X \sim B(20, 0'3)$
 - X mide el número de “éxitos” (contraer gripe), por lo que Y medirá el número de “fracasos” (no contraer gripe), por tanto:

$$X + Y = n$$

Distribución Binomial

- Parámetros:

- Función de cuantía

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x \in \{0, 1, 2, \dots, n\} \\ 0, & \text{resto} \end{cases}$$

- $E(X) = np$; $\text{Var}(X) = npq$

- $\psi(t) = \prod_1^n E(e^{tX_i}) = (p \cdot e^t + q)^n, -\infty < t < \infty$

- T^{ma}: Sean k v.a. X_1, X_2, \dots, X_k independientes cada una con distribución binomial $X_i \sim B(n_i, p)$, entonces

$$X = X_1 + X_2 + \dots + X_k \sim B(n_1 + n_2 + \dots + n_k, p)$$

- Ejemplo: nº de caras en el lanzamiento de 4 monedas + nº de caras en el lanzamiento de 6 monedas: binomial del lanzamiento de 10 monedas

Tabla Binomial

$$\text{Binomial } B(n, p) : F(k) = \sum_{i=0}^k \binom{n}{i} p^i q^{n-i}$$

n	k	0.01	0.05	0.10	0.15	0.20	0.25	0.30	1/3	0.35	0.40	0.45	0.50
1	0	0.9990	.9500	.9000	.8500	.8000	.7500	.7000	.6667	.6500	.6000	.5500	.5000
	1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	.9801	.9025	.8100	.7225	.6400	.5625	.4900	.4444	.4225	.3600	.3025	.2500
	1	.9999	.9975	.9900	.9775	.9600	.9375	.9100	.8889	.8775	.8400	.7975	.7500
	2		1	1	1	1	1	1	1	1	1	1	1
3	0	.9703	.8574	.7290	.6141	.5120	.4219	.3430	.2963	.2746	.2160	.1664	.1250
	1	.9997	.9928	.9720	.9393	.8960	.8438	.7840	.7407	.7183	.6480	.5748	.5000
	2	1	.9999	.9990	.9966	.9920	.9844	.9730	.9630	.9571	.9360	.9089	.8750
	3		1	1	1	1	1	1	1	1	1	1	1
4	0	.9606	.8145	.6561	.5220	.4096	.3164	.2401	.1975	.1785	.1296	.0915	.0625
	1	.9994	.9860	.9477	.8905	.8192	.7383	.6517	.5926	.5630	.4752	.3910	.3125
	2	1	.9995	.9963	.9880	.9728	.9492	.9163	.8889	.8735	.8208	.7585	.6875
	3		1	.9999	.9995	.9984	.9961	.9919	.9877	.9850	.9744	.9590	.9375
	4			1	1	1	1	1	1	1	1	1	1

Tabla Binomial

- La tabla representa la función de distribución $F(k) = P(X \leq k)$ de una v.a. binomial $X \sim B(n, p)$
- Alcanza valores hasta $n = 30$. Para valores mayores aproximar por Poisson, la Normal o, si es posible, hacer el cálculo directo.
- La prob. p aparece desde 0'01 hasta 0'50. Si $p > 0'50$ (no aparece en la tabla), entonces cambiar por Y ya que $X + Y = n$ y $p + q = 1$. Por tanto: $Y \sim B(n, q)$, donde $q < 0'50$.
- Para $P(X \leq k) = F(k) \rightarrow$ [tablas]
- Para $P(X > k) = 1 - P(X \leq k) = 1 - F(k) \rightarrow$ [tablas]
- Para $P(X = k) = P(X \leq k) - P(X \leq k - 1) = F(k) - F(k - 1) \rightarrow$ [tablas]
 - En este caso es fácil también usar $f(x)$ directamente.

Problema 5.1

- Sea $X \sim B(3, 0.2)$.
 - a) Calcular $P(X \leq 2)$.
 - b) Calcular $P(X > 2)$.
 - c) Calcular $P(X = 2)$.

Problema 5.2

- La probabilidad de que una persona adulta no sea hipertensa es de 0'7. Si tomamos 4 personas, calcular la probabilidad de que menos de 2 no sean hipertensas.

Problema 5.3

- Se lanzan 20 monedas; calcular:
 - a) Probabilidad de obtener 14 caras.
 - Resolver mediante tablas, función de cuantía y Laplace.
 - b) Probabilidad de obtener al menos 14 caras.

Distribución de Poisson

- La distribución de Poisson mide el número de ocurrencias por unidad de medida.
 - La v.a. es discreta y toma valores enteros $0, 1, 2, \dots$
 - Ejemplos: número de clientes que entran en un establecimiento por hora; defectos por metro en un rollo de tela; despegues de aviones por día; etc.
- El número de ocurrencias es proporcional a la unidad de medida.
- Las ocurrencias deben ser independientes por unidad de medida.

Distribución de Poisson

- Una v.a. X tiene una distribución de Poisson $X \sim P(\lambda)$ con media $\lambda > 0$ si es una v.a. discreta con función de cuantía:

$$f(x) = \begin{cases} e^{-\lambda} \cdot \frac{\lambda^x}{x!}, & x \in \{0, 1, 2, \dots\} \\ 0, & \text{resto} \end{cases}$$

- Ejemplo: el número de descargas por hora de un archivo en un servidor sigue una distribución de Poisson de media 20 por hora.
 - $X \sim P(20)$ en una hora
 - $X \sim P(40)$ en dos horas
 - Las descargas son independientes.

Distribución de Poisson

- Parámetros:
 - f.g.m.: $\psi(t) = e^{\lambda(e^t - 1)}$
 - Media: $E(X) = \psi'(0) = \lambda$
 - Varianza: $\text{Var}(X) = \psi''(0) - \psi'(0)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$
- T^{ma}: Sean k v.a. X_1, X_2, \dots, X_k independientes cada una con distribución de Poisson $X_i \sim P(\lambda_i)$, entonces
$$X = X_1 + X_2 + \dots + X_k \sim P(\lambda_1 + \lambda_2 + \dots + \lambda_k)$$
 - Ejemplo: el número de hombres que entran en una gasolinera por hora es $X \sim P(25)$ y el número de mujeres es $Y \sim P(30)$, luego el número de personas que entran en una gasolinera es $X + Y \sim P(55)$

Tabla de Poisson

$$\text{Poisson } P(\lambda) : F(k) = \sum_{i=0}^k e^{-\lambda} \frac{\lambda^i}{i!}$$

k / λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.9953	.9825	.9631	.9384	.9098	.8781	.8442	.8088	.7725	.7358
2	.9998	.9989	.9964	.9921	.9856	.9769	.9659	.9526	.9371	.9197
3	1	.9999	.9997	.9992	.9982	.9966	.9942	.9909	.9865	.9810
4		1	1	.9999	.9998	.9996	.9992	.9986	.9977	.9963
5				1	1	1	.9999	.9998	.9997	.9994
6							1	1	1	.9999
7										1
k / λ	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2
0	.3329	.3012	.2725	.2466	.2231	.2019	.1827	.1653	.1496	.1353
1	.6990	.6626	.6268	.5918	.5578	.5249	.4932	.4628	.4337	.4060
2	.9004	.8795	.8571	.8335	.8088	.7834	.7572	.7306	.7037	.6767
3	.9743	.9662	.9569	.9463	.9344	.9212	.9068	.8913	.8747	.8571
4	.9946	.9923	.9893	.9857	.9814	.9763	.9704	.9636	.9559	.9473
5	.9990	.9985	.9978	.9968	.9955	.9940	.9920	.9896	.9868	.9834
6	.9999	.9997	.9996	.9994	.9991	.9987	.9981	.9974	.9966	.9955
7	1	1	.9999	.9999	.9998	.9997	.9996	.9994	.9992	.9989
8			1	1	1	1	.9999	.9999	.9998	.9998
9							1	1	1	1

Tabla de Poisson

- La tabla representa la función de distribución $F(k) = P(X \leq k)$ de una v.a. de Poisson $X \sim P(\lambda)$
- Alcanza valores de $n = 21$. Para valores mayores aproximar por la Normal o, si es posible, hacer el cálculo directo.
- Para $P(X \leq k) = F(k) \rightarrow$ [tablas]
- Para $P(X > k) = 1 - P(X \leq k) = 1 - F(k) \rightarrow$ [tablas]
- Para $P(X = k) = P(X \leq k) - P(X \leq k - 1) = F(k) - F(k - 1) \rightarrow$ [tablas]
 - En este caso es fácil también usar $f(x)$ directamente.

Problema 5.4

- Sea $X \sim P(1'5)$.
 - a) Calcular $P(X \leq 3)$.
 - b) Calcular $P(X > 2)$.
 - c) Calcular $P(X = 2)$.

Problema 5.5

- El número de identificaciones con certificado en un servidor web sigue una distribución de Poisson de media 10 por minuto. Calcular:
 - a) Probabilidad de que en un minuto no se identifiquen más de 20 usuarios.
 - b) Probabilidad de que en dos minutos se identifiquen 8 usuarios.
 - c) Probabilidad de que en 15 segundos no acceda nadie.

Aproximación de la Binomial por Poisson

- Relación entre la Binomial y la de Poisson:
 - Dada una v.a. con distribución de Poisson, se toman n intervalos sobre la medida de manera que en cada intervalo sólo tengamos una ocurrencia. De esta forma, cada ocurrencia tendrá la misma probabilidad p , y como son independientes, la probabilidad de un intervalo será $\lambda = n \cdot p$.
 - $X \sim B(n, p) \rightarrow X \sim P(np)$
 - Se utiliza para Binomiales con n grande (>30).
 - La aproximación es buena para $p < 0'5$ y $np < 5$.

Problema 5.6

- La probabilidad de éxito de un correo electrónico malicioso (con virus) es de 0'02. Si se envía a 100 personas, calcular la probabilidad de que se infecten exactamente 5 incautos.