UNIVERSIDAD



Estadística Descriptiva e Introducción a la Probabilidad

Juan Antonio Maldonado Jurado

Departamento de Estadística e Investigación Operativa Universidad de Granada



Doble Grado en Ingeniería Informática y Matemáticas. 2014/15

Tema 6. Algunos modelos de distribuciones discretas

Introducción

Distribución degenerada

Distribución uniforme discreta

Distribución de Bernoulli

Distribución Binomial

Distribución Binomial Negativa

Distribución Hipergeométrica

Distribución de Poisson

Introducción: Modelos de distribuciones y sus aplicaciones

Una vez expuesta la teoría general sobre variables aleatorias y sus distribuciones de probabilidad, vamos a describir algunas distribuciones particulares que han demostrado, empíricamente, ser modelos apropiados para situaciones que ocurren en la vida real. A pesar de ello tales distribuciones presentan un carácter teórico en el sentido de que sus funciones de probabilidad o de densidad se deducen matemáticamente en base a ciertas hipótesis que se suponen válidas para los fenómenos aleatorios.

La elección de una distribución de probabilidad para representar un fenómeno de interés práctico debe estar motivada tanto por la comprensión de la naturaleza del fenómeno en sí, como por la posible verificación de la distribución seleccionada a través de la evidencia empírica. En todo momento debe evitarse aceptar de manera tácita una determinada distribución de probabilidad como modelo de un problema práctico.

Una distribución de probabilidad está caracterizada, de forma general, por una o más cantidades que reciben el nombre de parámetros de la distribución. Un parámetro puede tomar cualquier valor de un conjunto dado y, en ese sentido, se define una familia de distribuciones de probabilidad que tendrán la misma función genérica de probabilidad o función de densidad.



64 Distribución degenerada

La distribución discreta más sencilla es la correspondiente a una variable aleatoria degenerada o constante, es decir, la asociada a un experimento aleatorio que da lugar siempre al mismo resultado. Por tanto, dicha variable aleatoria tomará un único valor c.

Su función masa de probabilidad es

$$P[X = x] = \begin{cases} 1 & x = c \\ 0 & x \neq c \end{cases}$$

Algunas de sus características son:

► Función de distribución:

$$F(x) = P[X \le x] = \begin{cases} 0 & x < c \\ 1 & x \ge c \end{cases}$$

Distribución degenerada

Momentos no centrados:

$$m_k = E[X^k] = c^k P[X = c] = c^k, \qquad k = 1, 2, \cdots$$

Y, en particular, la MEDIA

$$E[X] = c$$
.

Momentos centrados:

$$\mu_k = E[(X - c)^k] = 0, \qquad k = 1, 2, \cdots$$

Y, en particular, la VARIANZA

$$Var[X] = 0.$$

Esta propiedad caracteriza a las distribuciones degeneradas; es decir, una variable aleatoria tiene varianza cero si y solamente si es degenerada en un punto (Propiedades de la varianza).

► Función generatriz de momentos:

$$M(t) = E[e^{tX}] = e^{tc} \quad \forall t \in \mathbb{R}.$$



Distribución uniforme discreta

Esta es la distribución de probabilidad de una variable aleatoria discreta que toma un número finito de valores que son equiprobables y se utiliza para modelizar variables aleatorias asociadas a experimentos aleatorios que tienen un número finito de posibles resultados equiprobables. Su función masa de probabilidad es

$$P[X = x_i] = \begin{cases} \frac{1}{n} & i = 1, 2, \dots, n \\ 0 & \text{en el resto} \end{cases}$$

Se dice entonces que la variable aleatoria X se distribuye uniformemente sobre los puntos x_1, x_2, \dots, x_n

Se dice entonces que la variable aleatoria
$$X$$
 se distribuye uniformemente sobre los puntos x_1 y se notará $X \sim U(x_1, x_2, \dots, x_n)$.

Algunas de sus características son:

Función de distribución:

$$F(x) = \frac{1}{n}(\text{Número de valores } x_i \leq x) = \begin{cases} 0 & \text{si } x < x_1 \\ \frac{1}{n} & \text{si } x_1 \leq x < x_2 \\ \vdots \\ \frac{n-1}{n} & \text{si } x_{n-1} \leq x < x_n \\ 1 & \text{si } x \geq x_n \end{cases}$$

Juan Antonio Maldonado

Curso Académ

Distribución uniforme discreta

Momentos no centrados:

$$m_k = E[X^k] = \frac{1}{n} \sum_{i=1}^n x_i^k, \qquad k = 1, 2, \cdots$$

Y, en particular, la MEDIA

$$E[X] = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}.$$

Momentos centrados:

$$\mu_k = E[(X - EX)^k] = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k, \qquad k = 1, 2, \dots$$

Y, en particular, la VARIANZA

$$\mathsf{Var}[X] = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

► Función generatriz de momentos:

$$M(t) = \mathsf{E}[\mathsf{e}^{tX}] = \frac{1}{n} \sum_{i=1}^n \mathsf{e}^{tx_i} \qquad \forall t \in \mathbb{R}.$$

62 Distribución uniforme discreta

En el caso particular $x_i = i, i = 1, 2, ..., n$

$$E[X] = \frac{1}{n} \sum_{i=1}^{n} i = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$$

$$E[X^2] = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$Var[X] = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2 - 1}{12}$$



🛱 Distribución de Bernoulli

Supongamos un experimento aleatorio que da lugar, únicamente, a dos posibles resultados que son mutuamente excluyentes y exhaustivos, que denotaremos por éxito (E), que será el suceso objeto de estudio, y fracaso (F), que se el complementario de E.

A este tipo de experimentos aleatorios se les llama experimentos o pruebas de Bernoulli. Asociado a un experimento o prueba de Bernoulli y a su correspondiente espacio muestral $\Omega = \{E, F\}$, se define la variable aleatoria con distribución de Bernoulli como

$$X = \begin{cases} 1 & \text{si ocurre el suceso } E \\ 0 & \text{si no ocurre el suceso } E \text{ (ocurre } F\text{)} \end{cases}$$

Si se denota por p a la probabilidad del suceso éxito (E), su función masa de probabilidad es

$$P[X = 1] = p$$

 $P[X = 0] = 1 - p$

$$X = \begin{cases} 1 & \text{si ocurre el suceso } E \\ 0 & \text{si no ocurre el suceso } E \end{cases}$$
 Si se denota por p a la probabilidad del suceso éxito (E), su función masa de probabilidad del suceso (E), su función fu



64 Distribución binomial

Se dice que una variable aleatoria X sigue una distribución binomial de parámetros n y p, $n \in \mathbb{N}, p \in (0,1)$ si modela el número de éxitos en n repeticiones independientes de un ensayo de Bernoulli con probabilidad p de éxito, manteniéndose ésta constante en las n repeticiones del experimento. Su función masa de probabilidad es

$$P[X = x] = \binom{n}{x} p^{x} (1-p)^{n-x}, \qquad x = 0, 1, \dots, n.$$

Se notará como $X \sim B(n, p)$.

$$\mathsf{E}[X] = np; \ \mathsf{Var}[X] = np(1-p)$$

Propiedad de simetría.- Si $X \sim B(n,p)$, entonces la variable aleatoria que contabiliza el número de fracasos, $Y = n - X \sim B(n, 1 - p)$ y, además

$$P[X = x] = P[Y = n - x]$$



Modelos probabilísticos discretos

Distribución Binomial

Distribución Binomial B(n,p)

Valores: $P\{X = k\} = {n \choose n} p^{n} (1-p)^{n-k}; k = 0,1,2,3,...,n$

	k p	0.81	0.05	0.19	8.15	6.20	6.25	0.30	113	9.35	8.40	8.45	0.49	0.50
*		1901	1966	2.840	8.759K	6.6900	4.000	6.400	0.448	3.42%	8.3804	43005	4.364	
		0.000	0.0800	9.7001	1 2501	6.3090	8.3950	6.4000	0.000	3-4251	8.4900	1.4950	5.4000	650
		0.000	9-9626	20466	9.3004	6 1900	6.6626	6 6900		31236	8.1606	83000	63604	0.30
,		0570	0-95N	9.7269	9.5041	6.5100 6.3860	64019	0.900	0.290	9.2546 0.8636	1.2500	3.1004 3.6004	6.1001	
		0.000	0.0306	0.00%	8.3001	£ 1000	1.000	0.400	0.000	2.250	8 Mag	0.1041	6.0006	
	1	1000	0.0001	1000	0.00M	1,1000	£ 81.00	1.600	0.002	10639	1,2501	83041	6.1176	0.05
	÷	1908	0.0001	2.8001	1.520	1 4000	0.00M	0.000	0.005	3135	1.1295	8.865	6.00/1	1.00
	1	0.088	9.17 8	9.26%	1,3695	5.6000	5.6219	0.676	0.000	2.3865	1,2630	1.2005	1,300	120
		0.008	0.0655	0.0495	1.3675	6 1836	6.7500	0.705	0.290	1.14%	1.1490	6.3626	6100	
		0.000	0.000	0.0000	0.26 81	6.1096	5.6900	0.076	0.088	3.11%	1.1536	1 2001	6,2600	0.70
		0.000	0-3000	9-0001	93000	£30m0	6.9064	6.684	04405	1-0456	8.0094	0.1410	6.6526	
v		0.00	OTTE	2.5905	14427	6.3077	6,2073	4 991		3.1101	1309	B-30C8	6.6045	
		0.000	9.2686	9.5249	8.3045	6.4096	6.3066	6.9600	0.8365	9.5434	9.2565	R2066	6.982	0.700
	2	0.000	0-027H	0.0039	8.1863	0.2041	0.2037	0.0000	OXBO	3.3304	11,2400	8.3309	6,3195	0.10
		0.000	0-0011	9.0001	0.0044	4.169.2	6.6079	6 606	0.1985	9.9641	8.2904	9.5063	6.3060	630
	4	0.000	0.0000	9.0000	9.3003	6.3004	0.0040	0.034	0.04.0	1068	1.3704	8.1526	6.970	0.79
		0000	0-0000	1000	83001	6 1006	6 1040	6.004	0.000	1-3001	9.3462	2010	61058	649
	:	0.960	9.7811 9.2621	9.35H 9.35H	8.3571	6.3951	8.1790 8.3590	0.779	0.00%	3-049H 3-2617	8.3467 8.1866	8.3017 8.1009	6.1014	666
	:	000% 87000	0.2671	9.3049 9.009H	5.1762	1,360	1,2000	0.000	0330	9.3200	1,1966	1,7790	1201	0.29
	1	0.000	9.0021	2016	0.790	1,300	6.1018	0.000	0330	9.2805	1,2710	8.3062 8.3062	6301	
	:	0000	0.0001	0000	0.3000	6,004	5.5000	0.000	0000	1200	8.1262	8.1001	5.289	0.75
		0.000	0.000	2000	1.3004	6.3055	5.004	5.000	0000	1000	1,3000	E 2000	0.0004	0.00
			0.000	1000	9.3000	6.3001	6.8000	0.000	000%	1000	1.3041	9.3065	5.000	
,	÷	0.600	1860	1430	1128	1000	4.58	1.004	1000	1000	1,000	£360	1.000	-220
	1	0.000	9259	0.1520	8,3966	6,3890	6,5005	6361	0.2940	3.7541	5.1006	8.3652	6.6906	0.00
		0.000	0.0000	0.0000	0.3001	6.3004	63063	0.000	0.00%	0.0004	0.3673	8.1620	0.0104	0.000
т		OKCF	0-ROH	2430	12021	4.1979	6.1004	6.0079	0.000	1-01 (9	8.3068	8.3084	6.1040	4-00
		0.6345	0.250	9-M(N	8.5647	6.3065	6.2500	0.1077	o yar	9.5579	11.30041	83548	6.0052	142
	2	0.000	0.000	9.7400	12170	6,2900	83089	0.390	0.750	9.2007	1.2090	8.1008	6.7103	0.31
		0-06m	0.0054	0.0621	9.3684	6.986	8.2005	0.294	0.156	9.2506	8 5643	9.2564	63055	0.24
	:	0.000	0.0004	0.0066	0.0101	6.1400	6.9900	0.585	0.1707	3.7675	0.2003	8.2867	6,2750	0.3%
	:	0000	0.000	0-000K	9.3002	6.3067	6.1004	1500	0000	9-360K	0.3413	8.1716	6.7006	0.75
		1000	1000	1000	1.300	£3001	5.1004	6600	1000	1000	0.3413	8 3008 8 3004	61001	0.50
		1000	1200	1000	1,000	6,3000	1.000	0.000	1000	1000	1,3007	8,3057	6.0000	100
	÷	6466	0.550	1.60%	1200	5 ((A)	6.0033	5 000A	1100	1100	8.3861	1.0045	6.000	
•														
	2	0-00K	0.0039	9 (33)	1,2567	6.3000	6 3006	4.208	0.280	9.2902	E-1662	8.10.00	6,67%	447
	,		0.007		5.7009		6,2006		0.750		1.2501		6.1/70	
		0.000	0.000	0.000	9.3006	6.1000	6.1004	0.0004	0.000	0.000	8.3088	B-3063	6.6153	660
		1000	0.000	1000	8.3006	6.3000	E3000	6.000	0.000	3-3001	8.3008	83000	6.9000	
		0-30A	2.340	2340	1.1964	6.074	4,1063	4,630	0445	1000	13061	1.000	6.6563	440
		0.00 M	9361	9.MN	9.5434	1.7504	6:001	0.654	1007	HRS	1.5401	0.000	6,0114	4-013
	2	0.000	0.0166	9.1687	0.2558	6.3000	0.2616	0.3006	0.1800	9.1707	1.1204	9.3053	6.000	0.00
		0-06M	0.0485	0-05N	0.1264	6.2548	6.2504	0.200	0.368	F2568	8.2950	8.1005	6:057	
		0000	0.001	0.010	0.0401	6,3861	6.1990	6,3866	OIIW	3 2577 3 HS86	8.2904	12014	62/00	638
	5	0.000	0.0001	1005	9.3052	6.1064	6.9562	6.659	0.186	2.000	8.3967	8.2540 8.1080	6.3456	0.36
		0.000	0.000	1000	8.30C1	£ 1006	6.0064	0.000	0.000	1000	1.1425	8.1586 8.1586	5 1000	0.38
		0.000	0.000	1000	8.3001	6.1004 6.1001	5.1004	C-0000	0000	1000	0.3425	1,1021	6.000	010
	:	0.000	1200	1000	1.3000	£ 1000	£ 1004	5 0004 5 0004	1000	1300	1304	8.3058 8.3062	6 1065	0.00
		0.000	0.000	2000	0.3000	£ 1000	1.000	0.0000	0.000	1,000	1.3001	1,3001	£ 1006	



6.5 Distribución binomial negativa

Consideremos un experimento aleatorio consistente en repeticiones independientes de ensayos de Bernoulli con probabilidad de éxito constante, hasta que aparezca el éxito r-ésimo. Es decir, en lugar de fijar el número de ensayos y observar el número de éxitos en esas realizaciones, se repiten las realizaciones hasta obtener un número determinado de éxitos y contabilizamos los fracasos.

Definimos la variable aleatoria con distribución binomial negativa como aquella que modela el número de fracasos antes de que aparezca el éxito r-ésimo.

La variable aleatoria X puede tomar los valores $x = 0, 1, 2, \dots$ y r será un entero positivo; es decir, r = 1, 2,

la función masa de probabilidad de esta variable aleatoria es

$$P[X = x] = {x + r - 1 \choose x} (1 - p)^x p^r, \qquad x = 0, 1, 2, \dots \quad (r = 1, 2, \dots; \quad 0$$

$$E[X] = \frac{r(1-p)}{p}; Var[X] = \frac{r(1-p)}{p^2}$$



6.6 Distribución Hipergeométrica

Supongamos una población de N individuos divididos en dos categorías de N_1 y $N_2 = N - N_1$ individuos cada uno. Se elige una muestra de n individuos de la población (sin reemplazamiento o simultáneamente). La variable aleatoria X que contabiliza el número de individuos de la 1^a categoría en la muestra se dice que tiene distribución hipergeométrica de parámetros N, N₁ y n v se nota

$$X \sim H(N, N_1, n), \qquad n, N_1, N \in \mathbb{N} - \{0\}, n, N_1 \leq N$$

La función masa de probabilidad de esta variable aleatoria es

La función masa de probabilidad de esta variable aleatoria es
$$P[X=x] = \frac{\binom{N_1}{x}\binom{N-N_1}{n-x}}{\binom{N}{n}} \qquad \text{máx}(0,n-(N-N_1)) \leq x \leq \min(n,N_1)$$

$$E[X] = np; \quad \text{Var}[X] = \frac{n(N-n)N_1(N-N_1)}{N^2(N-1)}$$

$$E[X] = np; Var[X] = \frac{n(N-n)N_1(N-N_1)}{N^2(N-1)}$$



Distribución de Poisson

Representa el número de ocurrencias de un determinado suceso durante un periodo de tiempo fijo o en una región fija del espacio, cuando el número de ocurrencias sigue unas determinadas pautas:

- El número de ocurrencias en un intervalo o región especificada debe ser independiente del número de ocurrencias en cualquier otro intervalo o región.
- Si se considera un intervalo de tiempo muy pequeño (o una región muy pequeña), la probabilidad de una ocurrencia es proporcional a la longitud del intervalo (al volumen de la región) y la probabilidad de dos o más ocurrencias es prácticamente nula (despreciable).

Con estas consideraciones, si por X_t denotamos el número de ocurrencias del suceso en un intervalo de longitud t(región de volumen t), puede probarse que

$$P\{X_t = k\} = \frac{e^{-\lambda t} (\lambda t)^k}{k!}$$
 $k = 0, 1, 2, ..., \lambda > 0$

La colección de variables aleatorias $\{X_t;\ t\geq 0\}$ constituyen lo que se denomina un PROCESO DE POISSON y cada variable X_t se dice que tiene una distribución de Poisson de parámetro λt .

cada variable X_t se dice En particular, la variable región de volumen unida Su función masa de properties $E[X] = \lambda; Var[X] = \lambda$ En particular, la variable aleatoria que especifica el número de ocurrencias en un intervalo de longitud unidad (o región de volumen unidad), se dice que tiene una distribución de Poisson de parámetro λ y se nota $X \sim \mathcal{P}(\lambda)$. Su función masa de probabilidad es

$$P[X = x] = e^{-\lambda} \frac{\lambda^x}{x!}, \qquad x = 0, 1, \dots, \ \lambda > 0$$

$$E[X] = \lambda; Var[X] =$$

Distribución de Poisson

Distribución de Poisson $P(\lambda)$

