Modelos de Distribución Discreta II

Introducción a la Estadística

DVM

Contenidos

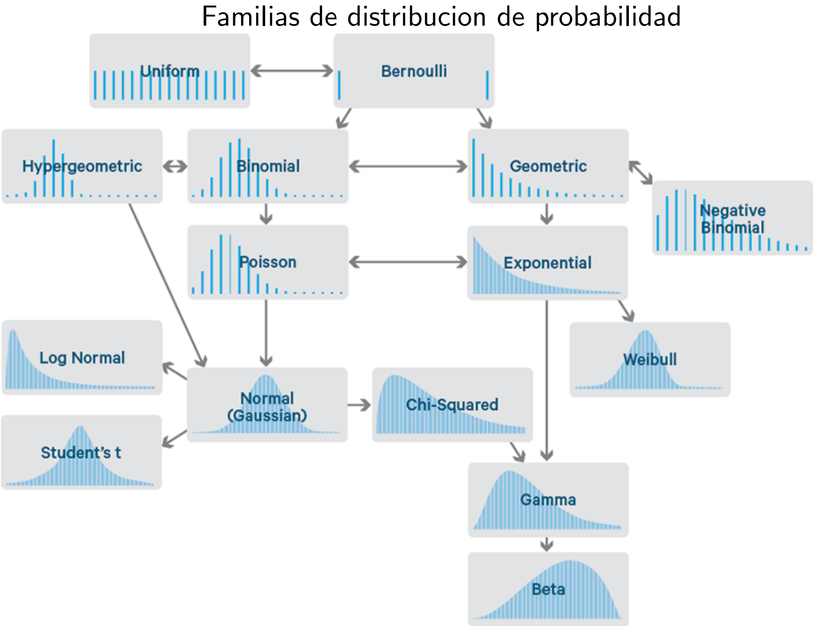
Familias de distribucion de probabilidad

– Cada experimento aleatorio genera al menos una distribución de probabilidad por lo que el estudio de todas las distribuciones de probabilidad es prácticamente imposible.
– Para simplificar, modelamos las distribuciones estadísticas a través de un miembro de una familia de distribuciones.
– Cada familia de distribuciones tiene un espacio de definición (conjunto de los números reales entre los que se reparte la probabilidad), y un espacio paramétrico (distintos valores que pueden tomar los parámetros).

Familia de distribución

Es el conjunto de funciones de distribución que tienen la misma fórmula pero en la que varía el valor del parámetro.

Familias de distribucion de probabilidad



Fuente

Familias de distribucion de probabilidad

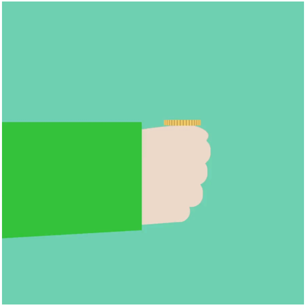
- Dada una situación real (fenómeno aleatorio) debemos elegir el modelo más idóneo que lo represente. No modelaremos el lanzamiento de una moneda de la misma manera que las alturas de las niñas de 14 años.
- Una vez hecha esta elección podremos calcular lo que queramos, con la finalidad de que nos informe sobre una población.

Familia de distribución

Es el conjunto de funciones de distribución que tienen la misma fórmula pero en la que varía el valor del parámetro.

Modelo de Bernoulli

- Lo usamos cuando estamos ante un experimento Bernouilli: Tenemos un espacio muestral con solo dos resultados posibles (exhaustivos) que son mutuamente excluyentes.
- La v.a. toma por tanto dos valores: 1 ó 0. En el caso del lanzamiento de una moneda X=1 si sale "cara" y X=0 si "cruz", por ejemplo. Siendo cara o cruz sucesos equiprobables (pero este no es necesariamente el caso siempre)



Modelo Bernoulli

Una v.a. X sigue una distribución Bernoulli de parámetro p (siendo p la probabilidad de éxito y q la de fracaso) tales que q+p=1 si su función de probabilidad es:

$$f(x) = p^x * q^(1 - x) \text{ para } x = 0 \text{ ó } 1$$

El valor de su media y varianza serán:

$$E(X) = p = \mu \text{ y } \sigma^2 = p * q$$

Se simboliza como X $\rightarrow B(p)$

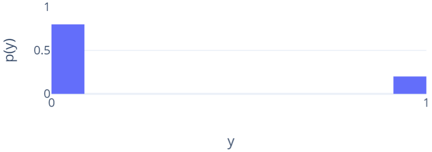
Ejemplos: (Pensemos en variables de tipo categorico con dos unicos vaolres en su rango)

- "Votar Sí" o " Votar No"
- "Pieza buena" o "Pieza Defectuosa"
- "Chico" o "Chica"
- etc.

Modelo Bernoulli

Ejemplo

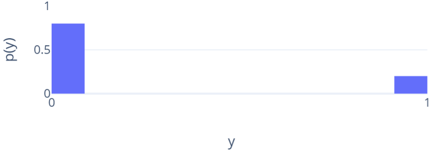
A.



B.

Modelo Bernoulli

A.



B.

A. Y : "*Sacar un diez sin estudiar*"
$$E(y) = 0.8 * 0 + 0.2 * 1 = 0.2$$
$$\sigma^2 = 0.8(0 - 0.2)^2 + 0.2(1 - 0.2)^2 = 0.16$$

B. X : "*Sacar cara al lanzar una moneda*"
$$E(x) = 0.5 * 1 + 0.5 * 0 = 0.5$$
$$\sigma^2 = 0.5(0 - 0.5)^2 + 0.5(1 - 0.5)^2 = 0.25$$

Modelo Binomial

Cuando el experimento aleatorio consiste en la n-repetición de un experimetno Bernoulli con parámetro p y las n pruebas son independientes entre sí.

La variable X (definida como el número de éxitos obtenidos en las n pruebas) sigue una distribución Binomial de parámetros (n, p). Con $0 \leq p \leq 1$, su función de probabilidad viene dada por :

$$f(x) = \binom{n}{x} p^x * q^{n-x} \text{ para } x = 0, 1, 2, ... \text{ y } n = 1, 2, 3.....$$

El valor de su media y varianza serán:

$$E(X) = \mu = n * p \text{ y } \sigma^2 = n * p * q$$

Se simboliza como X $\rightarrow B(n, p)$

Tabla Binomial Puntual | Tabla Binomial Acumulada

Modelo Binomial

Ejemplo

En un proceso de fabricación, la probabilidad de encontrar una pieza defectuosa es de 0,3. Si se toman 5 piezas , ¿cuál es la probabilidad de que salgan 0,1,2,3,4 o 5 piezas defectuosas?

X: "número de piezas defectuosas o D"

n=5

Tenemos que calcular :

Suceso		P(X)
X=0	(B, B, B, B, B)	(q^5)
X=1	(D, B, B, B, B)... $\rightarrow$ *hay* 5 *lotes*$({}^5C_1)$	$(q^4 p) * 5$
X=2	(D, D, B, B, B) ... $\rightarrow$ *hay* 10 *lotes*$({}^5C_2)$	$(q^3 p^2) * 10$
X=3	(D, D, D, B, B)... $\rightarrow$ *hay* 10 *lotes*$({}^5C_3)$	$(q^2 p^3) * 10$
X=4	(D, D, D, D, B) ... $\rightarrow$ *hay* 5 *lotes*$({}^5C_4)$	$(qp^4) * 5$
X=5	(D, D, D, D, D)	(p^5)

Buscamos primero n=5 y p=0,3 en la tabla de distribución binomial.

Modelo Binomial

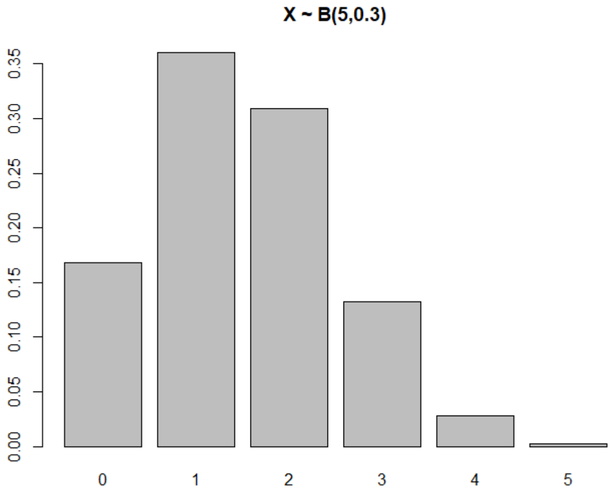
Ejemplo

Tabla de probabilidades puntuales de la distribución ***Binomial(n,p)***
$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

							p					
n	k	0,01	0,05	0,10	0,15	1/6	0,20	0,25	0,30	1/3	0,35	0,40
5	0	0,9510	0,7738	0,5905	0,4437	0,4019	0,3277	0,2373	0,1681	0,1317	0,1160	0,0778
	1	0,0480	0,2036	0,3281	0,3915	0,4019	0,4096	0,3955	0,3602	0,3292	0,3124	0,2592
	2	0,0010	0,0214	0,0729	0,1382	0,1608	0,2048	0,2637	0,3087	0,3292	0,3364	0,3456
	3	0,0000	0,0011	0,0081	0,0244	0,0322	0,0512	0,0879	0,1323	0,1646	0,1811	0,2304
	4	0,0000	0,0000	0,0005	0,0022	0,0032	0,0064	0,0146	0,0284	0,0412	0,0488	0,0768
	5	0,0000	0,0000	0,0000	0,0001	0,0001	0,0003	0,0010	0,0024	0,0041	0,0053	0,0102
6	0	0,9415	0,7351	0,5314	0,3771	0,3349	0,2621	0,1780	0,1176	0,0878	0,0754	0,0467
	1	0,0571	0,2321	0,3543	0,3993	0,4019	0,3932	0,3560	0,3025	0,2634	0,2437	0,1866
	2	0,0014	0,0305	0,0984	0,1762	0,2009	0,2458	0,2966	0,3241	0,3292	0,3280	0,3110
	3	0,0000	0,0021	0,0146	0,0415	0,0536	0,0819	0,1318	0,1852	0,2195	0,2355	0,2765

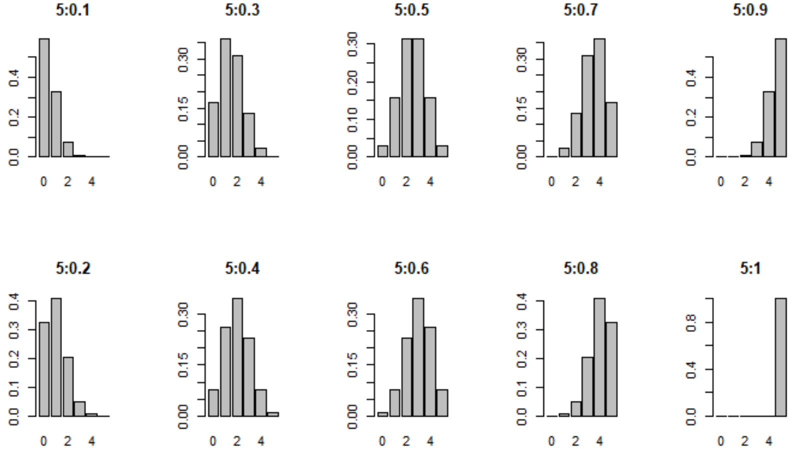
Modelo Binomial

Ejemplo



Modelo Binomial

Ejemplo



Modelo de Poisson

Esta variable aleatoria está asociada a experimentos donde se pretende contar el numero de ocurrencias de un evento en una unidad dada de tiempo, espacio, distanciao volumen. Además:

1. La aparición de sucesos puntuales por unidad de observación estable. Produce a largo plazo un número medio de sucesos constante por unidad de observación (tiempo, espacio, área, etc...)
2. Los sucesos que aparecen por unidad de observación son independientes de lo que pueden aparecer en la siguiente unidad observada.

Ejemplo:

- El numero de accidentes de trafico por día
- El numero de infectados

Modelo de Poisson

Una variable aleatoria X sigue una distribución de Poisson si su función de probabilidad es:

$$f(x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \text{ para } x = 0, 1, 2...$$

El valor de su media y varianza serán:

$$E(X) = \mu = \lambda \text{ y } \sigma^2 = \lambda$$

Se simboliza como X$\rightarrow P(\lambda)$

Tabla de Poisson Puntual
Tabla de Poisson Acumulada

Modelo de Poisson

Ejemplo

1.Las llamadas por averías en un puesto de servicio siguen una distribución Poisson de media 2 averías por semana. Calcular la probabilidad siguiente:

a) Ninguna avería en una semana.

b) Menos de cinco averías en una semana.

c) Menos de seis en un mes (cuatro semanas)

Modelo de Poisson

Ejemplo

Tenemos la variable aleatoria X: "Número de averías por semana con $E(X) = 2 = \lambda$

a) $P(X = 0) = \frac{e^{-2} * 2^0}{0!} = 0,1353$ o lo miramos directamente en la tabla.

b) $P(X \leq 5) = P(X = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4) + P(x = 5) = 0,9473$ o lo miramos directamente en la tabla (de prob. acumulada).

c) La media para cuatros semanas = 2*4 así que $\lambda = 8$; $P(X < 6) = P(X = 0) + P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4) + P(x = 5) = \frac{e^{-8} * 8^0}{0!} + ... + \frac{e^{-8} * 8^5}{5!} = 0,1912$

2. Supongamos que 300 erratas están distribuidas aleatoriamente en un libro de 500 páginas. Hallar la probabilidad de que una página dada tenga:

a) Exactamente 2 erratas

b) Dos o más erratas

Modelo de Poisson

Ejemplo

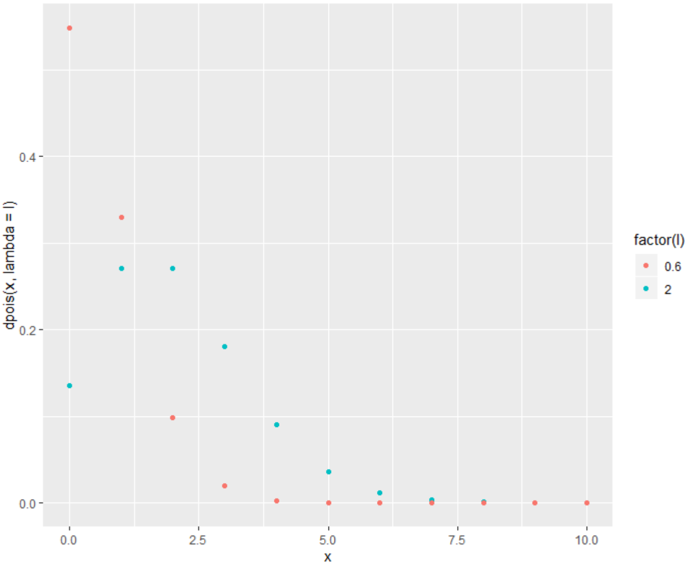
Tenemos la variable aleatoria X: "Número de erratas por página" con $E(X) = 3/5 = \lambda$

a) $P(X = 2) = \frac{e^{-0.6} * 0.6^2}{2!} = 0,0988$

b) $P(X \geq 2) = 1 - P(X < 2) = 1 - 0,8781 = 0,1219$

Modelo de Poisson

Ejemplo



Modelo de Poisson

Pequeña nota

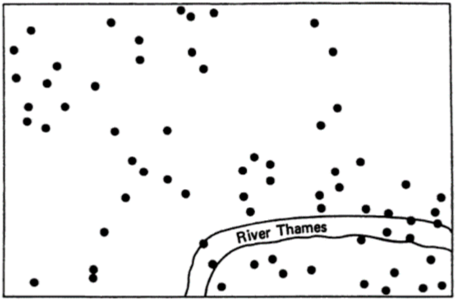


FIGURE 8.1. London V-1 and V-2 Rocket Impact Pattern
SOURCE: Based on Gilovich, 1991, p. 20; used by permission.

Hastie, Reid Dawes, Robyn. (2010). Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Marking

- Si λ es pequeña, los eventos son muy raros pero, aleatoriamente, pueden aparecer ocasionalmente formando clusters: "Poisson clumps"
- Esta clusterizazcion de Poisson se usa para explicar repentinos incrementos en la frecuencia de ciertos eventos como ataques de tiburones , coincidencias, cumpleaños...
- Se dice que Poisson dscubrio esto cuando estaba investigando porque mucho oficiales de la armad prusiana estaban muriendo por coces caballos.

Aproximacion Binomial a Poisson

Por lo general una variable aleatoria de Poisson puede estudiarse como una v.a. binomial donde el número de repetición del experimento n tiende a infinito.

$$X \rightarrow B(n, p) \xrightarrow[p \rightarrow 0]{n \rightarrow \infty} X \rightarrow P(\lambda)$$

En donde $n * p = \lambda$ en general con $np > 5$
Ejemplo:
Hacienda informa que el 5,5% de los contribuyentes comenten errores al rellenar la declaración de la renta si se eligen aleatoriamente 100 de ellos. ¿Cuál es la probabilidad de que menos de tres contengan errores?

Aproximacion Binomial a Poisson

Por lo general una variable aleatoria de Poisson puede estudiarse como una v.a. binomial donde el número de repetición del experimento n tiende a infinito.

$$X \to B(n,p) \xrightarrow[p \to 0]{n \to \infty} X \to P(\lambda)$$

En donde $n * p = \lambda$ en general con $np > 5$

Ejemplo:

Hacienda informa que el 5,5% de los contribuyentes comenten errores al rellenar la declaración de la renta si se eligen aleatoriamente 100 de ellos. ¿Cuál es la probabilidad de que menos de tres contengan errores?

np$= \lambda = 100 * 0,055 = 5,5$

$P(X < 3) = \sum_{x=0}^{2} \frac{e^{-5.5}5.5^x}{x!} = 0.0041 + 0.0225 + 0.0618 = 0.0884$

Distribución Hipergeometrica

Un experimento hipergeométrico es aquel que posee las siguientes propiedades:

1) Se selecciona una muestra de tamaño n **sin remplazamiento** de un total de resultados N.
2) Np resultados del total de N pueden clasificarse como éxitos y N-Np como fracasos.
3) La probabilidad no es constante en las n pruebas.

Distribución Hipergeometrica

Se dice que $X \rightarrow H(Np, N-Np, n)$ El número x de éxitos en n recibe el nombre de variable aleatoria hipergeométrica. Su función de probabilidad P(X=x) es:

$$f(x) = \frac{\binom{Np}{x}\binom{N-Np}{n-x}}{\binom{N}{n}}$$

Donde:
N: tamñano de la población
n: tamaño de la muestra
x: valor que toma la variable
Np: tamaño de uno de los grupos
$E(x) = \mu = np$ y $\sigma^2 = npq\frac{N-n}{N-1}$

Distribución Hipergeometrica

Ejemplo 1

1. En un estanque hay 500 peces. Para un estudio biológico se extrae a 50 de ellos, se les pone una chapa indentificativa y se les devuelve al estanque. Al cabo de un mes se extraen 10 peces para un control rutinario. Calcula la probabilidad de que 2 de ellos estén marcados. (suponemos que en ese tiempo no hay nacimientos ni muertes de peces).

Distribución Hipergeometrica

Ejemplo 1

1. En un estanque hay 500 peces. Para un estudio biológico se extrae a 50 de ellos, se les pone una chapa indentificativa y se les devuelve al estanque. Al cabo de un mes se extraen 10 peces para un control rutinario. Calcula la probabilidad de que 2 de ellos estén marcados. (suponemos que en ese tiempo no hay nacimientos ni muertes de peces).

- Entonces: $\frac{{}^{50}C_2 * {}^{450}C_8}{{}^{500}C_{10}}$

Distribución Hipergeometrica

Ejemplo 2

2.Se va a proceder a la siembra de una nueva variedad de semilla. De cada 10 que se plantan una no fructifica. De un lote de 40 semillas se extrae un lote de 10. ¿Cuál es la probabilidad de que haya 2 semillas que no fructifiquen?

Distribución Hipergeometrica

Ejemplo 2

2.Se va a proceder a la siembra de una nueva variedad de semilla. De cada 10 que se plantan una no fructifica. De un lote de 40 semillas se extrae un lote de 10. ¿Cuál es la probabilidad de que haya 2 semillas que no fructifiquen?

- Entonces: $\frac{{}^{4}C_{2} * {}^{36}C_{8}}{{}^{40}C_{10}}$

Distribución Hipergeometrica

Ejemplo 3

3.En un estado, la corte de apelaciones está formada por 7 jueces. Para un caso rutinario, tres de ellos se eligen al azar como jurado para escuchar el caso y emitir un dictamen. Si 5 de los 7 jueces son potencialmente proclives a cierto argumento legal, ¿Cuál es la probabilidad de que en el jurado aparezcan dos de dichos jueces? ¿Y que sean todos?

Distribución Hipergeometrica
Ejemplo 3

3.En un estado, la corte de apelaciones está formada por 7 jueces. Para un caso rutinario, tres de ellos se eligen al azar como jurado para escuchar el caso y emitir un dictamen. Si 5 de los 7 jueces son potencialmente proclives a cierto argumento legal, ¿Cuál es la probabilidad de que en el jurado aparezcan dos de dichos jueces? ¿Y que sean todos?

- $P(dos\ jueces\ proclives) = \frac{{}^{5}C_{2} * {}^{2}C_{1}}{{}^{7}C_{3}}$
- $P(todos\ jueces\ proclives) = \frac{{}^{5}C_{3}}{{}^{7}C_{3}}$