TEMA 5. ALGUNAS DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

5.1. Distribución Uniforme discreta
5.2. Distribución binomial
5.3. Distribución hipergeométrica
5.4 Distribución de Poisson
5.5. Distribución geométrica

1

Las distribuciones de probabilidad son funciones matemáticas que presentan comportamientos similares a fenómenos reales de la naturaleza.

Hablaremos pues de distribuciones especiales que han demostrado, **empíricamente**, ser modelos útiles para diversos problemas prácticos.

Se estudiaran los siguientes aspectos dentro de cada distribución:

-Tipo de experimento que le caracteriza

-Variable que la define

-Funciones de probabilidad y de distribución

-Interpretación de los parámetros que las determinan

-La esperanza matemática y la varianza

5.1. DISTRIBUCIÓN UNIFORME DISCRETA

Una variable aleatoria discreta sigue una distribución uniforme en n puntos si su función de probabilidad es:

$$p_k = \mathrm{P}[X = x_k] = \frac{1}{n} \qquad k = 1, \ldots, n$$

Ejemplo. En el experimento "tirar un dado" la variable X definida como "puntuación obtenida" sigue una distribución uniforme. Su función de probabilidad es: $p_k = P[X = k] = \frac{1}{6}, \quad k = 1, \ldots, 6$

Media: $E[X] = \sum_{k=1}^{n} x_k p_k = \sum_{k=1}^{n} x_k \frac{1}{n} = \frac{\sum_{k=1}^{n} x_k}{n}$

Varianza:

$$E\left[X^2\right] = \sum_{k=1}^{n} x_k^2 p_k = \sum_{k=1}^{n} x_k^2 \frac{1}{n} = \frac{\sum_{k=1}^{n} x_k^2}{n} \qquad Var[X] = \frac{\sum_{k=1}^{n} x_k^2}{n} - \left(\frac{\sum_{k=1}^{n} x_k}{n}\right)^2$$

3

Ejemplo. Obtener esperanza y varianza

X	$p_k = P[X = k]$	x_k^2
1	1/6	1
2	1/6	4
3	1/6	9
4	1/6	16
5	1/6	25
6	1/6	36
21	1	91

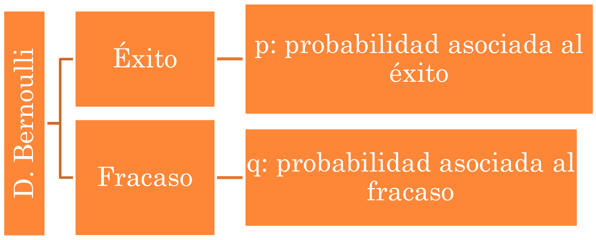
$$EX = \frac{21}{6} = 3.5$$
$$VarX = \frac{91}{6} - 3.5^2 = 2.916$$

4

5.2 DISTRIBUCIÓN BINOMIAL

DISTRIBUCIÓN DE BERNOULLI

Un experimento se dice que sigue una distribución de Bernoulli cuando presenta dos posibles resultados:



Como ambos sucesos son incompatibles se verifica que p+q=1.

A dicho experimento le asociamos la variable aleatoria X la cuál solamente toma dos valores:

$$X = \begin{cases} 1 & si\ ocurre\ Exito(A) \\ 0 & si\ ocurre\ Fracaso(\overline{A}) \end{cases}$$

5

Su distribución de probabilidad es:

$$\mathsf{P}[X=1]=p$$
$$\mathsf{P}[X=0]=q=1-p$$

Esta distribución se conoce como **Distribución de Bernoulli** de parámetros 1 y *p: B(1,p)*

Propiedades:

Media: $E[X]=\sum_{k=0}^{1} x_k p_k = (0\times q)+(1\times p)=p$

Varianza: $E\left[X^2\right]=\sum_{k=0}^{1} x_k^2 p_k = \left(0^2\times q\right)+\left(1^2\times p\right)=p$

$$\sigma^2[X]=E\left[X^2\right]-E[X]=p-p^2=p(1-p)=pq$$

Ejemplo. En el experimento "Tirar una moneda" , la variable X definida como "resultado" Sigue una distribución de Bernoulli.

X	$p_k = P[X=k]$
1	1/2 (cara)
0	1/2 (cruz)

$$EX = p = 0.5$$
$$Var\ X = pq = 0.5(1-0.5) = 0.25$$

6

Distribución binomial

La suma de n variables aleatorias independientes e idénticamente distribuidas según una distribución Bernoulli se conoce como **distribución Binomial**, $X \sim B(n,p)$

$$X = X_1 + X_2 + \ldots + X_n \qquad X_i \sim B(1,p),\ i = 1,\ldots,n$$

La v.a. X puede tomar todos los valores enteros comprendidos entre 0 y n, ambos inclusive, y representa el ***número de éxitos*** que se obtienen en n pruebas idénticas e independientes con una probabilidad de éxito de p.

Su distribución de probabilidad es $p_x = \mathrm{P}[X = x] = \binom{n}{x} p^x q^{n-x} \qquad x = 0,\ldots,n$

Media: $E[X] = E[X_1 + X_2 + \ldots + X_n] = E[X_1] + E[X_2] + \ldots + E[X_n] = p + p + \ldots + p = np$

Varianza: $\sigma^2[X] = \sigma^2[X_1 + X_2 + \ldots + X_n] = \sigma^2[X_1] + \sigma^2[X_2] + \ldots + \sigma^2[X_n]$

$$= pq + pq + \ldots + pq = npq$$

Propiedad de aditividad. Sean X e Y dos v.a. independientes con distribuciones de probabilidad $X \sim B(n,p)$ e $Y \sim B(m,p)$ entonces $X + Y \sim B(n+m,p)$

7

Ejercicio 1 relación

Experimento = Tirar moneda

n = 3 veces

$p = probabilidad\ exito\ (cara) = 0.5$

X = nº caras en 3 tiradas $\sim B(3{,}0.5)$

$P[X = 2] = 0.375$

(a R)

$EX = 3 * 0.5 = 1.5$

$VarX = 3 * 0.5 * (1 - 0.5) = 0.75$

Ejercicio 3 relación

Experimento = Contestar preguntas test

n = 10 preguntas

$p = probabilidad\ exito\ (acertar) = 0.25$

X = nº aciertos en 10 preguntas $\sim B(10{,}0.25)$

$d) P[X \geq 8] = 1 - P[X \leq 7] = 1 - 0.9996 = 0.0004$

(a R)

8

Ejemplo (libro J.A. Hermoso)

Un jugador profesional de póker que está pasando un fin de semana en Las Vegas decide jugar 10 partidas el sábado y 5 el domingo. Si la probabilidad de que gane una partida es del 40%. Calcule la probabilidad de que en el fin de semana gane en más de 3 ocasiones.

Solución:

X=*número de partidas ganadas el sábado* $\sim \mathcal{B}(10\,,\,p=0,40)$

Y=*número de partidas ganadas el domingo* $\sim \mathcal{B}(5\,,\,p=0,40)$

Aplicando la propiedad de aditividad:

Z=*número de partidas ganadas el fin de semana*=$X+Y \sim \mathcal{B}(10+5=15\,,\,p=0,40)$

$$P[X+Y>3]=1-P[X+Y\leq 3]=1-0.0905=0.9094$$

9

5.3. DISTRIBUCIÓN HIPERGEOMÉTRICA

Se selecciona una muestra (subconjunto) de una población finita con ***N*** elementos donde los individuos están clasificados en dos categorías (*éxito* y *fracaso*). Se puede hacer básicamente de dos formas:

Muestreo con reemplazamiento: los individuos se extraen, observan y devuelven a la población, de forma que la composición de la población es constante en cada extracción. Para calcular la probabilidad de que cierto número de individuos de la muestra presentan la característica *éxito* se utiliza la distribución *Binomial*.

Muestreo sin reemplazamiento: los individuos extraídos se dejan fuera de la población con lo que la composición de la población va cambiando con cada extracción (es equivalente extraer de uno en uno o en conjunto). Para calcular la probabilidad de que cierto número de individuos de la muestra presentan la característica *éxito* se utiliza la distribución *Hipergeométrica* .

La variable aleatoria ***X=número de éxitos en una muestra sin reemplazamiento de tamaño n*** sigue una **distribución de probabilidad Hipergeométrica** cuyas probabilidades asociadas son:

10

$$p_x = \mathrm{P}[X = x] = \frac{\binom{Np}{x}\binom{Nq}{n-x}}{\binom{N}{n}} \qquad X \sim H(N, n, p)$$

Donde Np es el número de elementos que presentan la característica éxito y Nq es el número de elementos que presentan la característica fracaso.

Media: $E[X] = np$

Varianza: $\sigma^2[X] = \frac{N-n}{N-1} npq$

Cuando N es infinito (en la práctica un valor muy elevado) las distribuciones *Hipergeométrica* y *Binomial* coinciden.

Ejercicio 12 relación

N = 3000 bujías

n = 10 bujías

$p = probabilidad\ exito\ (bien) = 1 - \frac{400}{3000} = 0.8667$

$a)$ X = nº bujias buenas en muestra sin reemplazamiento de tamaño 10 $\sim H(3000, 10, 0.8667)$

$P[X = 10] = 0.2385$

$b)$ X = nº bujias buenas en muestra con reemplazamiento de tamaño 10 $\sim B(10, 0.8667)$

$P[X = 10] = 0.2391$

5.4. DISTRIBUCIÓN POISSON

Una variable aleatoria X sigue una distribución de Poisson de parametro λ si puede tomar todos los valores enteros no negativos con probabilidades:

$$p_x = \mathrm{P}[X = x] = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0, 1, 2, \ldots$$

Una de las aplicaciones comunes de esta distribución es calcular la probabilidad del **número de ocurrencias de un suceso en un determinado periodo de tiempo**.

Propiedades:

Media: $E[X] = \lambda$ Varianza: $\sigma^2[X] = \lambda$

Propiedad de aditividad. Sean X e Y dos v.a. independientes con distribuciones de probabilidad $X \sim P(\lambda_1)$ e $Y \sim P(\lambda_2)$ entonces $X + Y \sim P(\lambda_1 + \lambda_2)$

Distribución de Poisson como límite de la distribución Binomial. Cuando en la distribución Binomial la probabilidad es muy pequeña y *n* es suficientemente grande, se puede aproximar por la distribución de Poisson de parámetros *np*

13

Ejercicio 5 relación

$\lambda = 4$ llamadas por minuto por termino medio

X = nº llamadas por minuto $\sim P(4)$

$P[X = 0] = 0.01831$

(a R)

$EX = 4$

$VarX = 4$

Ejercicio 6 relación

$\lambda = 10$ fallos por hora por termino medio

X = nº fallos por hora $\sim P(10)$

$P[X = 1] = 0.00045$

$P[X > 1] = 1 - P[X \leq 1] = 1 - 0.0005 = 0.9995$

(a R)

14

Ejercicio 9c relación

$\lambda = 2$ averías por semana por termino medio

X = nº averias por semana $\sim P(2)$

Y = nº averias en 4 semanas $\sim P(2 * 4 = 8)$

$P[Y < 6] = P[X \leq 5] = 0.1912$

15

5.5. DISTRIBUCIÓN GEOMÉTRICA

La variable aleatoria X=*número de fracasos antes de obtener el primer éxito* en repeticiones idénticas e independientes de un experimento. La variable X sigue una **distribución Geométrica de parametro p**, donde p es la probabilidad de éxito.

X puede tomar todos los valores enteros no negativos (0,1,2,..) con probabilidades:

$$p_x = \mathrm{P}[X = x] = (1-p)^x p = q^x p \qquad x = 0, 1, 2, \ldots$$

Media: $E[X] = \dfrac{1-p}{p} = \dfrac{q}{p}$

Varianza: $\sigma^2[X] = \dfrac{1-p}{p^2} = \dfrac{q}{p^2}$

Ejercicio 14 relación

X = nº suspensos antes de aprobar $\sim G(0.70)$

$P[X \leq 5] = 0.999271$

16