

DISTRIBUCIONES DE PROBABILIDAD



VARIABLES ALEATORIAS DISCRETAS

MATEMATICA PARA INGENIEROS

Distribución Uniforme



Sea X una **variable aleatoria discreta** cuyos k valores posibles son x_1, x_2, \dots, x_k , con igual probabilidad de ocurrencia, entonces

la **función de masa de X** es

$$p_X(x) = \begin{cases} 1/k & \text{para } x_1, x_2, \dots, x_k \\ 0 & \text{en otro caso} \end{cases}$$

y la **distribución** se llama **uniforme**

Ejemplos



- ❑ Cuando un foco se selecciona al azar de una caja que contiene un foco de 40 W , uno de 60 W , uno de 75 W y otro de 100 W
- ❑ Cuando se lanza un dado.
- ❑ Si se selecciona “al azar” uno de los k enteros positivos: $1, 2, \dots, k$, donde la frase “al azar” significa que los k enteros tienen la misma probabilidad de ser seleccionados
- ❑ Si se selecciona al azar un alumno entre 100 y definimos la variable aleatoria X como el número con que se identifica a cada alumno (a cada alumno se le hace corresponder un único número entre los enteros del 1 al 100)

❑ *No se puede asignar una distribución uniforme a una sucesión infinita de valores posibles pero se puede asignar una distribución de este tipo a cualquier sucesión finita de resultados igualmente probables.*

❑ **Representación gráfica** de la Distribución Uniforme

**gráfico de bastones
de puntos
histograma**

En el eje de abscisas representamos los valores que toma X
y el eje de ordenadas corresponde a los valores de $p_X(x)$

Distribución Uniforme sobre los enteros



Función de masa

$$p_{\mathbf{X}}(x) = \begin{cases} 1/k & \text{para } x = 1, 2, \dots, k \\ 0 & \text{en otro caso} \end{cases}$$

Esperanza

$$E(\mathbf{X}) = (k + 1) / 2$$

Varianza

$$\text{VAR}(\mathbf{X}) = (k^2 - 1) / 12$$

PROCESO DE BERNOUILLI

Toda experiencia aleatoria que presenta dos resultados posibles, uno que fijamos como “éxito” y lo contrario como “fracaso” se suele denominar de tipo Bernouilli.

Por ejemplo las siguientes experiencias son de tipo Bernouilli:

- Jugar a cara o cruz.
- El nacimiento de un niño (varón o mujer)
- Obtener un “4” en el lanzamiento de un dado.
- Obtener un número “par” en el lanzamiento de un dado.
- La fabricación de una pieza en una fabrica (aceptable o defectuosa).
- El resultado de una operación (éxito o fracaso).
- Aprobar o Desaprobar un examen.

Cuando se realizan n pruebas independientes de tipo Bernoulli decimos que es una experiencia aleatoria de tipo Binomial.

Por ejemplo las siguientes experiencias son de tipo Binomial:

- Jugar 10 veces a cara o cruz .
- Observar 30 nacimientos de un bebé (niña o niño)
- Obtener el “4” en el lanzamiento de 15 dados.
- La fabricación de 1000 piezas en una fabrica (aceptable o defectuosa).
- El resultado de 50 operaciones (éxito o fracaso).

Distribución Binomial



La **función de masa** para **X** siguiente

$$p_X(x) = \begin{cases} \binom{n}{x} p^x \cdot (1-p)^{(n-x)} & \text{si } x = 0, 1, 2, \dots, n \\ 0 & \text{en otro caso} \end{cases}$$

se llama **binomial** con **parámetros** **n** y **p**

se simboliza **$X \sim b(n, p)$**

Ejemplo



- ❑ Supóngase que una máquina produce un artículo defectuoso con probabilidad p ($0 < p < 1$) y produce uno no defectuoso con probabilidad $q = 1 - p$
- ❑ Que se examinan *independientemente* n artículos de los producidos por la máquina.
- ❑ Se define X como la variable aleatoria “**número de artículos defectuosos**” obtenidos en este experimento
- ❑ los valores posibles de X serán $0, 1, 2, \dots, n$ **artículos defectuosos** y para cada uno de ellos existe una probabilidad asociada.
- ❑ La probabilidad $P(X = x)$ es la **probabilidad de obtener cualquier sucesión ordenada de n artículos consistente de x artículos defectuosos y $n - x$ no defectuosos.**

□ p es la probabilidad de seleccionar **un artículo defectuoso**

Como los artículos se seleccionan en forma independiente **la probabilidad de elegir x artículos defectuosos es**

$$p^x = \underbrace{p \cdot p \dots p}_{x \text{ veces}}$$

empleando la regla multiplicativa para eventos independientes

□ $q = 1 - p$ es la probabilidad de seleccionar **un artículo no defectuoso**

La probabilidad de seleccionar $n - x$ artículos no defectuosos es

$$\underbrace{q \cdot q \dots q}_{n - x \text{ veces}} = q^{(n-x)} = (1 - p)^{(n-x)}$$

□ **La probabilidad de obtener alguna sucesión ordenada en particular de n artículos conteniendo x defectuosos y $n - x$ no defectuosos es**

$$p^x \cdot (1 - p)^{(n-x)}$$

Consideremos ahora el **número de sucesiones diferentes posibles de n artículos con x defectuosos y $n - x$ no defectuosos** .

Tenemos en cuenta que los artículos defectuosos no se distinguen entre sí y lo mismo sucede con los no defectuosos

□ Como las formas de arreglar r objetos distintos es $r!$; si separamos los n objetos en un lote de x objetos y otro de $n - x$, las formas de arreglar el lote de x objetos con el lote de $(n - x)$ objetos es $x!(n - x)!$

□ Si sucede que en cada grupo de objetos no podemos distinguir un objeto de otro entonces los n objetos separados en dos grupos uno de x objetos indistinguibles y otro de $(n - x)$ objetos distintos a los del otro lote pero indistinguibles entre sí podrán arreglarse en

$$\frac{n!}{(n - x)! x!} = \binom{n}{x} = \text{número combinatorio } n \text{ sobre } x$$

□ Se tienen $\binom{n}{x}$ sucesiones con x artículos defectuosos y $n - x$ no defectuosos y para cada una de ellas la probabilidad es $p^x \cdot (1 - p)^{(n-x)}$, entonces :

$$P(X = x) = \binom{n}{x} p^x \cdot q^{(n-x)}$$

Proceso Binomial



- ❑ El experimento consiste en n **pruebas repetidas** u observaciones seleccionadas de una *población infinita sin reemplazo* o de una *población finita con reemplazo*
- ❑ Cada prueba tiene **dos resultados posibles** que pueden clasificarse en **éxito** o **fracaso** .
- ❑ La probabilidad de éxito , representada por p permanece **constante** para todos los intentos .
- ❑ Los intentos repetidos son **independientes** , es decir , el resultado de cualquier prueba es independiente del resultado de cualquier otra prueba .
- Este **modelo** se aplica a poblaciones finitas o a poblaciones conceptualmente infinitas como las piezas que producirá una máquina o las personas que contraen una cierta enfermedad , de las que tomamos elementos al azar **con reemplazamiento**
- En esta situación el proceso generador debe ser **estable** , proporción de piezas defectuosas o de personas enfermas se mantiene constante al largo plazo y **sin memoria** , el resultado en cada momento es independiente de lo previamente ocurrido.



Ejemplo : Supongamos el proceso de seleccionar 3 artículos en forma independiente de un proceso que se supone produce un 10 % de piezas defectuosas y se los clasifica como D (Defectuosos) o N (No defectuosos)

- ❑ Un artículo defectuoso se considera un éxito.
- ❑ El número de éxitos es una variable aleatoria X que asume los valores enteros 0, 1, 2 y 3
- ❑ Los artículos se seleccionan en forma independiente de un proceso que supone un 10 % de piezas defectuosas

$$p = 0,10 \quad \text{y} \quad q = 1 - p = 0,90 .$$



El número de éxitos , X : “*cantidad de piezas defectuosas*” ,
en $n = 3$ experimentos es una variable aleatoria binomial
 $X \sim b (n = 3 , p = 0,1)$

$$p_X(x) = \begin{cases} \binom{3}{x} 0,1^x \cdot 0,9^{(3-x)} & \text{si } x = 0, 1, 2, \dots, n \\ 0 & \text{en otro caso} \end{cases}$$

Reemplazando resulta

x	0	1	2	3
$p_X(x)$	0,729	0,243	0,027	0,001

La probabilidad de obtener hasta dos artículos defectuosos

$$P(X \leq 2) = p_X(0) + p_X(1) + p_X(2) = 1 - p_X(3) = 0,999$$

Esperanza y varianza



- Dada una variable aleatoria con **Distribución Binomial**

$$X \sim b(n, p)$$

n es el **número de intentos**

p la **probabilidad de éxito** en un intento determinado,

la **esperanza** y la **varianza** de **X** son

$$E(X) = n p$$

$$VAR(X) = n p q$$

Si se debe extraer una muestra de tamaño $N = 100$ de una población de 1.000.000 de habitantes donde se conoce que 600.000 son varones



Obtenemos

- **La probabilidad de que el primer individuo seleccionado sea varón**

$$(600.000 / 1.000.000) = 0,6$$

- **La probabilidad de extraer un segundo individuo varón**

$$(599.999 / 1.000.000) = 0,5999$$

- **La de extraer un tercero**

$$(599.998 / 1.000.000) = 0,5999$$

- **y así sucesivamente**

Si comparamos estos cocientes, vemos que la probabilidad prácticamente no varía de repetición en repetición y podemos adaptar el experimento a un modelo de muestreo con reemplazo y se puede usar un *modelo binomial*

Consideremos , en cambio , otra situación ; en la reunión de una cooperadora escolar a la que asisten 20 personas , entre las que hay 14 mujeres ; se quieren seleccionar 5 personas para integrar una subcomisión



- Sea X : “ cantidad de mujeres que integran la comisión “ , entonces :
- La probabilidad de obtener la primera mujer
 $(14 / 20) = 0,70$
- La probabilidad de elegir la segunda mujer
 $(13 / 19) = 0,684$
- La probabilidad de elegir una tercera
 $(12 / 18) = 0,67$ y así sucesivamente

Vemos aquí que , la *probabilidad condicionada de extraer mujeres de esa población* , no se mantiene constante en cada repetición del experimento y no resulta posible, en consecuencia , aplicar el esquema binomial

Distribución Hipergeométrica



- Surge así la **Distribución Hipergeométrica** que es una distribución cercanamente relacionada a la distribución binomial pues la variable surge de la extracción de elementos de una determinada población a los que podemos clasificar en una de dos categorías mutuamente excluyentes denominadas éxito y fracaso.
- En la distribución hipergeométrica **no se requiere independencia** y se basa en un **muestreo** llevado a cabo **sin reemplazo** donde **la probabilidad cambia de un intento a otro**

Se desea conocer la probabilidad de que a la comisión de 5 personas la integren 3 mujeres

□ Esto es la probabilidad de seleccionar 3 mujeres de las 14 disponibles y 2 varones entre los 6 disponibles para constituir la comisión de cinco miembros.

□ Se tienen $\binom{14}{3}$ combinaciones de 3 mujeres tomados del grupo de 14 .

□ Se pueden seleccionar de $\binom{6}{2}$ maneras a 2 hombres de entre los 6 disponibles

□ El número total de modos de seleccionar 3 mujeres y 2 hombres en 5 intentos es el producto

$$\binom{14}{3} \binom{6}{2}$$

□ El número total de maneras de seleccionar 5 personas de las 20 disponibles es

$$\binom{20}{5}$$

□ La probabilidad de seleccionar 5 personas (sin reemplazo) de las cuales 3 son mujeres y 2 son hombres es :

$$p_x(3) = \frac{\binom{14}{3} \binom{6}{2}}{\binom{20}{5}}$$

En este problema X es la variable aleatoria que determina el número de éxitos de un experimento , en este caso es “ *el número de mujeres* “ que pueden integrar una comisión de 5 personas



Los valores posibles o el recorrido de X resultan : $REC X = \{ 0, 1, 2, 3, 4, 5 \}$ y el tamaño de la muestra es $n = 5$.

El éxito es obtener x mujeres y el fracaso es obtener $n - x$ hombres.

La selección se efectúa en un grupo de 20 personas , o sea que el tamaño de la población es $N = 20$.

La cantidad total de éxitos posibles es el número total de mujeres en la población y lo designamos con k , en este caso $k = 14$. La cantidad de fracasos posibles es el número de hombres del grupo , o sea $N - k$ en nuestro caso $N - k = 20 - 14 = 6$

$$p_X(x) = P(X=x) = \frac{\binom{14}{x} \binom{20-14}{5-x}}{\binom{20}{5}} \quad \text{si } x = 0, 1, 2, 3, 4, 5$$

Distribución Hipergeométrica de la VA X de parámetros $N = 20$, $n = 5$, $k = 14$

$$X \sim h(N, n, k) = h(20, 5, 14).$$

Distribución Hipergeométrica



La distribución de probabilidad de una variable aleatoria **hipergeométrica**

X “**número de éxitos en una muestra aleatoria de tamaño n seleccionada de N artículos totales, de los cuales k son considerados como éxitos y $N - k$ como fracasos**” es

$$p_X(x) = h(N, n, k) = \begin{cases} \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} & \text{si } x = 0, 1, 2, \dots, \min\{n, k\} \\ 0 & \text{en otro caso} \end{cases}$$

Esperanza y Varianza de $h(N, n, k)$



$$X \sim h(N, n, k)$$

Esperanza $E(X) = n \left(\frac{k}{N}\right)$

Varianza $VAR(X) = n \left(\frac{k}{N}\right) \left[\frac{(N-k)}{N}\right] \left[\frac{(N-n)}{N-1}\right]$

$(k/N) = p$ proporción de éxitos de la población

$(N-k)/N = q$ proporción de fracasos

$$E(X) = n p \quad VAR(X) = n p q \left[\frac{(N-n)}{(N-1)}\right]$$

$(N-n)/(N-1)$ **factor de corrección para poblaciones finitas**

Se observa que la esperanza es la misma que para la distribución binomial y la varianza se obtiene multiplicando la varianza de la distribución binomial por el **factor de corrección** $(N-n)/(N-1)$

Distribución de Poisson



Corresponde a la **variable aleatoria discreta** que mide el **número de resultados que ocurren en un intervalo de tiempo t** , o en una **región específica t**

$$p_x(x) = \begin{cases} \frac{e^{-\lambda t} (\lambda t)^x}{x!} & \text{para } x = 0, 1, 2, \dots \\ 0 & \text{en otro caso} \end{cases}$$

λ

número promedio de resultados por unidad de tiempo o región

El **intervalo de tiempo puede ser de cualquier duración** , por ejemplo , un minuto , un día , una semana , un mes o inclusive un año .



La variable aleatoria **X** puede representar

- ***el número de llamadas por hora que se reciben en una oficina .***
- ***el número de automóviles que llegan por día a una casilla de peaje .***
- ***el número de juegos pospuestos por lluvia durante la temporada de fútbol***
- ***el número de partículas radiactivas que pasa por un contador durante un intervalo de 3 milésimas de segundo.***

La **región específica** puede ser un segmento de línea , un área , un volumen o , tal vez , un pedazo de material



En este caso **X** puede representar :

- ***el número de ratas de campo por acre***
- ***el número de bacterias en un determinado cultivo***
- ***el número de pasas de uva en un pan de navidad***
- ***el número de errores de mecanografiado por página***

Proceso de Poisson



Satisface tres condiciones que surgen de la *observación de un fenómeno concreto durante un período de tiempo fijo o en una región específica*

Primera condición

La ocurrencia de un éxito en cualquier intervalo es estadísticamente independiente de aquella en cualquier otro intervalo

Proceso de Poisson



Segunda condición

La probabilidad de observar exactamente un éxito en un intervalo es estable.

La **probabilidad de una** ocurrencia durante cualquier **intervalo** de tiempo o región debe ser aproximadamente **proporcional a la longitud** de ese intervalo (o al tamaño de la región) y no depende del número de ocurrencias fuera de ese intervalo o región .

Tercera condición

La probabilidad de que haya dos o más ocurrencias en el intervalo de tiempo o región debe ser despreciable en comparación con la probabilidad de una ocurrencia .

Distribución de células de levadura en 400 cuadrículas de un aparato utilizado para este recuento llamado hemacitómetro



Cantidad de células por cuadrícula	Cantidad de Cuadrículas
0	75
1	103
2	121
3	54
4	30
5	13
6	2
7	1
8	0
9	1
Total	400

En esta Tabla de Distribución de Frecuencias para la variable aleatoria

“cantidad de células por cuadrícula”, se observa

- hay 75 cuadrículas que no presentan ninguna célula de levadura

- la mayor parte de las cuadrículas poseen 1 o 2 células.

- 17 cuadrículas contienen menos de 5 células de levadura.

Ejemplo

X: “cantidad de células por cuadrícula”



- La esperanza matemática o promedio de esta distribución es 1,8 células por cuadrícula .
- Se trata de un valor pequeño , considerando que la cantidad de células que podrían existir en la cuadrícula es muy grande . Esto nos lleva a pensar que estamos ante un acontecimiento relativamente raro.
- También podemos esperar que la existencia de células de levadura en un cuadrado es independiente de la existencia de este tipo de células en otros cuadrados .
- Así, este experimento origina para la variable aleatoria discreta una función de masa igual a la de Poisson.

El parámetro que caracteriza a esta distribución es

$$\mu = \lambda t = 1,8 \text{ células}, \quad \lambda = 1,8 \text{ células / cuadrícula}, \quad t = 1 \text{ cuadrícula}$$

$$P (X = 2) = (e^{-1,8} 1,8^2) / 2! = 0,2678$$

Este valor puede obtenerse empleando la tabla correspondiente , en donde se entra con los valores del parámetro y de la variable aleatoria ;1,8 y 2 respectivamente en este caso.

Esperanza y Varianza de una variable de Poisson



$$E(X) = \text{VAR}(X) = \mu = \lambda t$$

Ejemplo

El número promedio de partículas radiactivas que pasan a través de un contador durante un milisegundo en un experimento de laboratorio es 4. ¿Cuál es la probabilidad de que entren 6 partículas al contador en un milisegundo determinado?

Se trata de una distribución de Poisson de parámetro

$\mu = 4$ *y debemos calcular* $P(X = 6)$.

$$P(X = 6) = (e^{-4} 4^6) / 6! = 0,1042$$

Distribución de Poisson como límite de la Distribución Binomial



n (número de pruebas) en la Distribución Binomial es grande
 p (probabilidad de éxito) cercana a cero

*La Distribución de Poisson puede utilizarse ,
tomando el parámetro como $\mu = np$,
para aproximar Distribuciones Binomiales*

Ejemplo

En un proceso de manufactura en el cual se producen piezas de vidrio ,
ocurren defectos o burbujas. Se sabe que , en promedio una de cada mil
piezas tiene una o más burbujas . ¿Cuál es la probabilidad de que en una
muestra aleatoria de 8.000 piezas , menos de 7 de ellas tengan
burbujas ?

La variable aleatoria **X** : “ número de piezas con burbujas “ es binomial
 $n = 8000$ y $p = 0,001$

X : “ número de piezas con burbujas “ es Binomial

$$n = 8000 \text{ y } p = 0,001$$



$$P(X < 7) = P(X \leq 6) = \sum_{x=0}^6 \binom{8000}{x} 0,001^x 0,999^{n-x}$$

Dado que p se acerca a 0 y n es suficientemente grande se puede aproximar la distribución binomial a la de Poisson con parámetro $\mu = 8000 \cdot 0,001 = 8$ y obtener la probabilidad requerida del siguiente modo :

$$P(X < 7) = \sum_{x=0}^6 \left[\frac{(e^{-8} 8^x)}{x!} \right] = 0,3134$$

DISTRIBUCIÓN NORMAL.

Aproximación de una Binomial a la Normal



- Se puede demostrar (**teorema central del límite**) que una v.a. discreta con distribución binomial, $X \sim B(n, p)$ se puede aproximar mediante una distribución normal si n es suficientemente grande y p no está ni muy próximo a 0 ni a 1. Como el valor esperado y la varianza de X son respectivamente np y npq , la aproximación consiste en decir que

$$X \sim B(n, p) \text{ donde } \begin{cases} n > 30 \\ np > 5 \\ nq > 5 \end{cases} \rightarrow X \sim N(np, npq)$$

Se están aproximando probabilidades de una distribución discreta por probabilidades de una distribución continua, se debe aplicar un ***Factor de corrección por continuidad de 1/2, antes de calcular las probabilidades.*** Cuando el tamaño de la muestra es bien grande entonces el efecto de considerar el factor de corrección por continuidad es insignificante.

Fórmulas de aproximación Normal a la Binomial.

Si X es una Binomial con parámetros n y p , entonces

$$\text{i) } P(X = k) \cong P(k - .5 < X < k + .5) = P\left(\frac{k - .5 - np}{\sqrt{npq}} < Z < \frac{k + .5 - np}{\sqrt{npq}}\right)$$

$$\text{ii) } P(a < X < b) = P(a + .5 < X < b - .5) = P\left(\frac{a + .5 - np}{\sqrt{npq}} < Z < \frac{b - .5 - np}{\sqrt{npq}}\right)$$

$$\text{iii) } P(a \leq X \leq b) = P(a - .5 < X < b + .5) = P\left(\frac{a - .5 - np}{\sqrt{npq}} < Z < \frac{b + .5 - np}{\sqrt{npq}}\right)$$

Ejemplo. Durante cierta epidemia de gripe, enferma el 30% de la población. En un aula con 200 estudiantes de Medicina, ¿cuál es la probabilidad de que al menos 40 padezcan la enfermedad? Calcular la probabilidad de que haya 60 estudiantes con gripe.



- La v.a. X : el número de alumnos que padece la gripe es

$$X \sim \mathbf{B}(n = 200, p = 0,3)$$

cuya media es $\mu = 200 \times 0,3 = 60$ y cuya $\sigma = 200 \times 0,3 \times 0,7 = 42$

Realizar los cálculos con la ley binomial es muy engorroso, ya que intervienen números combinatorios de gran tamaño, y potencias muy elevadas. Por ello utilizamos la aproximación normal de X , teniendo en cuenta que se verifican las condiciones necesarias para que el error sea aceptable:

$$X \sim B(n, p) \text{ donde } \begin{cases} n = 200 > 30 \\ np = 60 > 5 \\ nq = 140 > 5 \end{cases} \rightarrow X \sim N(60, \sigma^2 = 42)$$

