

UNIDAD

6

Modelos discretos de probabilidad

Objetivos

Al finalizar la unidad, el alumno:

- distinguirá y resolverá ejercicios de cada uno de los modelos discretos: binomial, geométrico, binomial negativo, hipergeométrico y Poisson
- resolverá problemas de aplicación de variable aleatoria discreta e identificará el modelo empleado

Introducción

En la unidad anterior se analizó el concepto de variable aleatoria discreta, se mencionó su importancia y además se hizo énfasis en que las variables aleatorias no eran sólo una representación más de los eventos sino que introducen la *teoría de funciones* al estudio de las probabilidades; por tanto, se les heredan todas las propiedades y operaciones de las funciones. En conclusión, el estudio de las variables aleatorias se puede llevar a cabo de manera similar al de las funciones.

En la presente unidad se analizará una clasificación de las variables aleatorias discretas más comunes, se encontrarán su dominio, su rango y sus parámetros más comunes, como el valor esperado y la varianza.

Se comienza con el proceso de variable discreta más sencillo, llamado *proceso de Bernoulli*, que sirve de antecedente al modelo de variable discreta *binomial*. Posteriormente, se estudiará un modelo con pruebas independientes infinitas: el *modelo geométrica*; para finalizar, los experimentos aleatorios (de pruebas independientes) con el uso del *modelo binomial negativo* o de *Pascal*.

El estudio continúa dando un giro hacia los modelos cuyas pruebas en el experimento son dependientes: el *modelo hipergeométrico*, propicio en el estudio de las *técnicas de calidad*, para los casos en que se realizan muestreos aleatorios (sin reemplazo) en poblaciones divididas en artículos con y sin defectos.

Para finalizar la unidad, se estudia el *modelo de Poisson*, el cual se aplica en las líneas de espera, la teoría de inventarios, etcétera.

La forma de trabajo de esta unidad es la siguiente: en cada modelo se dan las definiciones y fórmulas correspondientes para las variables, cálculo de probabilidades, distribución de probabilidad, valor esperado y varianza y, por último, se resuelven ejemplos con base en los resultados obtenidos.

6.1 Modelo binomial

Un experimento a menudo consiste en pruebas repetidas, cada una con dos posibles resultados que se pueden etiquetar como *éxito* o *fracaso*. Por ejemplo, al lanzar una moneda, el resultado que sea objeto de estudio (cara águila) será considerado *éxito* y el otro resultado (cara sol) será *fracaso* con probabilidades p y $q = 1 - p$, respectivamente; si los ensayos que se repiten son independientes y la probabilidad de éxito permanece constante en cada ensayo, el proceso se denomina *proceso de Bernoulli*, es decir

Definición 6.1

Un experimento aleatorio se llama **de Bernoulli** cuando cumple las siguientes tres condiciones

1. El experimento consta de n (número finito) pruebas independientes.
2. Cada prueba tiene sólo dos resultados: éxito y fracaso.
3. La probabilidad de éxito en una prueba es p y la de fracaso es $q = 1 - p$ y se mantienen constantes de prueba en prueba.

Definición 6.2

A cada una de las pruebas efectuadas en un experimento de Bernoulli se les llama **ensayos de Bernoulli**.

Por **éxito en un ensayo** se entiende el *cumplimiento de la variable aleatoria*; es decir, si la variable X se define como: “cantidad de artículos defectuosos”, un éxito será un artículo defectuoso.

Se considera un conjunto de experimentos de Bernoulli. La variable que cuantifica el número X de éxitos en n experimentos de Bernoulli se denomina *variable aleatoria binomial*. La distribución de probabilidad de esta variable aleatoria define el modelo binomial:

Definición 6.3

Un experimento de Bernoulli puede convertirse en un experimento binomial si la variable aleatoria X representa la cantidad de éxitos en n ensayos de Bernoulli; es decir, si los n ensayos que se repiten son independientes, se genera el **modelo binomial** para la variable aleatoria X

$$B(k, n, p) = P(X = k) = C_k^n p^k q^{n-k}, k = 0, 1, 2, 3, \dots, n$$

A continuación se presentan algunos ejemplos en los que es posible comprobar si la variable aleatoria definida en el problema es o no binomial.

Ejemplo 1

1. Un sistema de tres radares para detectar carros a gran velocidad se instala en una carretera. Cada radar funciona de manera *independiente* con probabilidad de 0.99 de detectar un carro que viaje con gran velocidad. Se calcula la probabilidad de que un carro que viaja a gran velocidad por dicha carretera no sea detectado.

Considerando la variable aleatoria

X : “cantidad de radares que detectan el carro que viaja con gran velocidad”

se determina si este experimento es de tipo binomial.

Comprobación

Es necesario verificar que se cumplen las condiciones de un experimento binomial.

- el experimento consiste en tres ensayos, cada uno de ellos determina si el radar detecta o no al carro que viaja a gran velocidad. Por las condiciones del problema es posible observar que son independientes
- al pasar el carro a gran velocidad por un radar sólo puede ocurrir una de dos cosas: que sea o no detectado; es decir, un éxito o un fracaso
- el éxito (que sea detectado) de las condiciones del problema se conserva constante de radar en radar e igual a 0.99; de igual manera el fracaso es 0.01

A continuación se presentan dos ejemplos donde no se cumplen las condiciones de un experimento binomial (es decir, no se cumple alguna de las tres condiciones de Bernoulli).

2. En el ejemplo anterior se cambian las condiciones en la detección de los radares, de tal forma que en la detección éstos sigan siendo independientes, pero las probabilidades de detectar un carro a gran velocidad sean diferentes de radar en radar. Éste no es un experimento de Bernoulli, puesto que no cumple la condición 3: las probabilidades de éxito y fracaso son diferentes en los ensayos.
3. Si se cambia la condición de independencia de la detección de los radares, es decir, que la probabilidad de detectar del segundo radar dependa del resultado del primero y la del tercero, del segundo, no es un experimento binomial, puesto que no se cumplen las condiciones de Bernoulli (condición de ensayos independientes).

A continuación se presentan dos ejemplos en los que se puede observar la importancia de elegir los elementos de la muestra con y sin reemplazo.

4. Una urna contiene diez esferas, tres rojas y siete azules. Se extraen cuatro, una tras otra con reemplazo. Se define la variable aleatoria discreta

X : "cantidad de esferas rojas de las cuatro extraídas"

se determina si este experimento es de tipo binomial

Comprobación

Es necesario verificar *que se cumplen las condiciones* para que sea un experimento binomial.

Primero, no existe contradicción en el experimento con respecto a las cantidades, puesto que al permitirse el reemplazo se puede extraer cualquier cantidad de esferas rojas.

El experimento consiste en cuatro ensayos, en cada uno de ellos se determina si la esfera extraída es roja. Por las condiciones del problema, se puede observar que las extracciones son independientes, puesto que al volver a colocar la esfera en la urna, en la segunda extracción se tienen las mismas condiciones iniciales; es decir, el primer ensayo no influye en el segundo.

Al extraer una esfera de la urna sólo puede ocurrir que sea o no roja; es decir, un éxito o un fracaso. El éxito de que la esfera extraída sea roja se conserva constante de extracción en extracción e igual a 0.3. Por tanto, el fracaso es 0.7.

5. En el experimento anterior se cambia la condición y se extraen sólo tres, una tras otra, pero sin reemplazo; el experimento no es binomial, puesto que se altera la independencia, tal y como se mencionó en la unidad 4 sobre los eventos independientes.

De estos dos ejemplos, es posible observar que la condición con y sin reemplazo es fundamental para el experimento binomial.

Se simboliza

$P(X = k)$: "la probabilidad de que en el *experimento binomial* ocurran k éxitos de un total de n ensayos"

Esta probabilidad con frecuencia se simboliza por $B(k; n, p)$, la cual representa el *modelo binomial*, donde se quiere tener k éxitos de n ensayos con probabilidad de éxito p en cada ensayo. Ahora se presenta la fórmula para calcular las probabilidades:

Teorema 6.1

Dada una variable aleatoria binomial X , y $R_X = \{0, 1, \dots, n\}$, con éxito p y fracaso $q = 1 - p$, se cumple

$$B(k; n, p) = P(X = k) = C_k^n p^k q^{n-k}, \quad k = 0, 1, 2, 3, \dots, n$$

Demostración

Con el uso de la definición de una *variable aleatoria binomial*, se tienen n ensayos de Bernoulli independientes con probabilidad de éxito p y fracaso $q = 1 - p$, como nos interesa la probabilidad de que sean k éxitos, para cuando los primeros k ensayos sean exitosos

$$\underbrace{p \cdot p \cdot \dots \cdot p}_k \underbrace{q \cdot q \cdot \dots \cdot q}_{n-k} = p^k q^{n-k}$$

Faltaría conocer cuántos de tales productos pueden ocurrir en el experimento; de las técnicas de conteo (unidad 3) se sabe que podemos acomodar k y $n - k$ elementos iguales de

$$C_k^n = \frac{n!}{k!(n-k)!}, \text{ formas}$$

de donde se deduce

$$P(X = k) = C_k^n p^k q^{n-k}$$

En un modelo binomial la distribución de probabilidad está dada por la definición 6.4.

Definición 6.4

Se llama **distribución de probabilidad binomial** de una variable aleatoria binomial al conjunto de parejas $(k, B(k; n, p))$, para $k = 0, 1, \dots, n$.

Con el siguiente teorema se comprobará que efectivamente la definición anterior se trata de una distribución de probabilidad, es decir, la suma de todas sus probabilidades es igual a uno.

Teorema 6.2

Dada una variable aleatoria binomial X , con distribución $(k, B(k; n, p))$, para $k = 0, 1, \dots, n$, con éxito p y fracaso $q = 1 - p$, entonces

$$\sum_{k=0}^n B(k; n, p) = \sum_{k=0}^n C_k^n p^k q^{n-k} = 1$$

Demostración

La demostración se obtiene del binomio de Newton

$$(a+b)^n = \sum_{k=0}^n C_k^n a^k b^{n-k}$$

con $a = p$ y $b = q = 1 - p$. Se sustituye

$$\sum_{k=0}^n C_k^n p^k q^{n-k} = (p+q)^n = (1)^n = 1$$

Para terminar con el estudio del *modelo binomial*, se deducen las fórmulas para calcular los parámetros de *valor esperado* y la *varianza*.

Teorema 6.3

El **valor esperado** y la **varianza** de una variable aleatoria discreta X con distribución binomial, que consta de n ensayos, con éxito p y fracaso $q = 1 - p$, son

$$\begin{aligned} \text{a) } E(X) &= np \\ \text{b) } V(X) &= npq \end{aligned}$$

Demostración

- a) Con el uso de la definición de valor esperado de una variable aleatoria discreta para la **distribución binomial**, se tiene

$$E(X) = \sum_{k=0}^n k C_k^n p^k q^{n-k} = 0 C_0^n p^0 q^{n-0} + \sum_{k=1}^n k C_k^n p^k q^{n-k} = \sum_{k=1}^n k C_k^n p^k q^{n-k}$$

Al hacer el cambio de variable $k = m + 1$, cuando $k = 1 \Rightarrow m = 0$ y $k = n \Rightarrow m = n - 1$

$$E(X) = \sum_{m=0}^{n-1} (m+1) C_{m+1}^n p^{m+1} q^{n-(m+1)} = p \sum_{m=0}^{n-1} (m+1) C_{m+1}^n p^m q^{(n-1)-m}$$

De las propiedades de combinatorias, se tiene

$$(m+1) C_{m+1}^n = n C_m^{n-1}$$

donde

$$E(X) = np \sum_{m=0}^{n-1} C_m^{n-1} p^m q^{(n-1)-m}$$

Con el uso de

$$\sum_{m=0}^{n-1} C_m^{n-1} p^m q^{(n-1)-m} = (p+q)^{n-1} = 1$$

entonces

$$E(X) = np$$

- b) Con el teorema 5.2 para calcular la varianza

$$V(X) = E(X^2) - [E(X)]^2$$

$$\begin{aligned} V(X) &= \sum_{k=0}^n k^2 C_k^n p^k q^{n-k} - [E(X)]^2 = \sum_{k=0}^n (k^2 - k + k) C_k^n p^k q^{n-k} - (np)^2 \\ &= \sum_{k=0}^n k(k-1) C_k^n p^k q^{n-k} + \sum_{k=0}^n k C_k^n p^k q^{n-k} - (np)^2 \\ &= 0(0-1) C_0^n p^0 q^{n-0} + 1(1-1) C_1^n p^1 q^{n-1} + \sum_{k=2}^n k(k-1) C_k^n p^k q^{n-k} + npq \end{aligned}$$

Al hacer el cambio de variable $k = m + 2$, cuando $k = 2 \Rightarrow m = 0$ y $k = n \Rightarrow m = n - 2$

$$V(X) = \sum_{m=0}^{n-2} (m+2)(m+1)C_{m+2}^{n-2} p^{m+2} q^{n-(m+2)} = p^2 \sum_{m=0}^{n-1} (m+2)(m+1)C_{m+2}^{n-2} p^m q^{(n-2)-m}$$

De las propiedades de combinatorias, se tiene

$$(m+2)(m+1)C_{m+2}^{n-2} = n(n-1)C_m^{n-2}$$

$$V(X) = n(n-1)p^2 \sum_{m=0}^{n-2} C_m^{n-2} p^m q^{(n-2)-m} + np - (np)^2$$

como

$$\sum_{m=0}^{n-2} C_m^{n-2} p^m q^{(n-2)-m} = 1$$

entonces

$$V(X) = n(n-1)p^2 + np - (np)^2 = (np)^2 - np^2 + np - (np)^2 = np(1-p) = npq$$

Nota Para resolver problemas sobre modelos de variable discreta, se recomienda seguir los tres pasos siguientes:

- I. Definir la variable aleatoria en estudio.
- II. Identificar el modelo al que pertenece la variable definida.
- III. Aplicar las fórmulas correspondientes para el cálculo de probabilidades, valor esperado y varianza.

- Ejemplo 2
1. Un sistema de tres radares para detectar carros a gran velocidad se instala en una carretera. Cada radar funciona independientemente, con 0.99 de probabilidad de detectar un carro que viaja con gran velocidad. Considerando a la variable aleatoria X : “el número de radares que detectan al carro que viaja con gran velocidad”, se calcula
 - a) la distribución de probabilidad para X
 - b) el valor esperado y la varianza de X

Procedimiento

- I. En este ejemplo el primer punto para la solución de problemas ya se llevó a cabo, puesto que la variable ya se definió

X : “el número de radares que detectan al carro que viaja con gran velocidad”

- II. Identificación del modelo. Es de distribución binomial, puesto que en el ejemplo anterior 1, resultó que X tiene una *distribución binomial* con $R_X = \{0, 1, 2, 3\}$.
- III. Aplicación de las fórmulas. Se tiene

a) del teorema 6.1, para $p = 0.99$ y $q = 0.01$ resulta

$$P(X=0) = C_0^3 (0.99)^0 (0.01)^3 = 0.000001$$

$$P(X=1) = C_1^3 (0.99)^1 (0.01)^2 = 0.000297$$

$$P(X=2) = C_2^3 (0.99)^2 (0.01)^1 = 0.029403$$

$$P(X=3) = C_3^3 (0.99)^3 (0.01)^0 = 0.970299$$

b) del teorema 6.3, se obtiene

$$E(X) = np = 3(0.99) = 2.97$$

$$V(X) = npq = 3(0.99)(0.01) = 0.0297$$

2. Una urna contiene diez esferas, tres rojas y siete azules. Se extraen cuatro, una tras otra con reemplazo.

- a) determinar cuántas esferas de las extraídas se espera sean de color azul
- b) calcular la probabilidad de que al menos dos sean azules

Procedimiento

I. Definición de la variable

X : "cantidad de esferas azules de las cuatro extraídas"

II. Identificación del modelo. Ya se llevó a cabo, puesto que en el ejemplo 1, numeral 4, resultó que X tiene una distribución binomial con

$$R_X = \{0, 1, 2, 3, 4\}, p = \frac{7}{10} = 0.7 \text{ y } q = \frac{3}{10} = 0.3$$

III. Aplicación de las fórmulas

a) del teorema 6.3, $E(X) = np = 4(0.7) = 2.8$

b) $P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4)$ o por su complemento, resulta

$$P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)]$$

Con el teorema 6.1, para los cálculos

$$P(X = 0) = C_0^4 (0.7)^0 (0.3)^4 = 0.0081$$

$$P(X = 1) = C_1^4 (0.7)^1 (0.3)^3 = 0.0756$$

donde

$$P(X \geq 2) = 1 - [0.0081 + 0.0756] = 1 - 0.0837 = 0.9163$$

3. De una producción de tornillos 10% resulta con defectos. Si se toma una muestra de diez tornillos, se calcula la probabilidad de encontrar no más de dos tornillos defectuosos

Procedimiento

I. Definición de la variable

X : "cantidad de tornillos defectuosos en la muestra".

II. Identificación del modelo. Se tiene que 10% de la producción de tornillos es defectuosa, este porcentaje se considera invariable en los ensayos del experimento y por consiguiente los resultados serán independientes y la muestra es finita $n = 10$.

Cada que se analice un tornillo puede ocurrir sólo un caso, o es defectuoso o no, es decir, se tiene únicamente uno de los dos resultados.

El éxito $p = 0.10$ consiste en que el tornillo sea defectuoso, y el fracaso $q = 0.90$; ambos son constantes de ensayo en ensayo.

Como cumplió con las tres condiciones de Bernoulli, la variable tendrá una distribución de tipo binomial.

III. Aplicación de las fórmulas. Se tiene

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

Con el uso de la fórmula del teorema 6.1, en cada una de las probabilidades anteriores y efectuando los cálculos correspondientes, se tiene

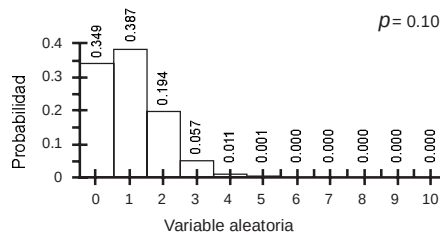
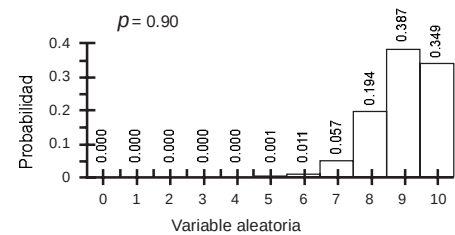
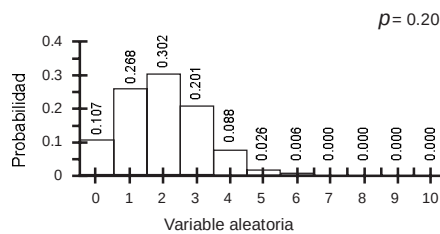
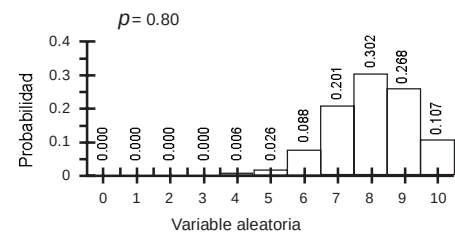
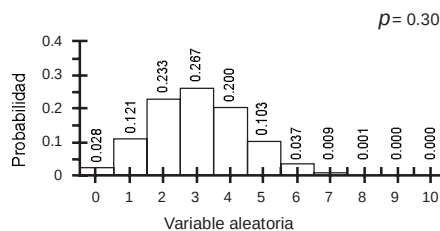
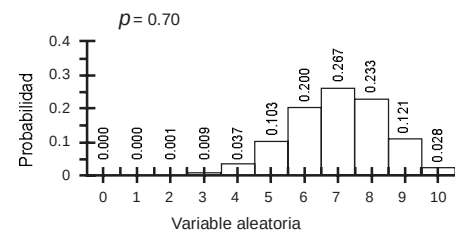
$$P(X \leq 2) = C_0^{10} (0.1)^0 (0.9)^{10} + C_1^{10} (0.1)^1 (0.9)^9 + C_2^{10} (0.1)^2 (0.9)^8 = 0.9298$$

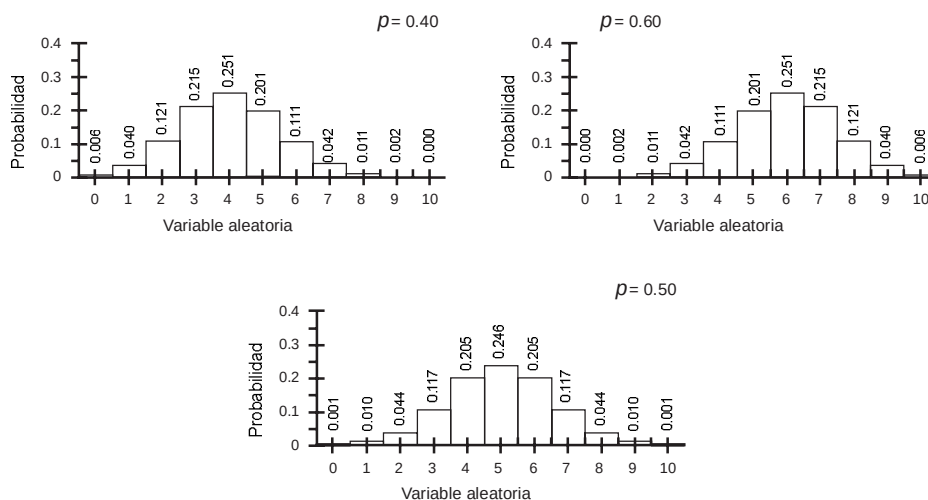
Se puede concluir que es poco probable ($1 - 0.9298 = 0.0702$) que se tengan más de dos tornillos defectuosos en una muestra de diez.

A continuación se presenta una serie de histogramas para la distribución binomial, con diez ensayos y diferentes valores de p , desde 0.1 hasta 0.9; donde podemos apreciar que la distribución binomial es más simétrica cuando el valor de p se aproxima a 0.5, mientras que para los valores más alejados se observa un sesgo en su comportamiento.

En estas gráficas se puede observar que la distribución es sesgada. Cuando $p < 0.50$ se tiene un sesgo a la derecha, mientras que en los valores de $p > 0.50$, el sesgo es a la izquierda.

Histogramas de la distribución binomial para $n = 10$





Estas gráficas ayudan a comprender más la simetría de la distribución binomial con respecto a p y $1-p$ cuando p tiende a 0.50.

Ejercicio 1

- La revisión aduanal en el aeropuerto se realiza aleatoriamente mediante un semáforo, si al pasar una persona se activa la luz roja, se revisan sus pertenencias; en caso de activarse la luz verde, el viajero pasa sin revisión. La luz roja aparece con 10% de frecuencia. Si se toma una muestra de 18 personas, calcula
 - la probabilidad de que tres o más sean revisadas
 - la probabilidad de que menos de cinco sean revisadas
 - de 100 personas, ¿cuántas se espera que sean revisadas?
- Si en general, quince de cada 100 hijos de padres alcohólicos nacen con deficiencias físicas o mentales
 - calcula la probabilidad de que de los próximos diez nacimientos (de padres alcohólicos), por lo menos dos niños resulten con deficiencias físicas o mentales
 - de los próximos 20 nacimientos (de padres alcohólicos), calcula cuántos niños se espera que no tengan deficiencias físicas o mentales
- Una máquina produce generalmente 5% de artículos defectuosos. Se toma una muestra al azar de ocho artículos. Si ésta produce más de dos objetos defectuosos, se revisará toda la producción.
 - calcula la probabilidad de que ocurra la inspección
 - calcula cuántos artículos se espera que no resulten defectuosos en una muestra de 50

4. Un examen consta de 20 preguntas de verdadero y falso. Un estudiante que no se ha preparado decide lanzar al aire una moneda para responder; anota verdadero si la cara de la moneda es sol y falso si es cara águila.
 - a) si para aprobar el examen tiene que contestar por lo menos 70% de las preguntas correctamente, calcula la probabilidad de que pase el examen
 - b) calcula la probabilidad de que conteste a lo más la mitad de las preguntas correctamente
5. De cierta población 10% sufre diabetes. Si se seleccionan 20 personas al azar
 - a) calcula la probabilidad de que al menos dos de estas personas sean diabéticas
 - b) calcula la cantidad de personas que se espera sean diabéticas

6.2 Modelo geométrico

El modelo binomial que se analizó proporciona respuesta a una gran cantidad de problemas con pruebas independientes. Sin embargo, en la práctica con gran frecuencia se encuentran otro tipo de problemas (también de ensayos independientes) que, a diferencia del modelo binomial, no tienen una cantidad finita de pruebas por realizar sino que el experimento se termina hasta que se obtiene el primer éxito. A este tipo de modelo se le llama **geométrico** (debido a su fórmula para calcular sus probabilidades).

Definición 6.5

Un experimento aleatorio se llama **geométrico** si cumple con cuatro condiciones

1. El experimento consta de ensayos independientes.
2. Cada ensayo tiene sólo dos resultados: éxito y fracaso.
3. La probabilidad de éxito en un ensayo es p y la de fracaso $q = 1 - p$, y se mantienen constantes de ensayo en ensayo.
4. El experimento termina cuando se obtiene el primer éxito en un ensayo.

Después de definir el experimento, se proporciona una definición para la variable aleatoria correspondiente.

Definición 6.6

Se llama **variable aleatoria geométrica** a la variable aleatoria discreta X , definida en un experimento geométrico, que representa la cantidad de pruebas necesarias hasta obtener el primer éxito.

A continuación se presentan algunos ejemplos de variable aleatoria geométrica.

Ejemplo 3

1. Al lanzar una moneda se define la variable aleatoria

X : "cantidad de lanzamientos hasta que resulte cara águila"

2. Si 35% de una población está a favor de un candidato, se puede definir la variable aleatoria

X : "cantidad de personas que se va a entrevistar al azar hasta obtener la primera que esté a favor del candidato"

3. Si una máquina de refrescos suministra un poco más de 200 ml por vaso y derrama 5% de refresco, se define la variable aleatoria como

X : "cantidad de vasos despachados hasta obtener uno que se derrame"

Se simboliza por $G(k; p) = P(X = k)$ a la probabilidad de que el primer éxito ocurra en el ensayo k . La fórmula para calcular las probabilidades de un modelo geométrico está dada en el siguiente teorema.

Teorema 6.4

Dada una variable aleatoria geométrica X , con éxito p y fracaso $q = 1 - p$, entonces

$$G(k; p) = P(X = k) = q^{k-1}p, k = 1, 2, 3, \dots$$

Demostración

Con el uso de la definición de variable aleatoria geométrica, se tiene que las primeras $k - 1$ pruebas son fracasos e independientes con probabilidades $q = 1 - p$. Mientras que la k -ésima prueba es el primer éxito, y también es independiente con probabilidad de éxito p . Donde

$$P(X = k) = \underbrace{q \cdot q \cdot \dots \cdot q}_{k-1 \text{ veces}} \cdot p = q^{k-1}p$$

Después de encontrar la fórmula para el cálculo de probabilidades, se define la distribución de probabilidades correspondiente.

Definición 6.7

Se llama **distribución de probabilidad** de una variable aleatoria **geométrica** X con éxito p , a las parejas $(k, G(k; p))$, para $k = 1, 2, \dots$

Nota

De la definición de variable aleatoria con distribución geométrica se debe observar que el rango de la variable, a diferencia de la binomial, comienza en uno y no termina, es decir, es infinito.

Después de definida la distribución geométrica, se verifica que la definición se refiere a una distribución de probabilidad.

Teorema 6.5

Dada una variable aleatoria geométrica X con distribución $(k, G(k; p))$ para $k = 1, 2, \dots$, con éxito p y fracaso $q = 1 - p$, entonces

$$\sum_{k=1}^{\infty} G(k; p) = \sum_{k=1}^{\infty} q^{k-1}p = 1$$

Demostración

Se obtiene con el uso de la progresión geométrica

$$1 + q + \dots + q^k = \frac{1 - q^{k+1}}{1 - q}, \text{ y el límite } \lim_{k \rightarrow \infty} q^k = 0, \text{ para } 0 \leq q < 1$$

$$\sum_{k=1}^{\infty} G(k; p) = \sum_{k=1}^{\infty} q^{k-1} p = p \lim_{N \rightarrow \infty} \sum_{k=1}^N q^{k-1} = p \lim_{N \rightarrow \infty} \frac{1-q^N}{1-q} = p \cdot \frac{1}{1-q} = p \cdot \frac{1}{p} = 1$$

Finalmente, se obtienen las fórmulas correspondientes a los cálculos del valor esperado y la varianza de una variable con distribución geométrica, esto último con el uso del siguiente teorema, el cual no se demostrará, debido a que es necesario tener conocimiento de series numéricas convergentes.

Teorema 6.6

Dada una variable aleatoria discreta X con distribución geométrica, con éxito p y fracaso $q = 1 - p$, entonces se cumple

$$E(X) = \frac{1}{p}$$

$$V(X) = \frac{1-p}{p^2}$$

En los *modelos geométricos* se presentan con frecuencia probabilidades de los tipos $P(X \leq k)$ o $P(X \geq k+1)$, por lo que es conveniente tener fórmulas adecuadas para sus cálculos.

Teorema 6.7

Dada una variable aleatoria discreta X con distribución geométrica, éxito p y fracaso $q = 1 - p$, entonces

$$P(X \leq k) = 1 - q^k, \quad P(X \geq k+1) = P(X > k) = q^k, \quad \text{para } k = 1, 2, \dots$$

Demostración

La primer fórmula, $P(X \leq k)$, se deduce de las definiciones de distribución geométrica, y la sumatoria de una progresión geométrica

$$P(X \leq k) = \sum_{i=1}^k p q^{i-1} = p \sum_{i=1}^k q^{i-1} = p \sum_{i=0}^{k-1} q^i = p \frac{1-q^k}{1-q} = 1 - q^k$$

La segunda fórmula se obtiene por el complemento de la primera

$$P(X \geq k+1) = P(X > k) = 1 - P(X \leq k) = 1 - (1 - q^k) = q^k$$

Para el cálculo de probabilidades de un *modelo geométrico* se verificarán las tres condiciones de Bernoulli o binomiales.

Ejemplo 4

1. Si 25% de una población está a favor de un candidato para las elecciones presidenciales, al momento de realizar entrevistas

- se obtiene la probabilidad de que la primer persona que esté a favor del candidato se encuentre después de la quinta persona entrevistada
- calcular cuántas personas se espera entrevistar hasta encontrar la primera que esté a favor del candidato

Procedimiento

I. Definición de la variable

X : “cantidad de personas que se va a entrevistar aleatoriamente hasta obtener la primera que esté a favor del candidato”

II. Identificación del modelo. Ya se llevó a cabo puesto que en el ejemplo 3, numeral 2, se obtuvo que X cumple con una variable geométrica con $p = 0.25$ y $q = 0.75$.

III. Aplicación de las fórmulas

a) del teorema 6.7, $P(X > 5) = q^5 = (0.75)^5 = 0.2373$

b) del teorema 6.6, $E(X) = 1/p = 1/0.25 = 4$

2. Un jugador de baloncesto acierta 80% de sus lanzamientos de tiros libres a la canasta, por partido. Se calcula la probabilidad de que en sólo uno de los siguientes cinco partidos anote su primer canasta de tiros libres después del segundo lanzamiento. Se supone que las condiciones de juego, de partido en partido, son independientes. Se define la variable aleatoria

X : “cantidad de los siguientes cinco partidos en que anota una canasta después del segundo lanzamiento”

De la condición de independencia, X tiene una distribución binomial con $n = 5$ y éxito p . Para encontrar el valor de p es necesario recordar su significado: p representa el éxito de X , es decir que en un partido el jugador anota una canasta después del segundo lanzamiento.

Para calcular el valor de p , primero se define la variable aleatoria

Y : “cantidad de lanzamientos en un partido hasta anotar su primer canasta”

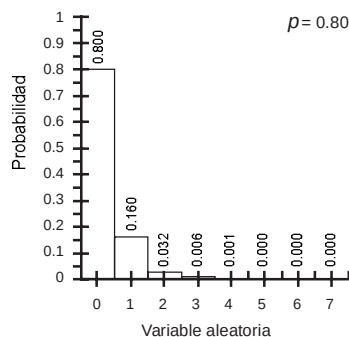
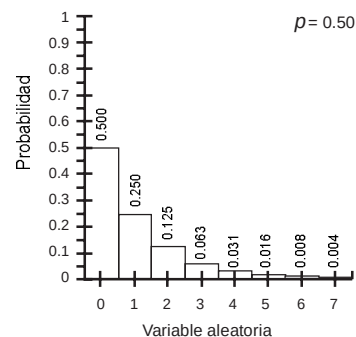
Como se puede observar, Y tiene distribución geométrica con $p_Y = 0.80$ (el subíndice se emplea para diferenciarla del éxito de X).

$$p = P(Y > 2) = q_Y^2 = (1 - p_Y)^2 = (1 - 0.80)^2 = 0.04$$

Finalmente, por definición de variable binomial

$$P(X = 1) = C_1^5 p^1 q^4 = 5(0.04)(0.96)^4 = 0.16987$$

Para conducir el estudio de la distribución geométrica, se analiza que ésta es sesgada hacia la derecha



Ejercicio 2

1. Una máquina de refrescos suministra un poco más de 20 ml por vaso y derrama 5% de los vasos despachados. Definimos a la variable aleatoria

X : "cantidad de vasos despachados hasta obtener el primero que se derramará"

Considera que la máquina despacha el líquido de manera independiente vaso con vaso y calcula la probabilidad de que el primer vaso que se derrame sea después del quinceavo.

2. Tres personas en una cafetería lanzan monedas al aire; la cara que resulte distinta pagará la cuenta. Si los tres resultados son iguales, se lanzan las monedas nuevamente hasta que resulte una distinta.
 - a) calcula la probabilidad de que se necesiten más de cuatro intentos para obtener un perdedor que pague la cuenta
 - b) determina en qué intento se espera tener al perdedor
3. Un inspector encontró que en seis de diez tiendas que visitó se presentan irregularidades. Si el inspector visita una serie de tiendas al azar, calcula la probabilidad de que
 - a) se encuentre la primera tienda con irregularidades después de revisar la cuarta tienda
 - b) determina cuántas tiendas se espera que visite para encontrar la primera con irregularidades
4. En un lote de artículos hay 3% de defectuosos. Si se toman artículos al azar, uno tras otro, hasta encontrar uno defectuoso, calcula la probabilidad de encontrar uno defectuoso después de inspeccionar cinco.
5. Se estima que 70% de una población de consumidores prefiere una marca particular de pasta de dientes, A, calcula la probabilidad de que al entrevistar a un grupo de consumidores
 - a) se tenga que entrevistar exactamente a tres personas para encontrar el primer consumidor que prefiere la marca A
 - b) se tenga que entrevistar por lo menos a diez personas para encontrar el primer consumidor que prefiere la marca A

6.3 Modelo de Pascal o binomial negativo

El modelo de Pascal es la combinación entre los modelos binomial y geométrico. En el modelo de Pascal los ensayos del experimento se realizan hasta obtener el n -ésimo éxito.

El modelo se formaliza con la siguiente definición.

Definición 6.8

Un experimento aleatorio se llama de **Pascal** o **binomial negativo**, cuando cumple las cuatro condiciones siguientes

1. El experimento consta de ensayos independientes.
2. Cada ensayo tiene sólo dos resultados; éxito y fracaso.
3. La probabilidad de éxito en un ensayo es p y la de fracaso $q = 1 - p$, y se mantienen constantes de ensayo en ensayo.
4. El experimento termina cuando ocurre el n -ésimo éxito.

Definición 6.9

A la variable aleatoria discreta X , definida en un experimento de Pascal que representa la cantidad de pruebas necesarias hasta obtener el n -ésimo éxito, se le llama **variable aleatoria de Pascal** o **binomial negativa**.

Ejemplo 5

1. Al lanzar una moneda se define a la variable aleatoria

X : “cantidad de lanzamientos hasta que resulten cinco caras águila”

2. En una población, 35% está a favor de un candidato para las elecciones presidenciales. Se define la variable aleatoria

X : “cantidad de personas que se entrevistarán al azar hasta obtener la décima que esté a favor del candidato”

3. Una máquina de refrescos suministra poco más de 20 ml por vaso y derrama 5%. Se define la variable aleatoria

X = “cantidad de vasos despachados hasta obtener el tercero derramado”

Se simboliza la probabilidad de que el n -ésimo éxito ocurra en el k -ésimo ensayo, $Pas(k; n, p) = P(X = k)$.

Teorema 6.8

Dada una variable aleatoria de Pascal X , con éxito p y fracaso $q = 1 - p$, entonces

$$Pas(k; n, p) = P(X = k) = C_{n-1}^{k-1} p^n q^{k-n}, k = n, n+1, n+2, n+3, \dots$$

Demostración

Con la definición de variable aleatoria de Pascal se tiene que en las primeras $k-1$ pruebas hay $n-1$ éxitos y $k-n$ fracasos, mientras que la k -ésima prueba es el n -ésimo éxito, todas ellas son independientes con probabilidad de éxito p y fracaso $q = 1 - p$. Del modelo binomial, se sabe que las primeras $k-1$ pruebas pueden ocurrir de

$$C_{n-1}^{k-1} p^{n-1} q^{k-n}$$

En vista de que la k -ésima prueba debe ser éxito, se tiene

$$Pas(k; n, p) = C_{n-1}^{k-1} p^{n-1} q^{k-n} p = C_{n-1}^{k-1} p^n q^{k-n}$$

Definición 6.10

Se llama **función de probabilidad de Pascal o binomial negativa** a

$$p(x) = \begin{cases} \text{Pas}(k; n, p) & x = n, n+1, n+2, \dots \\ 0, & x \neq n, n+1, n+2, \dots \end{cases}$$

y a las parejas correspondientes $(k, p(k))$, para $k = n, n+1, n+2, \dots$, se les llama **distribución de probabilidad de una variable aleatoria binomial negativa**.

A continuación, se enuncia, sin demostración, el teorema que muestra que efectivamente las parejas anteriores se refieren a una distribución de probabilidad.

Teorema 6.9

Dada una variable aleatoria discreta X con distribución binomial negativa $(k, p(k))$, para $k = n, n+1, n+2, \dots$, con éxito p y fracaso $q = 1 - p$, entonces

$$\sum_{k=n}^{\infty} \text{Pas}(k; n, p) = \sum_{k=n}^{\infty} C_{n-1}^{k-1} p^n q^{k-n} = 1$$

De forma similar, se formulará el teorema que muestre las fórmulas para calcular el valor esperado y la varianza de una distribución binomial negativa.

La demostración del teorema también se omitirá dada su extensión.

Teorema 6.10

Dada una variable aleatoria discreta X con distribución binomial negativa $(k, p(k))$, para $k = n, n+1, n+2, \dots$, con éxito p y fracaso $q = 1 - p$, entonces

$$E(X) = \frac{n}{p}$$

$$V(X) = \frac{n(1-p)}{p^2}$$

Ejemplo 6 1. En una población, 35% está a favor de un candidato para las elecciones presidenciales

- se calcula la probabilidad de que la tercer persona que esté a favor del candidato sea la quinta persona entrevistada
- se calcula cuántas personas se espera entrevistar para encontrar la tercera que esté a favor del candidato

Se define la variable aleatoria

X : "cantidad de personas que se va a entrevistar al azar hasta obtener la tercera que esté a favor del candidato"

X tiene distribución de Pascal; por tanto, de los teoremas 6.8 y 6.10, se tiene

- $\text{Pas}(5; 3, 0.35) = C_2^4 p^3 q^2 = 6(0.35)^3(0.65)^2 = 0.1087$
- $E(X) = n/p = 3/0.35 = 8.57$

2. Una máquina de refrescos suministra un poco más de 20 ml por vaso y derrama 5%. Se calcula la probabilidad de que el segundo vaso derramado sea el décimo despachado.

Se puede definir la variable aleatoria

X : "cantidad de vasos despachados hasta obtener el segundo derramado"

$$P_{X=10}(10; 2, 0.05) = C_1^9 p^2 q^8 = 9(0.05)^2 (0.95)^8 = 0.0149$$

Ejercicio 3

- Un contador encontró que nueve de diez auditorías a compañías contienen errores importantes. Si el contador revisa la contabilidad de una serie de compañías, calcula la probabilidad de que:
 - la tercera contabilidad con errores sustanciales sea la octava revisada
 - la segunda contabilidad con errores importantes se encuentre después de revisar la tercera
- Un explorador perforará una serie de pozos petroleros en cierta área hasta encontrar uno productivo. La probabilidad de que tenga éxito es 0.2, calcula la probabilidad de que el segundo pozo productivo se encuentre hasta el décimo pozo perforado.
- De los aspirantes para cierto trabajo industrial 30% tiene entrenamiento avanzado en programación. Los aspirantes son entrevistados uno tras otro y seleccionados al azar. Si una empresa necesita tres aspirantes con un entrenamiento avanzado en programación, calcula la probabilidad de encontrar el tercer aspirante con un entrenamiento avanzado en programación hasta la veinteava entrevista.
- Se sabe que una moneda está cargada de forma tal que, la probabilidad de que salga cara águila es cuatro veces la de que salga cara sol. Si la moneda se lanza varias veces, calcula la probabilidad de que se necesiten menos de cinco lanzamientos para obtener la segunda cara águila y cuántos lanzamientos se espera realizar para obtener la tercera cara sol.

6.4 Modelo hipergeométrico

Los dos modelos estudiados hasta ahora se refieren a pruebas independientes, pero, ¿qué pasa cuando las pruebas de los experimentos no son independientes. Por ejemplo, en una empresa es necesario efectuar chequeos constantes de la producción con el fin de llevar un buen control de calidad. Al realizarse el muestreo, éste tendrá que hacerse sin reemplazo; de este modo, se determina que las pruebas son *dependientes*. Por tanto, no es posible aplicar ninguno de los modelos estudiados. El problema anterior se soluciona con una nueva variable aleatoria a la que se llama *variable aleatoria hipergeométrica*.

Un modelo probabilístico será de tipo hipergeométrico cuando los experimentos que se realizan con respecto a un evento E son tales, que *sus pruebas no son independientes*. En estos modelos se consideran lotes de artículos, los cuales están constituidos de elementos divididos en dos clases. El experimento consiste en elegir una muestra del lote

sin *reemplazo* y calcular las probabilidades cuando sus elementos pertenezcan a una de las clases. Para formalizar el modelo, se tiene la siguiente definición.

Definición 6.11

Un experimento aleatorio se llama **hipergeométrico** si cumple las siguientes tres condiciones:

1. El experimento se realiza considerando un lote de tamaño N , en el cual sus elementos están divididos en dos clases de tamaños m y $N - m$.
2. Se toma una muestra de tamaño n , sin *reemplazo* del lote.
3. Se calculan las probabilidades de que k elementos de una de las clases estén en la muestra de tamaño n .

A las clases se les llama *éxitos* y *fracasos*, para conservar la terminología de los modelos anteriores.

Al introducir un modelo nuevo es necesario nombrar las variables aleatorias que sean necesarias para su estudio.

Definición 6.12

La variable aleatoria discreta X definida en un experimento *hipergeométrico* que representa a la cantidad de elementos que se encuentran en la muestra perteneciente a la clase de *éxitos*, se llama **variable aleatoria hipergeométrica**.

A continuación se presentan dos ejemplos de variables aleatorias hipergeométricas

Ejemplo 7

1. Una urna contiene quince esferas, cinco rojas y diez azules. Se toma una muestra sin *reemplazo* de cuatro esferas. Es posible definir la variable aleatoria

X : "cantidad de esferas azules de la muestra"

2. En un lote de 20 autos usados se tienen cinco descompuestos. Se toma una muestra sin *reemplazo* de tres autos. Es posible definir la variable aleatoria

X : "cantidad de carros descompuestos en la muestra"

Se simboliza por $H(k; N, n, m) = P(X = k)$ la probabilidad de que existan k *éxitos* en la muestra de tamaño n , tomada sin *reemplazo* de una población constituida únicamente de dos *dases* (*éxitos* y *fracasos*), y de tamaño N en la que se encuentran m elementos de la *dase* de *éxitos*.

A continuación se presenta una fórmula para calcular las probabilidades de *variables aleatorias hipergeométricas*.

Teorema 6.11

Dada una variable aleatoria hipergeométrica X , con m *éxitos* en una población de tamaño N , de la cual se elige una muestra al azar de tamaño n , entonces

$$H(k; N, n, m) = P(X = k) = \frac{C_k^m C_{n-k}^{N-m}}{C_n^N}, \text{ máx } \{n + m - N, 0\} \leq k \leq \text{mín } \{n, m\}$$

Demostración

Para obtener la fórmula, se emplea la definición clásica de probabilidad.

Para la cantidad de elementos del espacio muestral se tiene que una muestra de tamaño n se puede tomar sin reemplazo de un lote de tamaño N de C_n^N maneras.

Igualmente, la toma de k elementos de la clase de éxitos se puede realizar de C_k^m maneras, y finalmente los restantes $n - k$ elementos de la muestra se toman de la clase de fracasos de C_{n-k}^{N-m} maneras. Por tanto, del principio de multiplicación, la muestra que contenga k éxitos y $n - k$ fracasos se puede obtener de $C_k^m C_{n-k}^{N-m}$ maneras.

Con la definición clásica de probabilidad y dividiendo ambos resultados, se tiene

$$H(k; N, n, m) = P(X = k) = \frac{C_k^m C_{n-k}^{N-m}}{C_n^N}$$

Para concluir la demostración falta verificar que k sólo puede tomar valores en el rango $\max\{n + m - N, 0\} \leq k \leq \min\{n, m\}$.

La acotación anterior es válida puesto que k no puede ser mayor al tamaño de la muestra n ni tampoco mayor a la cantidad de elementos de la clase de los éxitos m . Por tanto, se concluye $k \leq \min\{n, m\}$.

La acotación siguiente se obtiene puesto que k no puede ser negativo ni menor a cero. Cuando $n > N - m$, k no puede ser menor a la cantidad $n - (N - m) = n + m - N$. Por tanto, se concluye $\max\{n + m - N, 0\} \leq k$.

Al combinar las dos acotaciones anteriores queda demostrado

$$\max\{n + m - N, 0\} \leq k \leq \min\{n, m\}$$

Después de encontrar la fórmula para el cálculo de probabilidades de las variables aleatorias hipergeométricas, se definirá su distribución de probabilidades.

Definición 6.13

Dada una población de tamaño N con m éxitos y de la cual se toma una muestra de tamaño n sin reemplazo, se llama **distribución de probabilidad de una variable aleatoria hipergeométrica** X a las parejas $(k, H(k; N, n, m))$, donde

$$\max\{n + m - N, 0\} \leq k \leq \min\{n, m\}$$

Nota

De la definición de variable aleatoria con **distribución hipergeométrica**, se debe observar que el rango de la variable no necesariamente inicia en cero o en uno.

Después de definir a la distribución hipergeométrica para comprobar si se trata de una distribución de probabilidad, se analizará el siguiente teorema. Su demostración se realiza con base en las combinatorias y el principio de multiplicación; aquí se omitirá por no tener mayor trascendencia.

Teorema 6.12

Dada una variable aleatoria hipergeométrica X con distribución $(k, G(k; p))$ con m éxitos, en un lote de tamaño N en el cual se elige una muestra sin reemplazo de tamaño n , entonces

$$\sum_{k=\max\{n+m-N, 0\}}^{\min\{n, m\}} H(k; N, n, m) = \sum_{k=\max\{n+m-N, 0\}}^{\min\{n, m\}} \left[\frac{C_k^m C_{n-k}^{N-m}}{C_n^N} \right] = 1$$

Finalmente, se presentan las fórmulas para calcular el valor esperado y la varianza de una variable aleatoria con *distribución hipergeométrica*, las cuales se encuentran en el siguiente teorema, donde también se omitirá su demostración.

Teorema 6.13

Dada una variable aleatoria hipergeométrica X con distribución $(k, G(k; p))$ y con m éxitos, en un lote de tamaño N en el cual se elige una muestra sin reemplazo, de tamaño n , entonces

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

$$V(X) = n \left(\frac{m}{N} \right) \left(1 - \frac{m}{N} \right) \frac{N-n}{N-1}$$

En la solución de problemas, a diferencia de las otras dos distribuciones, es más sencillo identificar a los modelos hipergeométricos por la condición de la toma sin reemplazo. Pero los pasos a seguir en la solución de problemas son los mismos: *definición de la variable, identificación y aplicación de fórmulas para los cálculos*

Ejemplo 8

1. Una caja contiene 20 discos duros para computadora, colocados en forma vertical y sin encimarse. Se supone que hay tres defectuosos; si se toman al azar cuatro de ellos, se calcula la distribución de probabilidad para

X : "cantidad de defectuosos en la muestra".

Procedimiento

- I. Definición de la variable. X ya está delimitada.
- II. Identificación del modelo. La muestra se toma sin reemplazo y las clases en que se divide el lote de discos son dos buenos y defectuosos. Por tanto, X tiene distribución de tipo hipergeométrico con $N = 20$ y cantidad de discos defectuosos $m = 3$. La muestra elegida es de tamaño $n = 4$.
- III. Aplicación de las fórmulas. Se tiene
Para el rango de X , se toma en cuenta

$$\max\{n+m-N, 0\} \leq k \leq \min\{n, m\}$$

Por tanto,

$$\max\{4+3-20, 0\} \leq k \leq \min\{4, 3\} \text{ esto es } 0 \leq k \leq 3$$

Las probabilidades se calculan con los resultados del teorema 6.11

$$P(X=0) = \frac{C_0^3 C_4^{17}}{C_4^{20}} = \frac{2\,380}{4\,845} \approx 0.4912, P(X=1) = \frac{C_1^3 C_3^{17}}{C_4^{20}} = \frac{2\,040}{4\,845} \approx 0.4211$$

$$P(X=2) = \frac{C_2^3 C_2^{17}}{C_4^{20}} = \frac{408}{4\,845} \approx 0.0842, P(X=3) = \frac{C_3^3 C_1^{17}}{C_4^{20}} = \frac{17}{4\,845} \approx 0.0035$$

2. En un lote de diez componentes electrónicos en buen estado se agregan tres defectuosos. Una persona compra cuatro de tales componentes para reparar televisores, se calcula la probabilidad de que la persona tenga que regresar a reclamar al vendedor por haber obtenido componentes defectuosos.

Procedimiento

I. Definición de la variable.

X : "cantidad de componentes defectuosos en la muestra"

- II. Identificación del modelo. Por las condiciones del problema se deduce que la muestra se tomó sin reemplazo; además de que el tamaño del lote es finito e igual a trece y sólo se tienen dos clases de componentes, buenos y defectuosos. De esto, se deduce que X es una variable hipergeométrica con $N = 13$, $n = 4$, $m = 3$.
- III. Aplicación de las fórmulas. Por las condiciones del problema, se sabe que la persona reclamará si un componente resulta defectuoso. Por tanto, la probabilidad que se debe calcular es

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{C_0^3 C_4^{10}}{C_4^{13}} \approx 1 - 0.2937 = 0.7063$$

Este resultado indica que probablemente el comprador regresará a reclamar.

3. Una de las máquinas para elaborar tornillos milimétricos se descompuso, por lo que una gran cantidad de tornillos resultó defectuosa. Para tratar de evitar pérdidas, en cada caja de 30 tornillos se colocan cinco defectuosos (25 sin defectos). El vendedor de tornillos comienza a recibir reclamos debido a las piezas defectuosas y decide cambiar de proveedor si al inspeccionar aleatoriamente seis tornillos de la siguiente caja resultan dos o más defectuosos. Se calcula la probabilidad de que el vendedor cambie de proveedor.

Procedimiento

I. Definición de la variable

X : "cantidad de defectuosos en la selección de seis tornillos"

- II. Identificación del modelo. Por las condiciones del problema, se deduce que la muestra se tomó sin reemplazo; además de que el tamaño del lote es finito e igual a 30 y sólo se tienen dos clases de componentes, con y sin defectos. Por tanto, X es una variable hipergeométrica, con $N = 30$, $n = 6$ y $m = 5$.
- III. Aplicación de las fórmulas. Se calcula la probabilidad de que en una caja se encuentren dos o más defectuosos en la inspección aleatoria de seis de ellos

$$P(X \geq 2) = 1 - [P(X = 0) + P(X = 1)] = 1 - \left[\frac{C_0^5 C_6^{25}}{C_6^{30}} + \frac{C_1^5 C_5^{25}}{C_6^{30}} \right] = 0.25435$$

es la probabilidad de que en una toma aleatoria de seis tornillos de una caja resulten dos o más defectuosos

Ejercicio 4

1. Supón que un radioreceptor contiene seis transistores, de los cuales dos son defectuosos. Se prueban tres transistores tomados al azar. Dada Y = "cantidad de defectuosos encontrados", calcula la distribución de probabilidad para Y .

2. En un lote de diez proyectiles se disparan cuatro al azar. Si el lote contiene cinco proyectiles que no disparan
 - a) calcula la probabilidad de que los cuatro disparen
 - b) calcula cuántos de los cuatro se espera que disparen
3. Para hacer un reporte de control de calidad sobre la fabricación de videos, de un lote de 25 se toma una muestra al azar de cinco de ellos y se prueban, en caso de que no se encuentren elementos defectuosos, el reporte se determina como satisfactorio. Calcula la probabilidad de que el reporte resulte satisfactorio si en el lote se encuentran cuatro videos defectuosos
4. En la aduana de un aeropuerto, debido a la gran afluencia de pasajeros, sólo se revisa a 10% de ellos a la salida. Si de un grupo de 20 turistas, doce tienen compras muy por arriba de la cantidad permitida, calcula la probabilidad de que dos personas revisadas tengan que pagar los impuestos correspondientes por exceso de compras permitidas
5. Se toman sin reemplazo ocho objetos de un lote con quince sin defectos y seis con defectos
 - a) calcula la probabilidad de que se encuentren dos defectuosos entre los ocho objetos de la muestra
 - b) calcula cuántos se espera que no tengan defectos

6.5 Modelo de Poisson

El último de los modelos probabilísticos discreto que se analizará es el *modelo de Poisson*.¹ Este modelo estudia los experimentos cuyos resultados tienen lugar en intervalos continuos,² de tiempo, áreas, volúmenes, etc. Antes de seguir, cabe mencionar que el modelo de Poisson es de variable aleatoria discreta, puesto que en sus experimentos sólo *interesa la cantidad de resultados* que pueden ocurrir en un intervalo (de los antes mencionados), *mas no la continuidad del intervalo*.

El modelo de Poisson tiene muchas aplicaciones: se emplea generalmente cuando se desea optimizar los tiempos, tanto de espera como de servicio; a este tipo de problemas se les llama *líneas de espera* o *teoría de colas*.

La formalización del modelo de Poisson, desde nuestro punto de vista, es una de las más complicadas (de los modelos discretos), ya que hace referencia a la *teoría infinitesimal*, por lo que se omitirán algunas de sus demostraciones.

Nota Para ejemplificar la definición de experimento de Poisson al hablar de intervalo, se hará referencia al tiempo (tomando en cuenta que en lugar de tiempo se podría tratar de un área, un volumen, etcétera).

¹ En honor al matemático francés Siméon-Denis Poisson, quien nació en Pithiviers, en 1781, y murió en París, en 1840. Fue uno de los creadores de la física matemática y autor de una serie de trabajos sobre mecánica celeste, elasticidad, capilaridad, cálculo de probabilidades y magnetismo.

² Debido a los intervalos continuos en los que ocurren los modelos de Poisson, éstos tienen estrecha relación con los modelos continuos de tipo exponencial; esto se analizará en la unidad 8.

Definición 6.14

Un **experimento de Poisson** debe cumplir las siguientes tres condiciones:

1. Los resultados de intervalos que no tienen puntos en común son independientes. Esto es, los resultados que ocurren en (t_1, t_2) son independientes de los que transcurran en el intervalo (t_3, t_4) , cuando los intervalos son disjuntos. Se dice que el experimento de Poisson, en su ejecución no tiene memoria.
2. La probabilidad de que un resultado ocurra en un intervalo de tiempo mucho muy pequeño $(t, t + \Delta t)$ es una cantidad de orden Δt . Esto es, la probabilidad de obtener exactamente un resultado en un intervalo pequeño es proporcional a la longitud del intervalo.
3. La probabilidad de que ocurra más de un resultado en el transcurso del intervalo $(t, t + \Delta t)$ es una cantidad mucho más pequeña que Δt . Esto significa que la probabilidad de obtener dos o más resultados en un intervalo pequeño es mínima.

De acuerdo con la metodología que se ha adoptado, se pasa a la definición de la variable aleatoria correspondiente, y los experimentos o procesos de Poisson.

Definición 6.15

A la variable aleatoria X definida en un experimento de Poisson que representa la cantidad de resultados que ocurren en el intervalo de tiempo (t_0, t) , se le llama **variable aleatoria de Poisson**.

En estas condiciones resulta que X es discreta con valores $0, 1, 2, 3, \dots$. Los intervalos dependen del experimento y pueden ser:

- un minuto, un día, una semana, un año, etcétera.
- un metro cuadrado o cúbico, una hectárea, etcétera.

A continuación se presentan algunos ejemplos de experimentos aleatorios que se consideran dentro de un modelo de Poisson.

Ejemplo 9

1. La cantidad de llamadas telefónicas a un conmutador en un intervalo de cinco minutos.
2. La cantidad de accidentes automotores mensuales en un cruce determinado.
3. La cantidad de carros que llegan a un estacionamiento en una hora determinada.
4. El número de partículas que pasan a través de un contador en un milisegundo.
5. La cantidad de errores de captura por página en un documento.
6. Cantidad de árboles infectados por ciertos gusanos en un área determinada.
7. Llegadas de clientes a una tienda durante un determinado intervalo de tiempo.

Se simboliza por $P(k; \lambda t) = P(X = k)$: “la probabilidad de que en el experimento de Poisson ocurran k resultados en un intervalo (t_0, t) ” (donde λ es un parámetro que será definido al final del teorema 6.15).

En el siguiente teorema se proporciona la fórmula para calcular probabilidades de modelos de Poisson; sin embargo, debido a su complejidad no se hará la demostración.

Teorema 6.14

Dada X como una variable aleatoria de **Poisson** en el intervalo (t_0, t) y $R_X = \{0, 1, 2, \dots\}$, (representando por t la longitud del intervalo (t_0, t)), entonces

$$P(k; \lambda t) = P(X = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad k = 0, 1, 2, \dots$$

De acuerdo con la metodología adoptada, a continuación se define la distribución de probabilidad correspondiente.

Definición 6.16

Se llama **distribución de probabilidad de Poisson** a las parejas $(k, P(k; \lambda t))$, para k igual a $0, 1, 2, 3, \dots$

En el siguiente y último teorema de la unidad se verifica que efectivamente la definición anterior se refiere a una distribución de probabilidad. Además se deducen las fórmulas correspondientes al valor esperado y la varianza de la variable.

Teorema 6.15

Dada X como una **variable aleatoria de Poisson** en un intervalo de longitud t y $R_X = \{0, 1, 2, \dots\}$, con parámetro λ , entonces

$$\begin{aligned}\sum_{k=0}^{\infty} P(k; \lambda t) &= 1 \\ \mu = E(X) &= \lambda t \\ \sigma^2 = V(X) &= \lambda t\end{aligned}$$

Demostración

La sumatoria se deduce de manera inmediata de la serie

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

puesto que

$$\sum_{k=0}^{\infty} P(k; \lambda t) = \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = (e^{-\lambda t}) e^{\lambda t} = 1$$

Para el valor esperado se empleará la serie

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

y el cambio de variable $k - 1 = m$.

$$E(X) = \sum_{k=0}^{\infty} k \left[\frac{(\lambda t)^k}{k!} e^{-\lambda t} \right] = \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} e^{-\lambda t} = \sum_{m=0}^{\infty} \frac{(\lambda t)^{m+1}}{m!} e^{-\lambda t} = \lambda t e^{-\lambda t} \sum_{m=0}^{\infty} \frac{(\lambda t)^m}{m!} = \lambda t e^{-\lambda t} (e^{\lambda t}) = \lambda t$$

Para la varianza se empleará el teorema 5.2

$$V(X) = E(X^2) - E^2(X) = E(X(X-1)) + E(X) - E^2(X)$$

Calculando $E(X(X-1))$ de la misma forma en que se realizó en el valor esperado

$$\begin{aligned}E(X(X-1)) &= \sum_{k=0}^{\infty} k(k-1) \left[\frac{(\lambda t)^k}{k!} e^{-\lambda t} \right] = \sum_{k=2}^{\infty} \frac{(\lambda t)^k}{(k-2)!} e^{-\lambda t} = \sum_{m=0}^{\infty} \frac{(\lambda t)^{m+2}}{m!} e^{-\lambda t} = \\ &= (\lambda t)^2 e^{-\lambda t} \sum_{m=0}^{\infty} \frac{(\lambda t)^m}{m!} = (\lambda t)^2 e^{-\lambda t} (e^{\lambda t}) = (\lambda t)^2\end{aligned}$$

Por consiguiente

$$V(X) = E(X(X-1)) + E(X) - E^2(X) = (\lambda t)^2 + (\lambda t) - (\lambda t)^2 = \lambda t$$

En el teorema 6.14 se presentó el parámetro λ , el cual se puede interpretar ahora, puesto que en el teorema 6.15 se demostró que $E(X) = \lambda t$; por tanto,

$$\lambda = \frac{E(X)}{t}$$

representa la razón esperada de resultados en el intervalo de estudio.

En caso de que $t = 1$ (una hora, un día, un metro, etc.), la fórmula anterior se reduce a $\lambda = E(X)$, y se emplea la fórmula simplificada para el cálculo de probabilidades

$$P(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Ejemplo 10

1. En una tienda los clientes llegan al mostrador conforme una distribución de Poisson con un promedio de diez cada hora. En una hora dada, se calcula la probabilidad de que lleguen al menos cinco clientes.

Procedimiento

- I. Definición de la variable

X : "cantidad de clientes que llegan a la tienda"

- II. Clasificación del modelo. El promedio es de diez clientes cada hora

$$\lambda = 10 \frac{\text{clientes}}{\text{hora}}$$

en un intervalo de una hora dada, es decir, $t = 1\text{h}$

- III. Aplicación de las fórmulas. Se emplea

$$P(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

con $k \geq 5$ y se calcula

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - [P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)] =$$

$$1 - \left[\frac{e^{-10}(10)^0}{0!} + \frac{e^{-10}(10)^1}{1!} + \frac{e^{-10}(10)^2}{2!} + \frac{e^{-10}(10)^3}{3!} + \frac{e^{-10}(10)^4}{4!} \right] = 1 - 0.0293 = 0.9707$$

La probabilidad es bastante grande, puesto que al considerar un valor esperado de diez clientes será muy probable que cinco o más clientes lleguen en el transcurso de una hora (ver los histogramas con diferentes valores de $E(X)$ al final del ejercicio siguiente).

2. Al revisar la calidad en el pulido de un lente, cierta compañía acostumbra determinar el número de manchas en la superficie considerando el lente defectuoso si tres o más de tales manchas, asperezas y otro tipo de defectos aparecen en él. Si el promedio es dos defectos por cm^2 , calcula la probabilidad de que un lente de cuatro cm^2 no sea considerado defectuoso.

Procedimiento

I. Definición de la variable aleatoria

X : "cantidad de defectos que aparecen en el lente".

II. Identificación del modelo. El promedio es dos defectos por cm^2 ; es decir,

$$\lambda = 2 \frac{\text{defectos}}{\text{cm}^2}$$

Para que un lente de 4 cm^2 sea revisado, se tiene que $t = 4 \text{ cm}^2$.

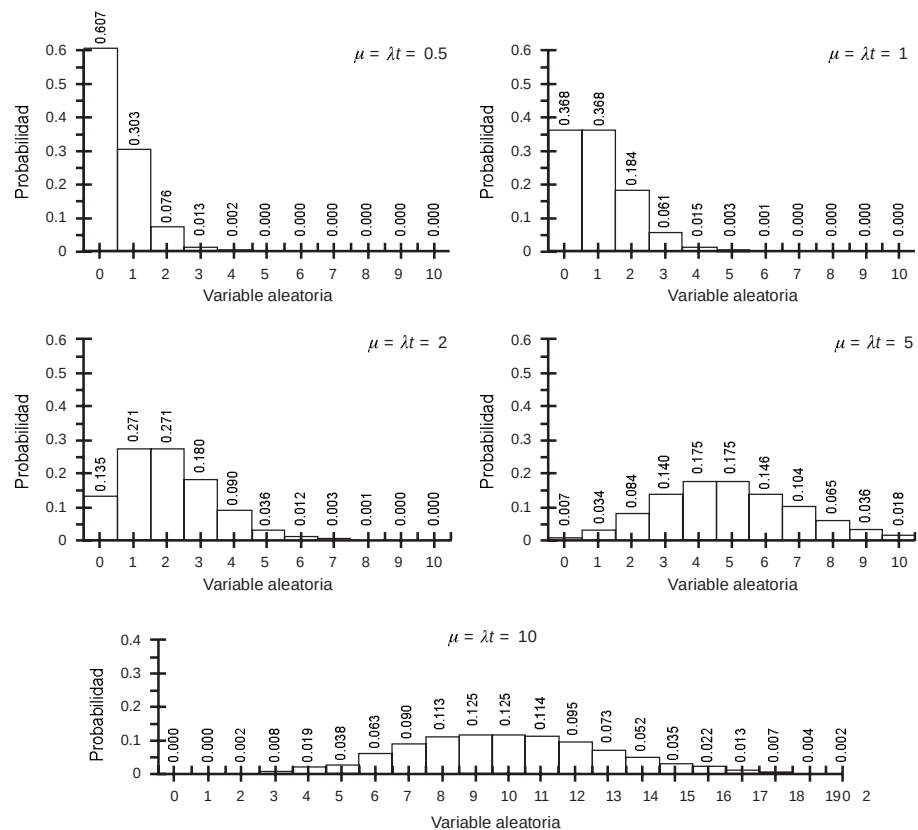
III. Aplicación de las fórmulas. Se tiene, por tanto $E(X) = \lambda t = 8$ defectos.

Para que un lente no sea considerado defectuoso debe tener menos de tres defectos. Por tanto, la probabilidad que se debe calcular es que un lente de 4 cm^2 tenga menos de tres defectos (es decir, esté en buen estado)

$$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) =$$

$$\frac{e^{-8}(8)^0}{0!} + \frac{e^{-8}(8)^1}{1!} + \frac{e^{-8}(8)^2}{2!} = 0.0138$$

A continuación se presentan algunos histogramas para la distribución de Poisson. En ellos se puede apreciar que la distribución de probabilidades se concentra alrededor del valor esperado. Es decir, con valores esperados pequeños, la distribución de probabilidad se concentra en los puntos iniciales; posteriormente, las probabilidades se aproximan a cero. En los histogramas de abajo se aprecia que, al aumentar el valor de μ , la distribución se aproxima a un modelo simétrico:



Ejercicio 5

1. Una secretaria promedia dos errores al escribir una página. Si los errores son independientes y siguen un proceso de Poisson, calcula la probabilidad de que cometa uno o más errores en la siguiente página que escriba.
2. Si el número de coches que llegan a un estacionamiento es de ocho cada hora y su llegada sigue el proceso de Poisson, calcula la probabilidad de que en un periodo de diez minutos lleguen al estacionamiento (comenta el resultado obtenido)
 - a) entre tres y seis automóviles
 - b) más de dos automóviles
3. Al revisar la calidad en el pulido de un lente, cierta compañía acostumbra determinar el número de manchas en la superficie, considerando el lente defectuoso si tres o más de tales manchas, asperezas y otro tipo de defectos aparecen en él. Si el promedio es de dos defectos por cm^2 , con distribución de Poisson
 - a) calcula la probabilidad de que un lente de 1 cm^2 no sea considerado defectuoso
 - b) calcula la probabilidad de que un lente redondo con un diámetro de 1 cm no se le catalogue como defectuoso
4. Desde 1996, el cierre de empresas por problemas financieros ha ocurrido, en promedio, a razón de 5.7 cierres por año. Supón que el número de cierres por año tiene distribución de Poisson, calcula la probabilidad de que ninguna empresa cierre durante un periodo de cuatro meses.
5. Supón que una cajera de un banco atiende en promedio a 4.5 clientes por cada diez minutos y que la cantidad de personas atendidas sigue un proceso de Poisson, calcula la probabilidad de que una cajera atienda a sólo dos clientes en el transcurso de los siguientes diez minutos.

Ejercicios propuestos

1. La probabilidad de que un motor, recién ajustado, tire aceite en los primeros 100 km por los retenes es de 0.05. Si diez automóviles se ajustan en un taller mecánico
 - a) calcula la probabilidad de que por lo menos dos tiren aceite por los retenes
 - b) de los siguientes 200 automóviles que se ajustaron en dicho taller, calcula cuántos se espera que tiren aceite por los retenes
2. Según las estadísticas de una ciudad, en cierta zona se cometen en promedio diez asaltos diarios a conductores de autos. Si los asaltos son independientes y se apegan a un proceso de Poisson
 - a) calcula la probabilidad de que en un día se cometan más de diez asaltos
 - b) calcula la probabilidad de que entre las 6:00 y 12:00 AM no se cometan asaltos
3. La probabilidad de que un estudiante de aviación apruebe el examen escrito para obtener su licencia de piloto es 0.6, calcula la probabilidad de que apruebe el examen en el tercer intento.

4. Si el costo de pasaje por persona en el transporte público es \$2.50 y cada vehículo transporta en promedio doce pasajeros cada 30 minutos, suponiendo que la cantidad de personas transportadas sigue una distribución de Poisson
 - a) calcula el ingreso esperado por día de trabajo de un chofer (un día de trabajo equivale a diez horas), si invierte 200 pesos diarios en gasolina
 - b) calcula la probabilidad de que en un intervalo de 30 minutos, transporte a lo más la mitad del promedio dado anteriormente
5. Una caja contiene cuatro naranjas y dos manzanas. Se toman tres frutas sin reemplazo. Si X es la variable aleatoria definida como el número de naranjas que se tomaron
 - a) calcula la probabilidad de que $P(X \geq 2)$
 - b) calcula la probabilidad anterior si se permite el reemplazo
6. En un almacén los clientes llegan al mostrador de caja en promedio de siete por hora, de acuerdo una distribución de Poisson. En una hora dada, calcula la probabilidad de que
 - a) no lleguen más de tres clientes
 - b) lleguen exactamente cinco clientes
7. En una población 40% es fumador. Si se toma una muestra de 20 personas al azar
 - a) calcula la probabilidad de que diez sean fumadores
 - b) calcula la probabilidad de que más de siete sean fumadores
8. La probabilidad de que un cliente acuda al mostrador de una tienda de abarrotes en cualquier periodo de un segundo es 0.1. Supón que los clientes llegan de manera aleatoria y, por tanto, las llegadas en cada intervalo de un segundo son independientes
 - a) calcula la probabilidad de que la primera llegada ocurra durante el tercer intervalo de un segundo
 - b) calcula la probabilidad de que la primera llegada ocurra después del tercer intervalo de un segundo
9. Tres personas lanzan una moneda al aire, el dueño de la moneda que resulte con cara distinta pagará la comida. Si los tres resultados son iguales las monedas se lanzan nuevamente, calcula la probabilidad de que se necesiten más de dos intentos para determinar al perdedor.
10. Un lote de 25 cinescopios de color se somete a un procedimiento de prueba de aceptación. Éste consiste en tomar cinco cinescopios sin reemplazo y probarlos; si dos o menos cinescopios fallan se acepta el lote, en caso contrario se rechaza. Supón que el lote contiene cuatro cinescopios defectuosos
 - a) calcula la probabilidad de que el lote pase la prueba
 - b) calcula cuántos de los cinco cinescopios se espera que no resulten defectuosos

11. Un examen de opción múltiple consta de 15 preguntas con cinco respuestas posibles cada una, de las cuales solamente una es correcta. Supón que uno de los estudiantes contesta el examen al azar, calcula la probabilidad de que conteste correctamente al menos diez preguntas. ¿Qué te indica el valor de la probabilidad obtenido, con respecto a las posibilidades de pasar del estudiante?
12. De un grupo de aspirantes para cierto trabajo industrial 30% tiene entrenamiento avanzado en programación. Los aspirantes son entrevistados uno tras otro, al azar. Calcula la probabilidad de que se encuentre al primer aspirante con entrenamiento avanzado en programación hasta la quinta entrevista.
13. Supón que la cajera de un banco atiende en promedio a 4.5 clientes cada diez minutos y que la cantidad de personas atendidas por la cajera sigue un proceso de Poisson. Supón, también, que la cajera es observada durante seis intervalos de diez minutos y se obtienen seis observaciones independientes, calcula la probabilidad de que en menos de dos de dichos intervalos, la cajera sólo atienda a tres clientes en el transcurso de diez minutos.
14. En una tienda se encontró que la venta de cierto artículo sigue un proceso de Poisson con promedio de cinco ventas al día, calcula la probabilidad de que en un día el artículo
 - a) sea pedido más de seis veces
 - b) si los pedidos diarios se suponen independientes, calcula la probabilidad de que tengan que pasar más de cuatro días para que soliciten más de seis pedidos

Autoevaluación

1. En un almacén, los clientes llegan al mostrador de caja conforme una distribución de Poisson con promedio de siete por hora. En una hora dada, calcula la probabilidad de que lleguen al menos dos clientes.
 - a) 0.8716
 - b) 0.7415
 - c) 0.6781
 - d) 0.9927
2. La probabilidad de que un motor, recién ajustado, tire aceite en los primeros 100 km por los retenes es de 0.05. Si diez automóviles se ajustan en un taller mecánico, calcula la probabilidad que menos de cuatro tiren aceite por los retenes.
 - a) 0.001
 - b) 0.999
 - c) 0.20
 - d) 0.80
3. Un profesor que imparte ecuaciones diferenciales encuentra que ocho de sus 27 alumnos no saben integrar por fracciones parciales. El coordinador académico examina a cuatro alumnos de ese grupo al azar, calcula la probabilidad de que al menos dos alumnos de esta muestra no sepan integrar por dicho método.

- a) 0.6626
 - b) 0.3374
 - c) 0.25
 - d) 0.476
4. La probabilidad de que un estudiante de aviación apruebe el examen escrito para obtener su licencia de piloto es 0.6, calcula la probabilidad de que apruebe el examen antes del cuarto intento.
- a) 0.936
 - b) 0.1296
 - c) 0.064
 - d) 0.8704
5. En una ciudad se efectuaron encuestas a un gran número de amas de casa para saber si por la noche el agua de su casa se salía de las cisternas, se encontró que aproximadamente 5% contestó afirmativamente. Calcula la probabilidad de que al inspeccionar 20 casas, por la noche, al menos en una de ellas se tire el agua.
- a) 0.3585
 - b) 0.4780
 - c) 0.6415
 - d) 0.5220
6. Una de las máquinas para elaborar tornillos milimétricos se descompuso, por lo que un gran número de tornillos resultó defectuoso. Para tratar de evitar pérdidas, en cada caja de 30 tornillos se colocan cinco defectuosos (25 sin defectos). El vendedor de tornillos comienza a recibir reclamos por los defectuosos, y decide cambiar de proveedor si antes del cuarto de los siguientes envíos recibe una caja donde al revisar aleatoriamente seis tornillos se encuentran dos o más defectuosos. Calcula la probabilidad de que el vendedor cambie de proveedor.
- a) 0.5854
 - b) 0.4146
 - c) 0.167
 - d) 0.833
7. Para llevar a cabo un reporte de control de calidad sobre la fabricación de videos, se analizan lotes independientes de quince videos, cada uno. Si de cada uno de estos lotes se elige una muestra aleatoria de tres de ellos y se prueban considerando que el productor ha puesto en cada lote dos defectuosos, el total de los videos se comprará, si después de analizar un lote no se han encontrado defectuosos. Calcula la probabilidad de que bajo estas condiciones se compre el total de videos.
- a) 0.3714
 - b) 0.20
 - c) 0.6286
 - d) 0.80

8. De un grupo de aspirantes para cierto trabajo industrial 30% tiene entrenamiento avanzado en programación. Los aspirantes son entrevistados uno tras otro, al azar, calcula la probabilidad de que se encuentre el tercer aspirante con entrenamiento avanzado en programación hasta la sexta entrevista.
- a) 0.917
 - b) 0.050
 - c) 0.093
 - d) 0.504
9. Una compañía quiere evaluar sus procedimientos de inspección en embarques de 50 artículos idénticos. El procedimiento consiste en tomar una muestra de cinco y aceptar el embarque si no se encuentran más de dos defectuosos, calcula qué proporción de embarques con 20% de artículos defectuosos será aceptada.
- a) 0.0483
 - b) 0.897
 - c) 0.103
 - d) 0.9517
10. El promedio de automóviles que entran al túnel de una montaña es de un coche cada dos minutos. Si la entrada de los coches al túnel sigue una distribución de Poisson, calcula la probabilidad de que el número de coches que entra al túnel durante un periodo de dos minutos exceda de tres.
- a) 0.981
 - b) 0.019
 - c) 0.568
 - d) 0.432

Respuestas de los ejercicios

Ejercicio 1

- 1.
- a) 0.2662
 - b) 0.9718
 - c) 10
- 2.
- a) 0.4557
 - b) 17
- 3.
- a) 0.0058
 - b) 47.5

4.

- a) 0.0577
- b) 0.5881

5.

- a) 0.6083
- b) 2

Ejercicio 2

1. 0.4633

2.

- a) 0.0039
- b) teórico 1.333

3.

- a) 0.0256
- b) teórico 1.67

4. 0.8587

5.

- a) 0.063
- b) $(0.3)^9$

Ejercicio 3

1.

- a) 0.00015
- b) 0.028

2. 0.0604

3. 0.0107

4.

- a) 0.9728
- b) 15

Ejercicio 4

1. $P(Y=0) = 0.2$, $P(Y=1) = 0.6$, $P(Y=2) = 0.2$,

2.

- a) 0.0238
- b) 2

3. 0.3830
4. 0.34737
5.
 - a) 0.36894
 - b) teórico 5.7, aproximado 6

Ejercicio 5

1. 0.8647
2.
 - a) 0.1502
 - b) 0.1506
3.
 - a) 0.6767
 - b) 0.7909
4. 0.1496
5. 0.1125

Respuestas de los ejercicios propuestos

1.
 - a) 0.0861
 - b) 10
2.
 - a) 0.4170
 - b) 0.0821
3. 0.096
4.
 - a) \$400
 - b) 0.0458
5.
 - a) 0.8
 - b) 0.7407

6.
a) 0.0818
b) 0.1277
7.
a) 0.1171
b) 0.5841
8.
a) 0.081
b) 0.729
9. 0.0625
10.
a) 0.9838
b) 4.2
11. 0.0001
12. 0.07203
13. 0.7319
14.
a) 0.2378
b) 0.3375

Respuestas de la autoevaluación

1. d)
2. b)
3. b)
4. a)
5. c)
6. a)
7. c)
8. c)
9. d)
10. b)