SESION 12

LA DISTRIBUCIÓN BINOMIAL

I. CONTENIDOS:

1. La distribución binomial.
2. Variables aleatorias en una distribución binomial.
3. Descripciones de la distribución binomial.
4. Distribución de Poisson.

II. OBJETIVOS:

Al término de la Clase, el alumno:

- Conocerá e interpretará las áreas bajo la curva de una distribución binomial.

III. PROBLEMATIZACIÓN:

Comenta las preguntas con tu Asesor y selecciona las ideas más significativas.

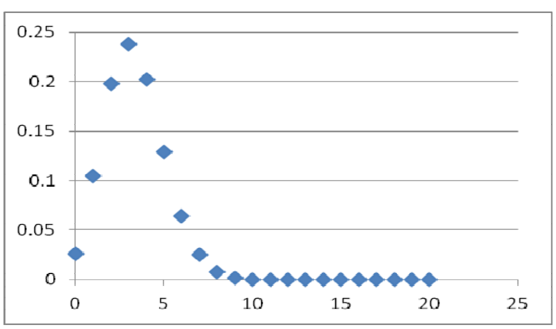
- Si se maneja una situación que se refiere a ensayos repetidos, donde sólo existen dos posibles resultados para cada ensayo, ¿cómo determinarías la probabilidad?
- En un problema en donde se requiere la probabilidad de que fallen o no un grupo de máquinas de un total de cien, ¿cuál sería el camino más adecuado para determinar la probabilidad?

IV. TEXTO INFORMATIVO-FORMATIVO:

1.1. La distribución binomial

La distribución binomial es una distribución de probabilidad exclusiva para variables aleatorias discretas que tienen sólo dos categorías mutuamente excluyentes, dónde una se considera como éxito y la otra como fracaso; además las probabilidades de cada categoría permanecen constantes en cada nueva elección. Fue desarrollada por Jacobo Bernoulli un matemático francés en 1713. Describe la distribución de variables aleatorias discretas en los llamados procesos de Bernoulli. Un proceso de Bernoulli es una serie de repeticiones de un experimento aleatorio en el que sólo puede haber dos resultados posibles, por ejemplo:

Se lanza un dado veinte veces, en cada una se registra la puntuación obtenida. Si la probabilidad teórica de obtener un dos en cualquier lanzamiento es de 1/6, ¿Cuál es la probabilidad de obtener un dos en 0,1, 2, 3,…, 20 ocasiones?



Como podemos apreciar en la gráfica, el obtener tres veces el número dos al lanzar un dado en veinte ocasiones es lo más probable (23.79%), pero la probabilidad de obtener diez o más números dos es prácticamente de cero (0.06%). Es más probable no obtener números dos al lanzar el dado veinte veces (2.61%), que obtener nueve veces el número dos en los citados veinte lanzamientos (0.22%)

2.1. Variables aleatorias en una distribución binomial

La variable de una distribución binomial debe cumplir con las siguientes características:

- *En cada una de las repeticiones solamente se pueden obtener dos resultados diferentes; es decir, la variable que se mide tiene sólo dos categorías. A una de ellas se le considera como éxito y a la otra como fracaso.*
- *El resultado de un ensayo es independiente del resto. Esto significa que los eventos son independientes tal como se explicó en la clase 10.*
- *La probabilidad de ocurrencia de cada categoría es la misma en todos los ensayos.*
- *Hay un número determinado de ensayos a realizar.*

3.1. Descripciones de la distribución binomial

A fin de facilitar el uso de la función de probabilidad binomial, se aconseja que se considere como éxito el obtener la categoría que se menciona en las preguntas del problema a resolver, la otra categoría será un fracaso si se obtiene en alguna elección.

Entonces, se hacen una serie de repeticiones de un fenómeno aleatorio que sólo puede tener dos categorías mutuamente excluyentes en la variable que se estudia, ese número de elecciones se representa con "n"; luego, de estas elecciones se espera que aparezca "n_1" veces la categoría que consideramos éxito, y "n_2" veces la otra categoría. También contamos con la probabilidad de ocurrencia de cada categoría, donde "p" es la probabilidad de éxito y "q" la de fracaso. Se utiliza la siguiente función:

$$P_{bin}(n_1) = \frac{n!}{n_1!n_2!} p^{n_1} q^{n_2}$$

La media y la desviación estándar de la distribución binomial se calcula con las fórmulas:

$$\overline{X}_{bin} = np$$

$$S_{bin} = \sqrt{npq}$$

Ejemplo 1 Una ensambladora de computadoras adquiere 400 memorias DDR. El lote de memorias está en el área de recepción del almacén. El criterio que se emplea para aceptar el lote consiste en que de 20 memorias, seleccionadas al azar, a lo sumo una esté con algún defecto.

Sin embargo, el lote tiene un 6% de memorias defectuosas. Si se sigue el protocolo descrito antes, ¿cuál es la probabilidad de que se acepte el lote? Para comenzar debemos identificar la variable, en este caso es "funcionamiento de la memoria", sus dos categorías son: "defectuosa" y "funcional"

Luego debemos identificar la categoría que es considerada como éxito; en las elecciones que se van a realizar, ¿qué se busca?, la respuesta es memorias defectuosas; entonces se considera como éxito a la categoría "defectuosa"

Después debemos conocer la probabilidad de ocurrencia de cada categoría. Como menciona el enunciado "el lote tiene un 6% de memorias defectuosas", entonces la probabilidad de éxito es de 0.06. Recuerda que en los cálculos no se deben utilizar probabilidades en porcentaje.
Por consiguiente, la probabilidad de fracaso es 0.94, pues es el complemento para el 100%.

El número de repeticiones es la cantidad de veces que se ensaya el fenómeno para encontrar los llamados "éxitos" que son simplemente una de dos categorías que pueden ocurrir, aquí se van a examinar 20 memorias entonces el número de repeticiones es tal.

Finalmente, como la cuestión expresa que "a lo sumo haya una memoria defectuosa", entonces debemos calcular por separado para: cero memorias defectuosas y una memoria defectuosa, que son los resultados que permitirán que el lote sea aceptado. Finalmente debemos sumar esas dos probabilidades de acuerdo con lo aprendido en el tema de la combinación de probabilidades. Aplicando la función, con los valores descritos:

Probabilidad de encontrar cero memorias defectuosas:

$$n = 20 \qquad n_1 = 0 \qquad n_2 = 20$$

$$P_{bin}(n_1) = \frac{n!}{n_1!n_2!} p^{n_1} q^{n_2} = P_{bin}(0) = \frac{20!}{0!20!}(0.06)^0 (0.94)^{20} = 0.2901$$

Probabilidad de encontrar una memoria defectuosa:

$$n = 20 \qquad n_1 = 1 \qquad n_2 = 19$$

$$P_{bin}(n_1) = \frac{n!}{n_1!n_2!} p^{n_1} q^{n_2} = P_{bin}(1) = \frac{20!}{1!19!}(0.06)^1 (0.94)^{19} = 0.3703$$

Sumando los resultados, la probabilidad de que el lote sea aceptado es del 66.04%

Ejemplo 2 ¿Cuál es la media y la desviación estándar de la distribución binomial anterior?

Aplicando las fórmulas que se han mostrado, para cada caso:

$$\overline{X}_{bin} = np = (20)(0.06) = 1.2$$

$$S_{bin} = \sqrt{npq} = \sqrt{(20)(0.06)(0.94)} = 1.062$$

Lo anterior significa que el promedio de memorias defectuosas que se han de encontrar es de 1.2, que rebasa lo que se tiene como valor máximo de memorias defectuosas para aceptar el lote. Por esa razón hay una probabilidad de casi el 34% de que no se acepte el lote de memorias.

4.1. Distribución de Poisson

Es una distribución de probabilidad discreta que expresa la probabilidad de que ocurra una cantidad determinada de eventos independientes, en un lapso de tiempo o magnitud espacial; considerando sólo a la media o promedio que corresponde a la ocurrencia de la categoría que se espera y el número de ocasiones que se espera aparezca dicha categoría de una variable discreta.

Por ejemplo, si se sabe que en promedio se ejecuta a una persona cada 40 minutos, en México (2010), a causa de la violencia, ¿cuál es la probabilidad de que en dos horas contiguas, mueran sólo dos personas por esta causa?

Aquí la media será 3, pues la pregunta se hace sobre un lapso de dos horas. Como en cada 40 minutos hay en promedio un ejecutado, entonces en dos horas habrá en promedio 3 ejecutados. Por otra parte, el número de ocasiones que se espera aparezca la categoría es de 2.
La función que permite calcular esto es:

$$f(x;\lambda) = \frac{\lambda^x}{e^\lambda x!}$$

Donde λ es la media o promedio que se ha registrado para la ocurrencia de la categoría, y x es el número de ocasiones que se espera aparezca la categoría.

El número e es llamado de Euler y se calcula con:

$$\sum_{n=0}^{\infty} \frac{1}{n!} \approx 2.7182818284$$

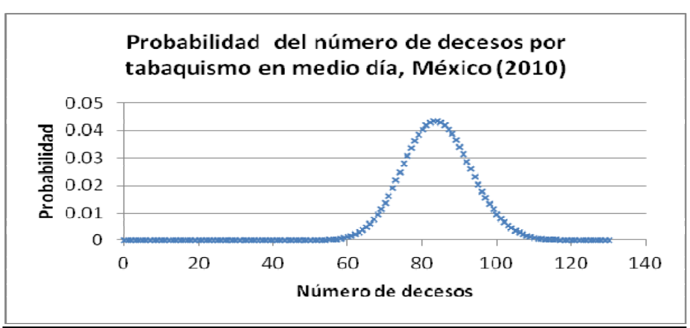
Para resolver el problema anterior se sustituyen los valores:

$$f(x;\lambda) = \frac{\lambda^x}{e^{\lambda} x!} = \frac{3^2}{(e^3)(2!)} = 22.40\%$$

Ejemplo 1 Se ha calculado que el número de muertes por causa del tabaquismo, en México (2010), es de 7 cada hora en promedio. Con esta información calcula la probabilidad de que en medio día mueran 70 o menos, 75 o menos, 80 o menos, 85 o menos, 90 o menos, 95 o menos, 100 o menos, por esta causa.

Se procede a calcular con la función de la distribución de Poisson, para 0, 1, 2,…100 muertes, con una media de 84, pues medio día son 12 horas y cada hora mueren 7. Después se suman las probabilidades, como se ha mostrado en la clase 2, para encontrar la probabilidad acumulada. Por razones de espacio no se muestran los 100 cálculos, pero sí los resultados:

Muertes en 12 horas	Probabilidad acumulada
70 o menos	6.725%
75 o menos	17.742%
80 o menos	35.709%
85 o menos	57.195%
90 o menos	76.372%
95 o menos	89.352%
100 o menos	96.108%



Como se puede ver en el gráfico, la probabilidad de que en medio día haya menos de 50 decesos por tabaquismo es prácticamente cero; pero aumenta la probabilidad desde 60 hasta 84 (que es la media) y luego desciende hasta 108 decesos. Después de 108 decesos, la probabilidad es casi cero.

V. ESTRATEGIAS CENTRADAS EN EL APRENDIZAJE:

A. Realiza las siguientes actividades.

1. En una cierta población, el 52% de los nacimientos registrados son de varones. Si tomamos cinco registros define la variable que te permita calcular las probabilidades que se piden a continuación (piensa que se trata de una población grande).

   a) Que dos registros correspondan a varones.
   b) Menos de tres sean varones.

2. Una caja contiene 30 baterías para radio, de las cuales cinco son defectuosas. De la caja se escogen al azar seis baterías; determina la probabilidad de que:

   a) Dos sean defectuosas.
   b) Ninguna sea defectuosa.
   c) Menos de tres sean defectuosas.

3. El promedio de personas que llegan a la ventanilla de un banco por minuto durante las horas hábiles es de una. Determina la probabilidad de que en un minuto dado:

   a) No aparezca clientes.
   b) Hay tres o más clientes.
   c) Hay tres o menos clientes.