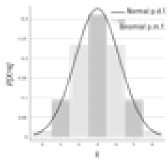
Probabilidad y Estadística 6

Actividades de Aprendizaje



Conceptos y definiciones de esta clase:

Distribución Binomial
Características
Utilidad

Esperanza Matemática de una distribución binomial

1.17 Distribución Binomial

La distribución binomial es una **distribución discreta** muy significativa que surge en muchas aplicaciones industriales de control de calidad, bioestadísticas, sistemas de producción, etc. Esta distribución aparece al realizar repeticiones independientes de un experimento que tenga **respuesta binaria**, como aquellas que dan "verdadero" o "falso", "correcto" o "incorrecto", "éxito" o "fracaso", "encendido" o "apagado" y otros por el estilo.

La variable discreta que cuenta el número de éxitos en **n pruebas** independientes de ese experimento, cada una de ellas con la misma **probabilidad de "éxito" igual a p**, sigue una distribución binomial de parámetros **n** y **p**. Este modelo se aplica a poblaciones finitas en las que se toman elementos al azar con reemplazo, y también a poblaciones conceptualmente infinitas, como por ejemplo las piezas que produce una máquina, siempre que el proceso de producción sea: **estable** (la proporción de piezas defectuosas se mantiene constante a largo plazo), y **sin memoria** (el resultado de cada pieza no depende de las anteriores).

Ejemplo 1

Supongamos que en el día de hoy Usted se dirige a clases como de costumbre, y cuando llega, el profesor les propone realizar una pequeña evaluación como para verificar que los alumnos hayan repasado los contenidos, tal como se los ha pedido la clase anterior. Y para hacerlo, decide que una buena herramienta es tomar un pequeño cuestionario de **cuatro preguntas que se contestan con opción múltiple, dando tres opciones para cada pregunta**.

La verdad es que Usted no ha tenido mucho tiempo de repasar durante la semana, pero se considera una persona afortunada que saldrá exitosa del examen, aún sin haber estudiado.

Deténgase un minuto. ***¿Se ha puesto a pensar qué probabilidades de éxito tiene?***

Bueno, vamos a resolver esta situación con lo recientemente aprendido.

Para comenzar, ***anote en un papel cuáles serían las respuestas que elegiría***

A priori sabemos que nuestras posibilidades van desde el fracaso total (las cuatro preguntas incorrectas) hasta un rotundo éxito (las cuatro preguntas respondidas correctamente).

Esta es la manera en que se vería la evaluación:

<u>Evaluación</u>

Pregunta 1: …

Respuestas: a) …
b) …
c)…

Pregunta 2: …

Respuestas: a) …
b) …
c)…

Pregunta 3: …

Respuestas: a) …
b) …
c)…

Pregunta 4: …

Respuestas: a) …
b) …
c)…

Ahora bien, de acuerdo a lo recién visto, tendremos que definir los dos parámetros **n** y **p**, a saber:

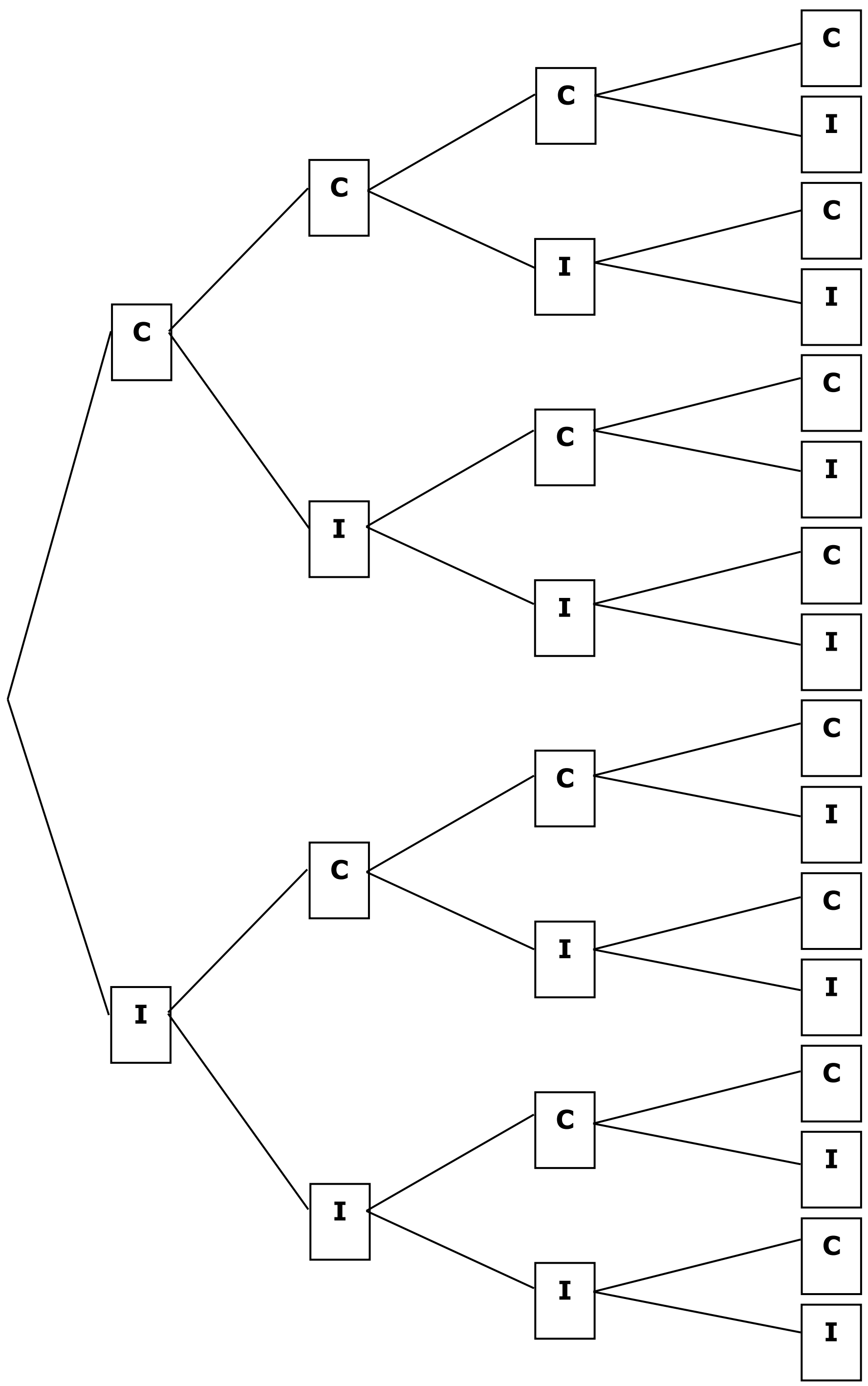
n será la cantidad de respuestas acertadas (respuestas correctas)

p será la probabilidad que tiene cada n

Con respecto a **n**, no habrá demasiadas dudas, sabemos que es un número entero, que en este caso tomará valores entre 0 (lo que anteriormente mencionamos como ninguna respuesta correcta), hasta un valor de 4 (las cuatro preguntas correctas).

Con respecto a **p**, tendremos que hallar las probabilidades de cada situación. Para comenzar, evaluemos cuáles son las posibilidades. Esto ha sido visto en unidades anteriores, y básicamente puede materializarse realizando un diagrama de árbol, o una tabla.

Si elegimos utilizar un diagrama de árbol, la situación se vería de la siguiente manera:



También podemos elegir otra manera de representación de esta situación. Por ejemplo, utilizar una tabla:

Pregunta 1	Pregunta 2	Pregunta 3	Pregunta 4	Total. Correctas
C	C	C	C	4
C	C	C	I	3
C	C	I	C	3
C	C	I	I	2
C	I	C	C	3
C	I	C	I	2
C	I	I	C	2
C	I	I	I	1
I	C	C	C	3
I	C	C	I	2
I	C	I	C	2
I	C	I	I	1
I	I	C	C	2
I	I	C	I	1
I	I	I	C	1
I	I	I	I	0

Cualquiera sea el método elegido, enseguida podrá apreciarse, para cada caso posible, cuántas veces se logra el objetivo (que es dar una respuesta correcta.

Ahora bien, como cada pregunta contiene tres respuestas, la probabilidad de responder correctamente es de 1/3, mientras que la de responder incorrectamente es de 2/3.

De esta manera, para el caso de que usted acierte las cuatro, la probabilidad es:

$$P(n=4)=\frac{1}{3}x\frac{1}{3}x\frac{1}{3}x\frac{1}{3}=\frac{1}{81}=0.0123\ldots$$

De manera similar, no acertar a ninguna tiene una probabilidad:

$$P(n=0)=\frac{2}{3}x\frac{2}{3}x\frac{2}{3}x\frac{2}{3}=\frac{16}{81}=0.1975\ldots$$

Y así podemos calculas las probabilidades para n=1, n=2 y n=3. **¡¡Pero atención!!** En estos casos la ocurrencia deja de ser 1 como en las dos anteriores.

Por ejemplo, la probabilidad de n=1 es:

$$P(n=1)=4\,x\frac{1}{3}x\frac{2}{3}x\frac{2}{3}x\frac{2}{3}=\frac{32}{81}=0.395\ldots$$

Nótese que se multiplica por 4, puesto que hay cuatro posibilidades de responder una sola respuesta correcta.

Una vez realizados todos los cálculos, podremos completar la siguiente tabla:

n	P(n)
0	0.1975
1	0.3951
2	0.2963
3	0.0987
4	0.0123
Total	0.9999

Cabe observar que se ha preferido tomar cuatro decimales para aproximar el total de probabilidades a 1. Se pueden utilizar menos decimales sin problema, siempre y cuando se redondeen correctamente los resultados y se verifique que la suma de probabilidades de por resultado 1.

Actividad para la Facilitación de los Aprendizajes

Nota importante para el alumno:

Es muy común, en este tipo de ejercicios confundir el tipo de respuesta con la cantidad de resultados. Recordemos que se trata de una distribución **binomial**, que tendrá **dos** opciones, por ejemplo "responder bien" y "responder mal", sin importar cuántos resultados tiene cada pregunta, si bien es este número el que dará mayores o menores probabilidades a cada situación.

Situación Problemática 13

En base al análisis realizado responda las siguientes preguntas:

a- Si la evaluación se aprueba con un mínimo de tres preguntas respondidas correctamente, ¿qué probabilidades tiene de aprobar?
b- ¿Cuál es el resultado más probable de obtener?
c- Si en la clase hay 23 alumnos y ninguno ha estudiado. ¿Cuántos habrán aprobado el examen "adivinando" las respuestas?
d- Suponiendo que las respuestas correctas son 1.c, 2.c, 3.a, 4.b. ¿Qué resultado obtuvo con el método de "adivinar" las respuestas?

Situación Problemática 14

Solicitar a varias personas (no menos de 10) que adivinen los resultados de esta evaluación, y luego construir una tabla con frecuencias relativas para n desde 0 a 4 y compararlo con los resultados obtenidos en nuestro análisis.

Situación Problemática 15

Repetir el ejercicio realizado como ejemplo, pero suponiendo 10 preguntas con posibilidades de verdadero (V) y falso (F), en el que se aprueba con 6 respuestas correctas. ¿Es más conveniente para un alumno que no estudió o menos conveniente?

Es natural que a estas alturas uno se pregunte cómo sabe si se encuentra frente a un experimento en el que podrá aplicar lo aprendido, o cómo diseñar correctamente un estudio para que puedan aplicarse estos conocimientos.

Para facilitar esta tarea enunciaremos brevemente algunas propiedades que se cumplen en un experimento de este tipo.

1. Hay sólo dos resultados posibles para la respuesta de cada prueba, que podemos generalizar como "correcto" o "incorrecto".
2. Se realiza una cantidad n de pruebas independientes unas de otras (los resultados de cada una de ellas no tienen influencia sobre el resultado de las restantes).

3. La probabilidad de éxito más la probabilidad de fracaso es 1.
4. La variable binomial aleatoria representa la cantidad de pruebas exitosas que se obtienen, y toma un valor de 0 a n.

Si analizamos nuestro ejemplo según estas propiedades, obtenemos que:

1. Se cumple porque hay sólo dos posibilidades para cada pregunta, aun cuando tenga tres posibilidades de respuesta. En otras palabras, la pregunta está bien o mal respondida.
2. Los resultados son independientes, porque no dependen de otros.
3. La probabilidad de éxito es 1/3 y la de fracaso es 2/3, y su suma es 1.
4. La cantidad de preguntas respondidas puede ir de 0 a 4.

La distribución discreta de probabilidad se suele denominar **distribución binomial** porque para *r=0, 1, 2, 3, …, n* corresponde a los sucesivos términos de la forma o expansión binomial. Todos conocemos la expansión del binomio:

$$(a+b)^2 = a^2 + 2\cdot a\cdot b + b^2$$

En general, si tenemos p = probabilidad del éxito y $q = 1 - p$ = probabilidad del fracaso, la expansión será:

$$(p+q)^n = p^n + C_n^1\cdot p\cdot q^{n-1} + C_n^2\cdot p^2\cdot q^{n-2} + \ldots + q^n$$

y los valores de las combinaciones C_n^1, C_n^2, C_n^3, etc. se llaman coeficientes binomiales.
Así, por ejemplo, si queremos saber cuál es la probabilidad de obtener 3 caras tirando 5 monedas, como p=q=0,5 hacemos la expansión correspondiente:

$$(p+q)^5 = p^5 + 5\cdot p^4\cdot q + 10\cdot p^3\cdot q^2 + 10\cdot p^2\cdot q^3 + 5\cdot p\cdot q^4 + q^5$$

buscamos el término correspondiente y obtenemos el cociente entre el coeficiente binomial correspondiente y la cantidad total de ensayos o tiradas de monedas y tenemos:

$$= 10\cdot p^3\cdot q^2 = 10\cdot\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^2 = 10\cdot\frac{1}{8}\cdot\frac{1}{4} = \frac{10}{32} = 0{,}31$$

Este cálculo corresponde con la fórmula de la Distribución Binomial.

$$P_{(3)} = C_5^3\cdot\left(\frac{1}{2}\right)^3\cdot\left(\frac{1}{2}\right)^{5-3} = \frac{5\,!}{2\,!\,3\,!}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2 = \frac{5.4.3.2.1}{3.2.1.2.1}\cdot\frac{1}{8}\cdot\frac{1}{4} = \frac{120}{384}\cdot\cdot = 0{,}3125$$

Generalizando:

$$P_{(r)} = C_n^r\cdot p^r\cdot q^{n-r}$$

Esperanza Matemática de una distribución binomial.

Sabemos que cuando una variable aleatoria o sea al azar, puede tomar solo un conjunto finito de valores, es decir valores que son números enteros (por ejemplo, cantidad de hijos en una familia, caras de un dado, etc.) hablamos de una variable discreta. En este caso la distribución de probabilidad de que una familia tenga 2 o más hijos y una probabilidad de que tenga más de 2 hijos y menos de tres. Siempre que tenemos una variable aleatoria discreta en la que haya valores de x_1 a x_n la serie de las probabilidades de ocurrencia para los *n* valores es igual a 1, es decir que:

$$p(x_1)+p(x_2)+p(x_3)+\ldots+p(x_n)=1$$

La media aritmética de una distribución de probabilidad se llama **esperanza matemática**. Es decir, el valor esperado para una variable aleatoria X cuyo símbolo es E(X) se obtiene sumando el producto de cada uno de los valores en que puede darse la variable por su probabilidad, es decir:

$$E(X)=\sum X\cdot p(X)$$

Por ejemplo, si tenemos la siguiente distribución.

X = número de hijos	Familia	p(X)	X.p(X)
1	10	10/120	(1*10) /120
2	20	20/120	(2*20) /120
3	30	30/120	(3*30) /120
4	25	25/120	(4*25) /120
5	20	20/120	(5*20) /120
6	15	15/120	(6*15) /120
	120		430

$$E(X)=\sum X\cdot p(X)=\frac{430}{120}=3{,}58$$

Veamos ahora un nuevo Problema ideado para el área de Informática, y donde integraremos varios conceptos vistos anteriormente.

Problema: Un sistema de comunicación binario envía dígitos "0" y "1". A causa del ruido del sistema, a veces se envía un 0 y en cambio se recibe un 1, y viceversa. Se sabe que la probabilidad de que un 0 se transmita incorrectamente (no correctamente) es 0.06. Por otro lado, la probabilidad de que un 1 se transmita correctamente es 0.9 y la probabilidad de enviar un 0 es 0.45.

Comenzamos definiendo los sucesos de "enviar" y "recibir". Llamamos:

$$E_0=\{\textit{se envió un 0}\}$$

$$E_1=\{\textit{se envió un 1}\}=\overline{E_0}$$ (Complemento de E_0)

$$R_0=\{\textit{se recibió un 0}\}$$

$$R_1 = \{se\ recibió\ un\ 1\} = \overline{R_0}$$ (Complemento de R_0)

Analizando el enunciado que un 0 se transmita incorrectamente, implica que se recibió un 1. Es decir, estamos ante la probabilidad de recibir un 1 dado que se envió un 0. En símbolos: $P\left(R_1 \middle| E_0\right) = 0.06$.

Luego, que un 1 se transmita correctamente implica que se recibió un 1 dado que se envió un 1. En símbolos:

$P\left(R_1 \middle| E_1\right) = 0.9$.

Por último, la probabilidad de enviar un 0 es en símbolos: $P\left(E_0\right) = 0.45$

Dada la siguiente Tabla de Contingencia

	E_0	**E_1**	
R_0	*0.423*	*0.055*	*0.478*
R_1	*0.027*	*0.495*	*0.522*
	0.45	*0.55*	***1***

a) Calcule la probabilidad de que, en el envío de un bit (dígito binario), la transmisión sea correcta. Defina el suceso, observe la tabla y recuerde propiedades.

b) Suponiendo que el sistema de comunicación envía una cadena de dos bits. Para mejorar el canal de comunicación, se ha decidido analizar <u>la cantidad de bits correctamente transmitidos</u>. ¿Cuál es la variable aleatoria en análisis?

c) Construya una tabla y represente gráficamente la función de distribución de probabilidad

d) ¿Cuántos bits se espera que hayan sido transmitidos correctamente?

e) Observe la tabla y la gráfica construida en el inciso c), ¿se identifica con alguna distribución de probabilidad conocida? Justifica.