TEMA 5 MODELOS DE VARIABLES ALEATORIAS
DISTRIBUCIÓN BINOMIAL

Pablo Buenestado

Curso 2019-2020 Primavera

Departamento de Matemáticas (UPC)

Índice

Esquema

Con frecuencia un experimento consta de pruebas repetidas, cada una con dos resultados posibles que se pueden denominar éxito o fracaso. La aplicación más evidente tiene que ver con la prueba de artículos a medida que salen de una línea de ensamble, donde cada prueba o experimento puede indicar si un artículo está o no defectuoso. Podemos elegir definir cualquiera de los resultados como éxito. El proceso se conoce como proceso de Bernoulli y cada ensayo se denomina experimento de Bernoulli.

Por ejemplo, si extraemos cartas de una baraja y éstas no se reemplazan, cambian las probabilidades en la repetición de cada ensayo; es decir, la probabilidad de seleccionar una carta de corazones en la primera extracción es $1/4$, pero en la segunda es una probabilidad condicional que tiene un valor de $13/51$ o $12/51$, dependiendo de si resulta un corazón en la primera extracción; entonces éste ya no sería considerado un conjunto de experimentos de Bernoulli.

Esquema

Definición

En términos estrictos el proceso de Bernoulli se caracteriza por lo siguiente:

1. El experimento consta de ensayos repetidos.
2. Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso.
3. La probabilidad de un éxito, que se denota con p, permanece constante de un ensayo a otro.
4. Los ensayos repetidos son independientes.

Esquema

Consideremos el conjunto de experimentos de Bernoulli en el que se seleccionan tres artículos al azar de un proceso de producción, luego se inspeccionan y se clasifican como defectuosos o no defectuosos. Un artículo defectuoso se designa como un éxito. El número de éxitos es una variable aleatoria X que toma valores integrales de cero a 3. Los ocho resultados posibles y los valores correspondientes de X son

Resultado	NNN	NDN	NND	DNN	NDD	DND	DDN	DDD
x	0	1	1	1	2	2	2	3

Como los artículos se seleccionan de forma independiente y suponiendo que el proceso genera un 25% de artículos defectuosos

$$P(NDN) = P(N) \cdot P(D) \cdot P(N) = \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} = \frac{9}{64}$$

Cálculos similares dan las probabilidades para los otros resultados posibles. En consecuencia, por ejemplo para $X = 2$, dado que hay tres formas de obtener 2 artículos defectuosos, tenemos

$$P(X = 2) = 3 \cdot \frac{9}{64} = \frac{27}{64}$$

La distribución de probabilidad de X es, por lo tanto,

x	0	1	2	3
$f(x)$	$\frac{27}{64}$	$\frac{27}{64}$	$\frac{9}{64}$	$\frac{1}{64}$

El número X de éxitos en n experimentos de Bernoulli se denomina variable aleatoria binomial. La distribución de probabilidad de esta variable aleatoria discreta se llama distribución binomial y sus valores se denotarán como $b(x; n, p)$, ya que dependen del número de ensayos y de la probabilidad de éxito en un ensayo dado. Por consiguiente, para la distribución de probabilidad de X el número de productos defectuosos es

$$P(X = 2) = f(2) = b\left(2; 3, \frac{1}{4}\right) = \frac{9}{64}$$

Generalizamos ahora la ilustración anterior con el fin de obtener una fórmula para $b(x; n, p)$.

Esto significa que deseamos encontrar una fórmula que dé la probabilidad de x éxitos en n ensayos para un experimento binomial.

Empezamos por considerar la probabilidad de x éxitos y $n - x$ fracasos en un orden específico.

Como los ensayos son independientes, podemos multiplicar todas las probabilidades que corresponden a los diferentes resultados. Cada éxito ocurre con probabilidad p y cada fracaso con probabilidad $q = 1 - p$. Por lo tanto, la probabilidad para el orden específico es $p^x \cdot q^{n-x}$.

Ahora debemos determinar el número total de puntos muestrales en el experimento que tienen x éxitos y $n - x$ fracasos. Este número es igual al número de particiones de n resultados en dos grupos con x en un grupo y $n - x$ en el otro, y se escribe $\binom{n}{x}$.

Como estas particiones son mutuamente excluyentes, sumamos las probabilidades de todas las diferentes particiones para obtener la fórmula general o simplemente multiplicamos $p^x \cdot q^{n-x}$ por $\binom{n}{x}$.

Esquema

Definición

Distribución binomial

Un experimiento de Bernoulli puede tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1 - p$. Entonces, la distribución de probabilidad de la variable aleatoria binomial X, el número de éxitos en n ensayos independientes, es

$$b(x; n, p) = \binom{n}{x} p^x \cdot q^{n-x}; \qquad x = 0, 1, 2 \ldots, n$$

Expresamos la variable aleatoria binomial como $X \sim B(n, p)$ y el conjunto de valores de la misma es $X = \{0, 1, 2 \ldots, n\}$

Esquema

Definición

Si $X \sim B(n, p)$, entonces la media y la varianza son, respectivamente:

$$\mu = E(X) = n \cdot p$$

$$\sigma^2 = Var(X) = n \cdot p \cdot (1 - p)$$

Para la variable aleatoria X que cuenta el número de artículos defectuosos en una muestra de 3 artículos elegidos al azar, dónde la probabilidad de Bernoulli de pieza extraer un artículo defetuoso es de 0.25, entonces $X \sim B(3, 0.25)$, ¿cuánto valen la media y la varianza de X?

Solución

$$\mu = 3 \cdot \frac{1}{4} = \frac{3}{4} \qquad \sigma^2 = 3 \cdot \frac{1}{4} \cdot \left(1 - \frac{1}{4}\right) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$$

Esquema

EJERCICIO 1

La probabilidad de que un paciente se recupere de una rara enfermedad sanguínea es de 0.4. Si se sabe que 15 personas contrajeron la enfermedad, ¿cuál es la probabilidad de que sobrevivan al menos 10?

Solución

Sea X el número de personas que sobreviven, asumiendo independencia, como la probabilidad individual de recuperación es la misma para todos, 0.4, y seleccionamos al azar 15 enfermos, podemos afirmar que $X \sim B(n, p)$ con $n = 15$ y $p = 0.4$. Para responder a la pregunta hemos de buscar la probabilidad que recoge el conjunto de al menos 10 personas, que se corresponde con el subconjunto de $\{X \geq 10\} = \{10, 11, \ldots 15\}$, o sea $P(X \geq 10)$:

$$P(X \geq 10) = \sum_{x=10}^{15} \binom{15}{x} 0.4^x \cdot 0.6^{15-x} = 0.0338$$

EJERCICIO 2

La probabilidad de que un paciente se recupere de una rara enfermedad sanguínea es de 0.4. Si se sabe que 15 personas contrajeron la enfermedad. Siendo X el número de personas que no sobreviven, calcula su media y su desviación estándar.

Solución

X es una distribución binomial con parámetros $n = 15$ y $p = 0.6$, es decir $X \sim B(15, 0.6)$, buscamos sus medidas características:

$$\mu = E(X) = n \cdot p = 15 \cdot 0.6 = 9$$

Es decir, esperamos que no sobrevivan 9 de los 15 enfermos.

$$\sigma = \sqrt{Var(X)} = \sqrt{\sigma^2} = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{15 \cdot 0.6 \cdot 0.4} = 1.897$$

EJERCICIO 3

Una cadena grande de tiendas al detalle le compra cierto tipo de dispositivo electrónico a un fabricante, el cual le indica que la tasa de dispositivos defectuosos es de 3%.

a) El inspector de la cadena elige 20 artículos al azar de un cargamento. ¿Cuál es la probabilidad de que haya al menos un artículo defectuoso entre estos 20?

Solución

Declaramos X como la variable aleatoria que cuenta el número de dispositivos defectuosos de los 20. Siendo 20 extracciones de Bernoulli, entonces $X \sim B(n,p)$ con $n = 20$ y $p = 0.03$. Así $X = \{0, 1, \ldots, 20\}$.

La probabilidad solicitada se corresponde con el subconjunto de $\{X \geq 1\} = \{1, \ldots 20\}$, por consiguiente

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{20}{0} 0.03^0 \cdot (1 - 0.03)^{20} = 0.4562$$

EJERCICIO 3 (Continuación)

b) La cadena recibe 10 cargamentos en un mes y el inspector prueba aleatoriamente 20 dispositivos por cargamento. ¿Cuál es la probabilidad de que haya exactamente tres cargamentos que contengan al menos un dispositivo defectuoso de entre los 20 seleccionados y probados?

Solución

En este caso cada cargamento puede o no contener al menos un artículo defectuoso. Por lo tanto, el hecho de probar el resultado de cada cargamento puede considerarse como un experimento de Bernoulli con $p = 0.4562$ del inciso anterior. Si suponemos la independencia de un cargamento a otro, y si se denotamos con Y el número de cargamentos que contienen al menos un artículo defectuoso, Y sigue otra distribución binomial $Y \sim B(10, 0.4562)$. Por lo tanto

$$P(Y = 3) = \binom{10}{3} 0.4562^3 \cdot (1 - 0.4562)^7 = 0.1602$$

EJERCICIO 4:

En cierta comunidad rural con infinidad de pozos, se conjetura que hay impurezas en el 30% del total de pozos de agua potable. Para obtener información sobre la verdadera magnitud del problema se determina que debe realizarse algún tipo de prueba. Como es muy costoso probar todos los pozos del área, se eligen 10 al azar para someterlos a la prueba.

¿Cuál es la probabilidad de que menos de 6 pozos tengan impurezas?

Solución

Sea X la variable aleatoria que cuenta el número de pozos con impurezas de una muestra de 10 pozos, asumiendo una probabilidad de 0.3 de que un pozo tenga impurezas, podemos utilizar el modelo binomial $X \sim B(10, 0.3)$ y la probabilidad demandada es $P(X < 6)$:

$$P(X < 6) = P(X \leq 5) = \sum_{x=0}^{5} \binom{10}{x} 0.3^x \cdot 0.7^{10-x} = 0.9527$$

EJERCICIO 5:

Consideramos la idea de que el 30% de los pozos de cierta comunidad rural tienen impurezas en el agua potable es sólo una conjetura del consejo local del agua. Supongamos que se eligen 10 pozos de forma aleatoria y resulta que 6 contienen impurezas. ¿Qué implica esto respecto de la conjetura?

Solución

Primero debemos preguntar: ''Si la conjetura es correcta, ¿podríamos haber encontrado 6 o más pozos con impurezas''

$$P(X \geq 6) = \sum_{x=6}^{10} \binom{10}{x} 0.3^x \cdot 0.7^{10-x} = 0.0473$$

En consecuencia, es poco probable (4.7% de probabilidad) que se encontrara que 6 o más pozos contenían impurezas si sólo 30% de ellos las contienen. Esto pone seriamente en duda la conjetura y sugiere que el problema de la impureza es mucho más grave.

Esquema

Problema 1

Un ingeniero de control de tráfico señala que el 75% de los vehículos que pasan por un punto de verificación son de ese estado. Se pide:

a. ¿Cuál es la probabilidad de que menos de 4 de los siguientes 9 vehículos sean de otro estado?
b. Calcula la media de vehículos que no son de ese estado que pasan por el punto de verificación para una muestra de 20 vehículos.

Problema 2

Un estudio a nivel nacional que examinó las actitudes hacia los antidepresivos reveló que aproximadamente 70% de los encuestados cree que "los antidepresivos en realidad no curan nada, sólo disfrazan el problema real". De acuerdo con este estudio, se pide:

a. ¿cuál es la probabilidad de que al menos 3 de las siguientes 10 personas seleccionadas al azar tengan esta opinión?
b. ¿cuál es la probabilidad de que menos 3 de las siguientes 10 personas seleccionadas al azar no tengan esta opinión?
c. Calcula la media y la varianza de los apartados anteriores.

Problema 3

Se sabe que 60% de los ratones inoculados con un suero quedan protegidos contra cierta enfermedad. Si se inoculan 5 ratones, calcula la probabilidad de que

a. ninguno contraiga la enfermedad;
b. menos de 2 contraigan la enfermedad;
c. la media de ratones que quedan protegidos contra la enfermedad para una muestra de 5 inoculados.

Problema 4

Supongamos que los motores de un avión operan de forma independiente y que tienen una probabilidad de falla de 0.4. Se supone que un avión tiene un vuelo seguro si funcionan al menos la mitad de sus motores. Si un avión tiene 4 motores y otro tiene 2, ¿cuál de los dos tiene la probabilidad más alta de un vuelo exitoso?

Problema 5

Unas figurillas de porcelana se venden a 10 euros si no tienen imperfección, y a 3 euros si la presentan. Entre las figurillas de cierta compañía, el 90% no tiene imperfecciones y el 10% sí tiene. En una muestra de 100 figurillas ya vendidas, sea Y el ingreso ganado por su venta y X el número de éstas que no presenta imperfecciones.

a. Expresa Y como una función de X.
b. Determina μ_Y.
c. Determina σ_Y.