Funciones de distribución discreta

Gamaliel Moreno Chávez

MCPI

Ago-Dic
2020

A menudo las observaciones que se generan mediante diferentes experimentos estadísticos tienen el mismo tipo general de comportamiento. En consecuencia, las variables aleatorias discretas asociadas con estos experimentos se pueden describir esencialmente con la misma distribución de probabilidad y, por lo tanto, es posible representarlas usando una sola fórmula. De hecho, se necesitan sólo unas cuantas distribuciones de probabilidad importantes.

Con frecuencia un experimento consta de pruebas repetidas, cada una con dos resultados posibles que se pueden denominar éxito o fracaso. El proceso se conoce como proceso de Bernoulli y cada ensayo se denomina experimento de Bernoulli.

El proceso de Bernoulli

Solo hay dos resultados: un "éxito"($X = 1$) con probabilidad p o un "fracaso"($X = 0$) con probabilidad $q = 1 - p$. El valor de la variable aleatoria X se utiliza como indicador del resultado. Por ejemplo, en un solo lanzamiento de moneda, $\mathrm{X} = 1$ está asociado con la aparición, o la presencia de la característica, de una cara, y $\mathrm{X} = 0$ con una cruz, o la ausencia de una cara.

Distribución binomial

La función de probabilidad de esta variable aleatoria se puede expresar como

$$p_1 = P(X = 1) = p, p_0 = P(X = 0) = q,$$

El valor esperado de X es

$$\mu = E(x) = (0)(q) + (1)(p) = p,$$

Y la varianza

$$\sigma^2 = V(X) = E(X^2) - \mu^2 = (0^2)(q) + (1^2)(p) - p^2 = pq.$$

Distribución binomial

Supongamos ahora que estamos interesados en variables aleatorias asociadas con repeticiones independientes de experimentos de Bernoulli, cada una con una probabilidad de éxito, p. Considere primero la distribución de probabilidad de una variable aleatoria X que es el número de éxitos en un número fijo de ensayos independientes, n. Si hay k éxitos y $n-k$ fracasos en n ensayos, entonces cada secuencia de 1 y 0 tiene la probabilidad $P(X=k)=p^k q^{n-k}$. El número de formas en que se pueden organizar x éxitos en n ensayos es la expresión binomial.

$$\frac{n!}{k!(n-k)!} \quad \text{también se expresa con} \quad \binom{n}{k}$$

Distribución binomial

Dado que cada una de estas secuencias mutuamente excluyentes ocurre con probabilidad $p^k q^{n-k}$, la función de probabilidad de esta variable aleatoria viene dada por

$$b(k; n, p) = p_k = \binom{n}{k} p^k q^{n-k} \quad k = 0, 1, 2, \ldots, n$$

Se le conoce como binomial porque el término $(n+1)$ tiene correspondencia con la expansión del binomio $(p+q)^2$

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

La media y la varianza son np y npq respectivamente, la cual es n veces la de Bernoulli.

Con frecuencia nos interesamos en problemas donde se necesita obtener $P(X < r)$ o $P(a \leq X \leq b)$. Estas se obtienen con sumatorias binomiales

Ejercicios

1. Considere 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. Considere que la probabilidad de defecto es 0.25. Encuentre los ocho resultados posibles, los valores correspondientes de X, tabule la función de distribución y calcule la probabilidad de que sean exactamente dos defectuosos usando la función de distribución.
2. La probabilidad de que cierta clase de componente sobreviva a una prueba de choque es de 3/4. Calcule la probabilidad de que sobrevivan exactamente 2 de los siguientes 4 componentes que se prueben.

Ejercicios

3. 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 a) sobrevivan al menos 10, b) sobrevivan de 3 a 8, y c) sobrevivan exactamente 5?
4. 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? b) Suponga que el detallista recibe 10 cargamentos en un mes y que 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?

Ejercicios

5. Se conjetura que hay impurezas en 30 % del total de pozos de agua potable de cierta co- munidad rural. 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. a) Si se utiliza la distribución binomial, ¿cuál es la probabilidad de que exactamente 3 pozos tengan impurezas, considerando que la conjetura es correcta? b) ¿Cuál es la probabilidad de que más de 3 pozos tengan impurezas?
6. Calcule la media y la varianza de la variable aleatoria binomial del ejemplo 3

Distribución multinomial

El experimento binomial se convierte en un **experimento multinomial** si cada prueba tiene más de dos resultados posibles.

Distribución multinomial

En general, si un ensayo dado puede tener como consecuencia cualquiera de los k resultados posibles $E_1, E_2, \ldots, E_k$ con probabilidades $p_1, p_2, \ldots, p_k$, la distribución multinomial dará la probabilidad de que E_1 ocurra x_1 veces, E_2 ocurra x_2 veces... y E_k ocurra x_k veces en n ensayos independientes, donde $x_1 + x_2 + \cdots + x_k = n$.

Distribución multinomial

Para derivar la fórmula general procedemos como en el caso binomial. El número total de ordenamientos esta dado por

$$\binom{n}{x_1, x_2 \ldots, x_k} = \frac{n!}{x_1! x_1! \cdots x_k!}$$

Como todas las particiones son mutuamente excluyentes y tienen la misma probabilidad de ocurrir, obtenemos la distribución multinomial multiplicando la probabilidad para un orden específico por el número total de particiones.

Distribución multinomial

Distribución multinomial

Si un ensayo dado puede producir los k resultados $E_1, E_2, \ldots, E_k$ con probabilidades $p_1, p_2, \ldots, p_k$, entonces la distribución de probabilidad de las variables aleatorias $X_1, X_2, ldots, X_k$, que representa el número de ocurrencias para $E_1, E_2, \ldots, E_k$ en n ensayos independientes, es

$$f(x_1, x_2, \cdots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2 \ldots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$$

con

$$\sum_{i=1}^{k} x_i = n \quad \sum_{i=1}^{k} p_i = 1$$

Ejemplo
La complejidad de las llegadas y las salidas de los aviones en un aeropuerto es tal que a menudo se utiliza la simulación por computadora para modelar las condiciones "ideales". Para un aeropuerto específico que tiene tres pistas se sabe que, en el escenario ideal, las probabilidades de que las pistas individuales sean utilizadas por un avión comercial que llega aleatoriamente son las siguientes:

$$\text{Pista 1: p } 1 = 2/9;\ \text{Pista 2: p } 2 = 1/6;\ \text{Pista 3: p } 3 = 11/18$$

¿Cuál es la probabilidad de que 6 aviones que llegan al azar se distribuyan de la siguien- te manera?

Pista 1: 2 aviones; Pista 2: 1 avión; Pista 3: 3 aviones

Distribución multinomial

Solución: Si usamos la distribución multinomial, tenemos

$$
\begin{aligned}
f(2,1,3;\frac{2}{9},\frac{1}{6},\frac{11}{18},6) &= \binom{6}{2,1,3}(2/9)^2(1/6)^1(11/18)^3 \\
&= \frac{6!}{2!1!3!}\frac{2^2}{9^2}\frac{1}{6}\frac{11^3}{18^3} = 0.1127
\end{aligned}
$$

Distribución hipergeométrica

Los tipos de aplicaciones de la distribución hipergeométrica son muy similares a los de la distribución binomial. Nos interesa el cálculo de probabilidades para el número de observaciones que caen en una categoría específica. La distribución hipergeométrica no requiere independencia y se basa en el muestreo que se realiza sin reemplazo.

Distribución hipergeométrica

En general, nos interesa la probabilidad de seleccionar x éxitos de los k artículos considerados éxitos y $n - x$ fracasos de los $N - k$ artículos que se consideran fracasos cuando una muestra aleatoria de tamaño n se selecciona de N artículos. Esto se conoce como un experimento hipergeométrico; es decir, aquel que posee las siguientes dos propiedades:

- De un lote de N artículos se selecciona una muestra aleatoria de tamaño n sin reemplazo.
- k de los N artículos se pueden clasificar como éxitos y $N - k$ se clasifican como fracasos.

la distribución de probabilidad de la variable hipergeométrica se conoce como distribución hipergeométrica, y sus valores se denotan con $h(x; N, n, k)$

Si deseamos calcular la probabilidad de obtener 3 cartas rojas en 5 extracciones de una baraja ordinaria de 52 cartas.Si se sacan 5 cartas al azar, nos interesa la probabilidad de seleccionar 3 cartas rojas de las 26 disponibles y 2 de las 26 cartas negras de que dispone la baraja.

$$\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = \frac{(26!/3!23!)(26!/2!24!)}{52!/5!47!} = 0.3251$$

Distribución hipergeométrica

La distribución de probabilidad de la variable aleatoria hipergeométrica X, el número hipergeométrica de éxitos en una muestra aleatoria de tamaño n que se selecciona de N artículos, en los que k se denomina éxito y N – k fracaso, es

$$h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

donde $mx\{0, n-(N-k)\} \leq x \leq mn\{n, k\}$.

Distribución hipergeométrica

La media y la varianza de la distribución hipergeométrica $h(x; N, n, k)$ son

$$\mu = \frac{nk}{N} \quad \sigma^2 = \frac{N-n}{N-1} n \frac{k}{N} (1 - k/N)$$

Ejercicios.

1. En una empresa donde trabajan 20 personas, hay 7 que fuman, si se seleccionan a 4 personas al azar ¿Cuál es la probabilidad que al menos una fume?
2. Lotes con 40 componentes cada uno que contengan 3 o más defectuosos se consideran inaceptables. El procedimiento para obtener muestras del lote consiste en seleccionar 5 componentes al azar y rechazar el lote si se encuentra un componente defectuoso. ¿Cuál es la probabilidad de, que en la muestra, se encuentre exactamente un componente defectuoso, si en todo el lote hay 3 defectuosos?

Distribución binomial negativa

Consideremos un experimento con las mismas propiedades de un experimento binomial, sólo que en este caso las pruebas se repetirán hasta que ocurra un número fijo de éxitos. Por lo tanto, en vez de encontrar la probabilidad de x éxitos en n pruebas, donde n es fija,ahora nos interesa la probabilidad de que ocurra el $k-$ ésimo éxito en la $x-$ ésima prueba. Los experimentos de este tipo se llaman experimentos binomiales negativos.

Distribución binomial negativa

Como ejemplo, considere el uso de un medicamento que se sabe que es eficaz en el 60 % de los casos en que se utiliza. El uso del medicamento se considerará un éxito si proporciona algún grado de alivio al paciente. Nos interesa calcular la probabilidad de que el quinto paciente que experimente alivio sea el séptimo paciente en recibir el medicamento en una semana determinada. Si designamos un éxito con E y un fracaso con F, un orden posible para alcanzar el resultado que se desea es EFEEEFE, que ocurre con la siguiente probabilidad

$$(0.6)(0.4)(0.6)(0.6)(0.6)(0.4)(0.6) = (0.6)^5(0.4)^2.$$

$$P(X = 7) = \binom{6}{4}(0.6)^5(0.4)^2 = 0.1866.$$

Distribución binomial negativa

Distribución binomial negativa

Si ensayos independientes repetidos pueden dar como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1\breve{}p$, entonces la distribución de probabilidad negativa de la variable aleatoria X, el número del ensayo en el que ocurre el $k-$ ésimo éxito, es

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k} \quad x = k, k+1, k+2, \ldots$$

Ejercicios

1. En la serie de campeonato de la NBA, el equipo que gane 4 de 7 juegos será el ganador. Suponga que los equipos A y B se enfrentan en los juegos de campeonato y que el equipo A tiene una probabilidad de 0.55 de ganarle al equipo B.
    a. ¿Cuál es la probabilidad de que el equipo A gane la serie en 6 juegos?
    b. ¿Cuál es la probabilidad de que el equipo A gane la serie?
    c. Si ambos equipos se enfrentaran en la eliminatoria de una serie regional y el triunfador fuera el que ganara 3 de 5 juegos, ¿cuál es la probabilidad de que el equipo A gane la serie?

Distribución geométrica

Si consideramos el caso especial de la distribución binomial negativa, donde $k = 1$, tenemos una distribución de probabilidad para el número de ensayos que se requieren para un solo éxito. Un ejemplo sería lanzar una moneda hasta que salga una cara. Nos podemos interesar en la probabilidad de que la primera cara resulte en el cuarto lanzamiento. En este caso la distribución binomial negativa se reduce a la forma

$$b^*(x; 1, p) = pq^{x-1}$$

A este caso especial se le conoce como distribución geométrica

Distribución geométrica

Distribución geométrica

Si pruebas independientes repetidas pueden tener como resultado un éxito con probabilidad p y un fracaso con probabilidad $q = 1\breve{}p$, entonces la distribución de probabilidad de la variable aleatoria X, el número de la prueba en el que ocurre el primer éxito, es

$$g(x; p) = pq^{x-1}, \quad x = 1, 2, 3, \ldots$$

Ejemplo Se sabe que en cierto proceso de fabricación uno de cada 100 artículos, en promedio, resulta defectuoso. ¿Cuál es la probabilidad de que el quinto artículo que se inspecciona, en un grupo de 100, sea el primer defectuoso que se encuentra?

Solución:

Si utilizamos la distribución geométrica con $x = 5$ y $p = 0.01$, tenemos

$$g(5, 0.01) = (0.01)(0.99)^4 = 0.0096$$

Proceso de Poisson

1. El número de resultados que ocurren en un intervalo o región específica es independiente del número que ocurre en cualquier otro intervalo de tiempo o región del espacio disjunto. De esta forma vemos que el proceso de Poisson no tiene memoria.
2. La probabilidad de que ocurra un solo resultado durante un intervalo de tiempo muy corto o en una región pequeña es proporcional a la longitud del intervalo o al tamaño de la región, y no depende del número de resultados que ocurren fuera de este intervalo de tiempo o región.
3. La probabilidad de que ocurra más de un resultado en tal intervalo de tiempo corto o que caiga en tal región pequeña es insignificante.

Distribución de Poisson

La distribución de probabilidad de la variable aleatoria de Poisson X, la cual representa el número de resultados que ocurren en un intervalo de tiempo dado o región específicos y se denota con t, es

Distribución de Poisson

$$p(x; \lambda t) = \frac{e^{\lambda} t (\lambda t)^x}{x!}, \quad x = 1, 2, 3, \ldots$$

donde λ es el número promedio de resultados por unidad de tiempo, distancia, área o volumen y e = 2.71828...

La distribución de probabilidad de la variable aleatoria de Poisson X, la cual representa el número de resultados que ocurren en un intervalo de tiempo dado o región específicos y se denota con t, es

Distribución de Poisson

$$p(x; \lambda t) = \frac{e^{\lambda} t(\lambda t)^{x}}{x!}, \quad x = 1, 2, 3, \ldots$$

donde λ es el número promedio de resultados por unidad de tiempo, distancia, área o volumen y e = 2.71828...

Tanto la media como la varianza de la distribución de Poisson $p(x; \lambda t)$ son λt.

Distribución de Poisson

Ejercicios

1. Durante un experimento de laboratorio el número promedio de partículas radiactivas que pasan a través de un contador en un milisegundo es 4. ¿Cuál es la probabilidad de que entren 6 partículas al contador en un milisegundo dado?
2. El número promedio de camiones-tanque que llega cada día a cierta ciudad portuaria es 10. Las instalaciones en el puerto pueden alojar a lo sumo 15 camiones-tanque por día. ¿Cuál es la probabilidad de que en un día determinado lleguen más de 15 camiones y se tenga que rechazar algunos?