Probabilidad y Estadística (93.24)
Variables aleatorias discretas

Índice

1. Repaso de algunos conceptos **2**

2. Guía de ejercicios **6**

3. Respuestas **14**

4. Ejercicios resueltos **16**

1. Repaso de algunos conceptos

Variable aleatoria discreta.

Se dice que una variable aleatoria X es discreta si toma valores en un conjunto discreto denominado *recorrido* y denotado como $\mathcal{R}_X$. El comportamiento de una variable aleatoria discreta (v.a.d.) está determinado por su *distribución de probabilidad* o *función de masa de probabilidad* p_X definida como

$$\mathrm{p}_X(x) = \mathrm{P}(X = x) \qquad \forall x \in \mathcal{R}_X. \tag{1}$$

La distribución de probabilidad cumple con la condición

$$\sum_{x \in \mathcal{R}_X} \mathrm{p}_X(x) = 1. \tag{2}$$

Otra forma de equivalente de definir el comportamiento de una v.a.d. es a través de la *función de distribución de probabilidad acumulada* definida como

$$\mathrm{F}_X(x) = \sum_{\{y \in \mathcal{R}_X :\, y \leq x\}} \mathrm{p}_X(y) \qquad \forall x \in \mathbb{R}. \tag{3}$$

Es fácil mostrar que F_X cumple siempre las siguientes condiciones:

- Es monótona no decreciente (porque $\mathrm{p}_X(x) \geq 0 \ \forall x \in \mathcal{R}_X$).
- $\mathrm{F}_X(-\infty) = \lim\limits_{x \to -\infty} \mathrm{F}_X(x) = 0$.
- $\mathrm{F}_X(+\infty) = \lim\limits_{x \to +\infty} \mathrm{F}_X(x) = 1$.

Se define el *valor esperado de una función $g(\cdot)$ de una variable aleatoria discreta* como

$$\mathrm{E}\left[g(X)\right] = \sum_{x \in \mathcal{R}_X} g(x)\mathrm{p}_X(x), \tag{4}$$

siempre que la suma en el lado de la derecha converja. En particular, los *momentos* de una v.a.d. vienen definidos por

$$\mathrm{E}\left[X^k\right] = \sum_{x \in \mathcal{R}_X} x^k \mathrm{p}_X(x) \qquad \forall k \in \mathbb{N}. \tag{5}$$

Cuando existe $\mathrm{E}[X]$, también se definen los *momentos centrados* como

$$\mathrm{E}\left[(X - \mathrm{E}[X])^k\right] = \sum_{x \in \mathcal{R}_X} (x - \mathrm{E}[X])^k \mathrm{p}_X(x) \qquad \forall k \in \mathbb{N}. \tag{6}$$

Dos parámetros importantes de toda v.a.d. son su *valor esperado* o *media* y su *varianza* dados por

$$\mu_X = \mathrm{E}[X] = \sum_{x \in \mathcal{R}_X} x\mathrm{p}_X(x), \tag{7}$$

$$\sigma_X^2 = \mathrm{Var}[X] = \mathrm{E}\left[(X - \mu_X)^2\right] = \mathrm{E}\left[X^2\right] - \left(\mathrm{E}[X]\right)^2, \tag{8}$$

respectivamente.

Teoría de la decisión

Una aplicación importante de la matemática es en teoría de la decisión, esto es, el estudio del comportamiento de un agente (un individuo, por ej.) frente a eventos inciertos. ¿Qué decisión tomar? ¿Cuál es la más conveniente? ¿Cómo se encuentra un camino óptimo a seguir? Estas son algunas de las preguntas contestadas por la teoría de la decisión.
Determinar la decisión óptima depende del problema, el acercamiento personal al mismo, etc. Un criterio muy común es el de maximización del valor esperado: se toma aquella decisión que maximiza el valor esperado de la ganancia o retorno.

Variable aleatoria Bernoulli

Sea X la variable aleatoria discreta correspondiente al éxito $(=1)$ o fracaso $(=0)$ en un experimento con probabilidad p de éxito. Luego, se dice que X tiene una distribución Bernoulli con parámetro p, $X \sim \text{Bernoulli}(p)$. (Ver Ej. 38 de la guía 1).
El recorrido de X es $\mathcal{R}_X = \{0, 1\}$. La función de masa de probabilidades está dada por:

$$\mathrm{P}(X=0) = (1-p), \qquad \mathrm{P}(X=1) = p. \tag{9}$$

Los primeros dos momentos son:

$$\mu_X = \mathrm{E}[X] = p, \qquad \sigma_X^2 = \mathrm{Var}[X] = p(1-p). \tag{10}$$

Variable aleatoria binomial

Sea X la variable aleatoria discreta correspondiente al número de éxitos en una secuencia de n experimentos independientes cuyos únicos resultados posibles son éxito o fracaso. Cada uno de los experimentos tiene una probabilidad p de éxito. Luego, se dice que X tiene una distribución binomial con parámetros n y p, $X \sim \text{Binomial}(n, p)$. (Ver Ej. 38 de la guía 1).
El recorrido de X es $\mathcal{R}_X = \{0, 1, \cdots, n\}$. La función de masa de probabilidad está dada por:

$$\mathrm{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad \text{para } k = 0, 1, \cdots, n. \tag{11}$$

Los primeros dos momentos son:

$$\mu_X = \mathrm{E}[X] = np, \qquad \sigma_X^2 = \mathrm{Var}[X] = np(1-p). \tag{12}$$

Hecho: $X \sim \text{Bernoulli}(p) \leftrightarrow X \sim \text{Binomial}(1, p)$.

Variable aleatoria Poisson

Sea X una variable aleatoria discreta con recorrido $\mathcal{R}_X = \mathbb{N}^0$ $(= \{0, 1, 2, \cdots\})$. Se dice que X tiene una distribución Poisson con parámetro λ, $X \sim \text{Poisson}(\lambda)$, si su función de masa de probabilidad está dada por:

$$\mathrm{P}(X=k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad \text{para } k = 0, 1, \cdots \tag{13}$$

Los primeros dos momentos son:

$$\mu_X = \mathrm{E}[X] = \lambda, \qquad \sigma_X^2 = \mathrm{Var}[X] = \lambda. \tag{14}$$

Variable aleatoria hipergeométrica

Considere una población con N elementos, M de los cuales son de un tipo dado. Se toma una muestra de tamaño n al azar sin reposición. Sea X la variable aleatoria discreta correspondiente al número de elementos en la muestra que son del tipo especial. Luego, se dice que X tiene una distribución hipergeométrica con parámetros N, M y n, $X \sim \text{Hipergeo}(N, M, n)$. El recorrido de X es $\mathcal{R}_X = \{\text{máx}\{0, n-(N-M)\}, \cdots, \text{mín}\{n, M\}\}$. La función de masa de probabilidad está dada por:

$$\mathrm{P}(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \qquad \text{para } k \in \mathcal{R}_X. \tag{15}$$

Los primeros dos momentos son:

$$\mu_X = \mathrm{E}[X] = n\frac{M}{N}, \qquad \sigma_X^2 = \mathrm{Var}[X] = n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}. \tag{16}$$

Variable aleatoria geométrica

Sea X la variable aleatoria discreta correspondiente al número de fracasos hasta obtener el primer éxito en una secuencia de experimentos independientes cuyos únicos resultados posibles son éxito o fracaso. Cada uno de los experimentos tiene una probabilidad p de éxito. Luego, se dice que X tiene una distribución geométrica con parámetro p, $X \sim \text{Geométrica}(p)$.
El recorrido de X es $\mathcal{R}_X = \mathbb{N}^0$. La función de masa de probabilidad está dada por:

$$\mathrm{P}(X = k) = p(1-p)^k \qquad \text{para } k = 0, 1, \cdots, n. \tag{17}$$

Los primeros dos momentos son:

$$\mu_X = \mathrm{E}[X] = \frac{1-p}{p}, \qquad \sigma_X^2 = \mathrm{Var}[X] = \frac{1-p}{p^2}. \tag{18}$$

Es interesante notar que

$$\mathrm{P}\big(X \geq k+m \,|\, X \geq k\big) = \mathrm{P}(X \geq m) \qquad \forall k, m \in \mathbb{N}^0. \tag{19}$$

Es decir: si la espera hasta el primer éxito ha sido al menos k, la probabilidad de esperar al menos otros m experimentos es la misma que la probabilidad de que la espera sea de al menos m desde el inicio. Esto se denomina la propiedad de la *falta de memoria* de una variable aleatoria geométrica.

Variable aleatoria binomial negativa

Sea X la variable aleatoria discreta correspondiente al número de fracasos hasta obtener r éxitos en una secuencia de experimentos independientes cuyos únicos resultados posibles son éxito o fracaso. Cada uno de los experimentos tiene una probabilidad p de éxito. Luego, se dice que X tiene una distribución binomial negativa con parámetros r y p, $X \sim \text{BinoNeg}(r, p)$.
El recorrido de X es $\mathcal{R}_X = \mathbb{N}^0$. La función de masa de probabilidad está dada por:

$$\mathrm{P}\left(X = k\right) = \binom{k+r-1}{k} p^r (1-p)^k \qquad \text{para } k = 0, 1, \cdots, n. \tag{20}$$

Los primeros dos momentos son:

$$\mu_X = \mathrm{E}\left[X\right] = r\frac{1-p}{p}, \qquad \sigma_X^2 = \mathrm{Var}\left[X\right] = r\frac{1-p}{p^2}. \tag{21}$$

Hecho: $X \sim \text{Geométrica}(p) \leftrightarrow X \sim \text{BinoNeg}(1, p)$.

Aproximación binomial a la hipergeométrica

Sean $X_N \sim \text{Hipergeo}(N, M(N), n)$ variables aleatorias tales que

$$\lim_{N\to\infty} \frac{M(N)}{N} = p, \tag{22}$$

con $p \in (0, 1)$. Luego

$$\lim_{N\to\infty} \mathrm{P}\left(X_N = k\right) = \mathrm{P}\left(Y = k\right) \qquad \forall k \in \mathcal{R}_Y, \tag{23}$$

donde $Y \sim \text{Binomial}(n, p)$.
Es decir que, cuando el tamaño de la población es grande en relación al tamaño de la muestra ($N \gg n$), una variable aleatoria hipergeométrica puede aproximarse por una binomial. Intuitivamente, no hace mucha diferencia muestrear con o sin reposición.

Aproximación Poisson a la binomial

Sean $X_n \sim \text{Binomial}(n, p(n))$ variables aleatorias tales que

$$\lim_{n\to\infty} n \times p(n) = \lambda, \tag{24}$$

con $\lambda > 0$. Luego

$$\lim_{n\to\infty} \mathrm{P}\left(X_n = k\right) = \mathrm{P}\left(Y = k\right) \qquad \forall k \in \mathcal{R}_Y, \tag{25}$$

donde $Y \sim \text{Poisson}(\lambda)$.
Es decir que, cuando el número de experimentos es grande, una variable aleatoria binomial puede aproximarse por una Poisson.

2. Guía de ejercicios

1. Considere el experimento de lanzar un par de dados no cargados.
Sean las variables aleatorias X: suma de los puntos obtenidos e Y: máximo puntaje de los dos obtenidos. Obtenga las tablas de la función de probabilidad y la función de distribución de estas variables aleatorias discretas. Calcule las siguientes probabilidades:

   a) $\mathrm{P}(X < 7)$ $\quad \mathrm{P}(2 \leq X < 8)$ $\quad \mathrm{P}(X \geq 9)$;

   b) $\mathrm{P}(Y \leq 3)$ $\quad \mathrm{P}(2 < Y \leq 5)$ $\quad \mathrm{P}(Y > 3)$.

2. Suponga que la demanda diaria de un artículo es una variable aleatoria D con recorrido $\mathcal{R}_D = \{1, 2, 3, 4\}$ y función de probabilidad $\mathrm{p}_D(r) = C\, 2^r / r!$.

   a) Calcule la constante C.

   b) Obtenga y represente gráficamente la función de distribución F_D.

   Resolución aquí.

3. *a*) Determine el valor esperado y la varianza de las variables aleatorias X e Y del ejercicio 1.

   b) Determine el valor esperado $\mathrm{E}[D]$ de la variable aleatoria D de la demanda del ejercicio anterior.

4. Se selecciona al azar una muestra sin reemplazo de 3 artículos de un total de 10 de los cuales 2 son defectuosos. Si X es la variable aleatoria: número de artículos defectuosos en la muestra, obtenga la distribución de probabilidades de X y calcule sus parámetros característicos. ¿Cómo cambia el problema si el muestreo es con sustitución?

5. Un fabricante de controladoras de discos rígidos somete cada unidad a una prueba rigurosa. De las controladoras recién ensambladas, el 84 % pasa la prueba sin ninguna modificación. Las que fallan en la prueba inicial son reelaboradas; de éstas, el 75 % pasa una segunda prueba. Aquellas controladoras que fallan en la segunda prueba se rehacen por segunda ocasión y se vuelven a probar; el 90 % de ellas pasan la prueba y el resto se desarman.
Se define X como la variable aleatoria que corresponde al número de veces que debe reprocesarse una controladora seleccionada al azar.

   a) Especificar el recorrido de X y obtener la distribución de probabilidad.

   b) Calcular el valor esperado de X. ¿Cómo se interpreta este número?

   c) Calcular la varianza y el desvío estándar de X.

   d) ¿Cuál es el porcentaje de controladoras que se desarman?

   Resolución aquí. Video con el ejercicio explicado.

6. En una estación de servicio, la distribución de clientes que llegan cada 15 minutos tiene la siguiente distribución de probabilidades:

Nro. de clientes	0	1	2
probabilidad	0.2	0.5	0.3

La probabilidad de que un cliente pague con tarjeta de crédito es 0.25, y cada cliente tiene una forma de pago independiente de la de los demás.

a) Obtenga la distribución de probabilidades de la variable aleatoria Y que indica la cantidad de clientes que en 15 minutos pagan con tarjeta de crédito.

b) Calcule el valor esperado y la dispersión de la variable aleatoria Y.

7. Una organización de consumidores que evalúa automóviles nuevos reporta regularmente el número de defectos importantes en cada automóvil examinado. Sea X el número de defectos importantes en un automóvil seleccionado al azar de un cierto tipo y $\mathrm{F}_X(x)$ la función de distribución de probabilidad correspondiente.

$$\mathrm{F}_X(x) = \begin{cases} 0 & x < 0 \\ 0.06 & 0 \leq x < 1 \\ 0.19 & 1 \leq x < 2 \\ 0.39 & 2 \leq x < 3 \\ 0.67 & 3 \leq x < 4 \\ 0.92 & 4 \leq x < 5 \\ 0.97 & 5 \leq x < 6 \\ 1 & x \geq 6 \end{cases}$$

a) Calcular las siguientes probabilidades directamente de la función dada: $\mathrm{p}_X(2) = \mathrm{P}(X = 2)$, $\mathrm{P}(X > 3)$, $\mathrm{P}(2 \leq X \leq 5)$, y $\mathrm{P}(2 < X < 5)$.

b) Representar gráficamente $\mathrm{p}_X(x)$.

c) ¿Cuál es el número esperado de defectos importantes que se espera al examinar un automóvil, seleccionado al azar, del tipo considerado en este problema?

8. La probabilidad de falla de un cierto tipo de componentes electrónicos es de 0.10. Una compañía produce con ellos dos tipos de circuitos, denominados I y II respectivamente. El circuito tipo I consiste en un paralelo de dos componentes. El circuito tipo II está armado con una serie de dos componentes. De la producción total se elige un circuito de cada tipo. Sea X la variable aleatoria que indica el número de circuitos que funciona cuando se prueban ambos.

a) Definir la función de probabilidad p_X, hallar la función de probabilidad y representar ambas funciones en forma gráfica.

b) Hallar el valor esperado de X.

Resolución aquí. Video con el ejercicio explicado.

9. Para la situación planteada en el problema de la demanda diaria D suponga que la utilidad U de la venta por artículo es 5 a (la unidad monetaria) y que cualquier artículo que no se venda produce una pérdida de 3 a. El fabricante produce diariamente k artículos.

a) Obtenga la distribución de probabilidades de la utilidad neta diaria U para distintos valores de k (1,2,3,4,5,...) y elija el valor de k que maximice la utilidad esperada $E[U]$.

b) Suponga que de alguna manera se pudiese determinar cuál será la demanda de un día (pagando por la información, desarrollando algún tipo de campaña, etc). ¿Cuánto pagaría como máximo para obtener la información?
Sugerencia: calcule la ganancia esperada considerando que se fabrica exactamente lo que se demanda.

10. Un contratista debe elegir entre dos obras. La primera promete una ganancia de \$ 240000 con una probabilidad de 0.75 o una pérdida de \$ 60000 (por inconvenientes varios) con una probabilidad de 0.25. La segunda obra promete una ganancia de \$ 360000 con probabilidad 0.5 o una pérdida de \$ 90000 con probabilidad 0.5.

a) ¿Cuál debería elegir el contratista si quiere maximizar la ganancia esperada?

b) ¿Cuál sería la obra que elegiría si su negocio anduviera mal y quebrara a menos que lograra una ganancia de $ 300000 en su próxima obra?

11. Un vendedor de diarios compra cada periódico a 40 centavos y lo vende a 1 peso y no puede devolver los diarios que no vendió. La demanda diaria es independiente de la del día anterior y tiene la siguiente distribución de probabilidades:

cantidad demandada	63	64	65	66	67	68	69	70
probabilidad	0.01	0.04	0.06	0.08	0.15	0.28	0.22	0.16

¿Cuántos diarios debe adquirir diariamente si desea maximizar la ganancia esperada? (Tenga en cuenta que la insatisfacción de la demanda no está penalizada.)

Resolución aquí.

12. De un mazo de naipes ordinarios (52 naipes) bien barajado se extraen cinco cartas, al azar (sin sustitución), para una mano de poker.

a) Obtenga la distribución de probabilidades del número de diamantes D en la mano.

b) Determine el valor esperado y el desvío estándar de D.

c) Calcule la probabilidad de sacar por lo menos un trébol.

d) Calcule la probabilidad de sacar por lo menos dos ases.

13. En una semana de trabajo se realizaron 50 facturas en un comercio. En 5 de ellas se cometió un error.

a) Se eligen 10 facturas al azar. Determinar la distribución de probabilidades del número de facturas que contienen el error buscado. Representar gráficamente la función de probabilidad del número de facturas que contienen el error buscado. Calcular el valor esperado y la varianza del número de facturas que contienen el error buscado

b) Suponga que se eligen 10 facturas al azar y no se conoce cuantas facturas hay con errores entre las 50 de esa semana. Se extrae la muestra y se obtiene una factura con error. Se conviene en estimar la cantidad de facturas con error como el valor que maximiza la probabilidad de obtener una factura con error entre las 10. Probar (confeccionando una tabla) que la estimación de la cantidad de facturas con error entre las 50 es 5.

14. De un grupo de N personas se eligen al azar a n de ellas.

a) Calcular la probabilidad de que se encuentren las r más altas entre las n elegidas ($r < n < N$).

b) La mitad de las N personas (N es par) habla inglés. Calcular la probabilidad de que la mitad de las n (n también es par) personas de la muestra aleatoria hable inglés.

c) Suponga que $N = 100$. Verificar que si el muestreo se realiza con sustitución la probabilidad de que la mitad de las n (n es par) personas de la muestra aleatoria hable inglés viene dada por

$$b(n) = \binom{n}{\frac{n}{2}} 0.5^n. \tag{26}$$

Si $hg(n)$ es la probabilidad de que la mitad de las n (n es par) personas de la muestra aleatoria hable inglés, para el caso real (muestreo sin sustitución), verificar (confeccionando una tabla) que se cumple $|hg(n) - b(n)| < 0.1\, hg(n)$ si $n < 18$.

15. Un inspector de control de calidad está inspeccionando artículos recientemente producidos para buscar defectos. El inspector inspecciona un artículo en una serie de etapas independientes, cada una con duración fija. Cuando una falla está presente, la probabilidad de que sea detectada durante cualquiera de las etapas es 0.90. Sea X la cantidad de etapas necesarias para detectar una falla presente. Obtener el recorrido de X, la función de probabilidad y calcular $\mathrm{E}[X]$ y $\mathrm{Var}[X]$.
Son conocidas:

$$\sum_{k=1}^{\infty} k\, q^{k-1} = \frac{1}{(1-q)^2}, \qquad \sum_{k=1}^{\infty} k^2\, q^{k-1} = \frac{1+q}{(1-q)^3}, \quad |q| < 1.$$

16. El contenido de bolillas rojas de la caja C_1 se forma como sigue: se arroja un dado y se colocan tantas bolillas rojas como indica el dado; luego se extraen dos bolillas (sin reemplazo) de una caja C_2, que contiene 3 blancas y 7 rojas, y se introducen en C_1. Suponga que la caja C_1 está inicialmente vacía. Obtenga la función de probabilidad del número R de bolillas rojas que quedan finalmente en C_1. Determine $E[R]$.

17. Considere la variable aleatoria X: *cantidad de espadas en una mano de truco.*

   a) Obtenga la función de probabilidad $\mathrm{p}_X(x)$ y la función de distribución $\mathrm{F}_X(x)$.

   b) Para jugar se paga \$ 2 y el premio es recibir en \$ el cuadrado de la cantidad de espadas obtenidas. Si Y es la ganancia neta calcule el valor esperado $E[Y]$.

18. La probabilidad de que un cierto examen dé una reacción positiva es igual a 0.4 y las reacciones son independientes. Obtenga la función de probabilidad del número N de reacciones negativas antes de la primera reacción positiva. ¿Cuál es la probabilidad de que ocurran menos de cinco reacciones negativas consecutivas antes de la primera positiva?

19. Sobre una placa **I** inciden partículas que, al chocar, pueden rebotar con probabilidad p_1 ó atravesarla con probabilidad $p_2 = 1 - p_1$. Las partículas que atraviesan la placa **I** impactan contra la placa **II** (paralela a la **I**) con probabilidad 1 de rebotar y por lo tanto vuelven a golpear contra el reverso de la placa **I** donde valen las probabilidades p_1 de rebotar en la **I** y p_2 de atravesarla hacia el medio de ingreso. Obtenga la función de probabilidad de la cantidad de veces N que una partícula rebota contra la placa **II**.

20. Un hombre tiene un llavero con 6 llaves. Ha olvidado cuál es la de su casa y las va probando una por una.

   a) Sea X el número de intentos necesarios hasta abrir la puerta. Halle la función de probabilidad de X y su valor esperado.

   b) Supongamos que este hombre está totalmente borracho y en cada intento vuelve a elegir una llave al azar de entre las 6, en lugar de separar las que ya probó. Halle la función de probabilidad de X, y extraiga conclusiones sobre los beneficios de la sobriedad.

21. Se determinó, a partir de numerosas experiencias previas, que de cada 5 fusibles que produce una máquina 1 es defectuoso. Calcular la probabilidad de que en una muestra aleatoria (extraída de un lote de gran tamaño) de 4 fusibles se obtengan:

   a) uno defectuoso;

   b) como máximo dos defectuosos;

   c) ninguno defectuoso;

   d) los cuatro defectuosos.

Resolución aquí.

22. Supongamos que diez trabajadores van a usar intermitentemente energía eléctrica y que nos interesa estimar la carga total con la que se va a operar. Supondremos que, en un instante dado, cada trabajador tiene aproximadamente la misma probabilidad p de requerir una unidad de energía. Si, en promedio, un trabajador usa la energía durante 12 minutos por hora, tomaremos $p = 0.2$ (ya que 12 minutos es el 20 % de una hora) .
Calcular la probabilidad de que 7 o más trabajadores requieran energía eléctrica al mismo tiempo (lo que equivale a calcular la probabilidad de sobrecarga cuando el suministro se ajusta a 6 unidades de energía).

23. Los motores de una producción se someten a dos controles independientes A y B. Se sabe, de experiencias previas, que el 1 % de los motores de este tipo falla en la prueba A y el 2 % en la prueba B.

   a) Calcular la probabilidad de que un motor probado falle en por lo menos uno de los controles.

   b) Si se selecciona al azar 20 de esos motores (se supone que el total de los motores es mucho mayor que 20), calcular la probabilidad de que fallen en por lo menos uno de los controles:

      1) ningún motor;

      2) a lo sumo el 10 % de los motores de la muestra.

   Representar gráficamente la función de probabilidad $\mathrm{p}_X(x)$ de la variable aleatoria X : número de motores que fallan en por lo menos uno de los controles.

24. La máquina A produce diariamente el doble de artículos que la máquina B; el 4 % de los artículos producidos por la máquina A tiende a ser defectuoso, mientras que para la máquina B el porcentaje de defectuosos es del 2 %. Se combina la producción diaria de ambas máquinas y se toma una muestra aleatoria de 10 artículos.

   a) ¿Cuál es la probabilidad de que la muestra contenga exactamente 2 artículos defectuosos?

   b) Determinar el valor esperado y el desvío estándar del número de artículos defectuosos en esa muestra aleatoria de 10 artículos.

   Resolución aquí. Video con el ejercicio explicado.

25. Una producción de lámparas eléctricas tiene un porcentaje de defecto del 1 %. Se prueba una muestra de 100 lámparas y si hay a lo sumo una defectuosa se acepta la producción; si hay dos o más defectuosas se toma otra muestra de 100 lámparas. Si en esta muestra hay a lo sumo una defectuosa se acepta la producción y en caso contrario se rechaza. Calcular la probabilidad de aceptación de la producción.

26. Una proporción p de un gran número de artículos en un lote es defectuosa. Se extrae una muestra de m artículos y si ésta no contiene items defectuosos, el lote es aceptado, en tanto que si contiene más de dos items defectuosos, el lote es rechazado. Si la muestra contiene uno o dos artículos defectuosos, se extrae una muestra independiente de tamaño m y si el número combinado de defectuosas en las dos muestras no excede de dos, el lote es aceptado. Calcular la probabilidad de aceptar este lote.

27. Un comprador y un vendedor acuerdan un plan de muestreo de tamaño $n = 10$ para hacer el control de calidad. El comprador acepta el lote del que se extrae la muestra sólo si el número de artículos defectuosos es 0, 1 o 2. En caso contrario la rechaza. Llamemos p a la probabilidad de que un artículo sea defectuoso y supongamos que el tamaño de la población de la que se extrae la muestra es lo suficientemente grande para suponer que las observaciones muestrales son independientes. Calcular la probabilidad de que el comprador acepte el lote para los siguientes valores de p: (a) 0.1, (b) 0.3 y (c) 0.5. Representar gráficamente la probabilidad de aceptación como función de p.

28. Sea p la probabilidad de que cualquier símbolo particular de un código se transmita erróneamente a través de un sistema de comunicaciones. Suponga que en diferentes símbolos ocurren errores independientemente uno de otro. Suponga también que con probabilidad p_2 un símbolo erróneo se corrige al recibirse. Sea X_n el número de símbolos correctos recibidos en un bloque de mensaje formado por n símbolos (una vez que el proceso de corrección haya terminado).

    a) Demostrar que X_n es una variable aleatoria discreta con distribución binomial.

    b) Determinar el valor esperado y la varianza de X_n si $n = 10$, $p = 0.02$, y $p_2 = 0.95$.

    Resolución aquí. Video con el ejercicio explicado.

29. Se supone que un proceso de fabricación de fusibles produce un 1 % de fusibles defectuosos. El proceso se controla periódicamente examinando 10 fusibles, y, si alguno falla, se detiene el proceso para revisarlo.

    a) ¿Cuál es la probabilidad de detener el proceso con el sistema de control implementado?

    b) ¿Cúantos fusibles deberán probarse (en vez de 10) si se desea que la probabilidad de detener el proceso para revisarlo sea 0.95 cuando esté produciendo 5 % de fusibles defectuosos (que es un porcentaje de defectos muy por encima del aceptable)?

    c) Con el tamaño de muestra obtenido en el punto (b), ¿cuál es la probabilidad de detener el proceso cuando esté produciendo al 1 % de fusibles defectuosos?

    Resolución aquí. Video con el ejercicio explicado.

30. Una partícula esta sometida a la acción de impulsos aleatorios, pudiéndose mover en la dirección del eje x. Supongamos conocida la probabilidad de un salto unitario hacia la derecha es p y la probabilidad de un salto unitario hacia la izquierda $q = 1 - p$. Este proceso aleatorio se denomina un *Random Walk* o caminata aleatoria unidimensional.

    a) Obtener una expresión para la probabilidad que la partícula alcance cierta abscisa x al cabo de n *saltos*. Suponga que la posición inicial es 0.
    Sugerencia: Si la partícula se encuentra en la posición x tras n saltos de los cuales r fueron a la derecha y l a la izquierda entonces se tiene: $r + l = n$ y $r - l = x$.

    b) Sea X_n la posición de la partícula al cabo de n *saltos*. Si $p = 0.5$ y $X_0 = 0$ calcular las siguientes probabilidades:

        1) $\mathrm{P}(X_n \geq 0)$ para n tomando los valores de 1 a 4.
        2) $\mathrm{P}(|X_n| \leq 2)$ para n tomando los valores de 1 a 4.

    Resolución aquí. Video con el ejercicio explicado.

31. Un dispositivo está compuesto de un gran número de elementos que trabajan independientemente con igual probabilidad (pequeña) de falla de cada elemento durante un tiempo T. Hallar el promedio de elementos que fallan en el tiempo T, si la probabilidad de que en ese tiempo falle por lo menos un elemento es igual a 0.999.

32. Una batería de artillería dispara a un blanco y se sabe que la probabilidad de acertar es p =0.01. ¿Cuántos disparos tendrá que realizar para tener probabilidad mayor que 90 % de dar en el blanco por lo menos una vez? Resolver usando las distribuciones binomial y de Poisson. Generar una tabla con la razón de las respuestas obtenidas con una y otra distribución para valores de p desde 0.1 hasta 0.01 en pasos de 0.005.

    Resolución aquí.

33. El número promedio de personas que realizan transacciones en un cajero automático cada 10 minutos es de 3.4 personas. El número de personas que acude al cajero automático se puede suponer una variable aleatoria con distribución de Poisson. Calcular la probabilidad de que en 10 minutos cualesquiera, se realicen:

    a) menos de 2 transacciones.

    b) más de 2 transacciones.

34. Una fábrica envía al depósito 500 artículos. La probabilidad de deterioro de un artículo en el camino es igual a 0.002. Hallar la probabilidad de que en el camino se deterioren:

    a) exactamente 3 artículos;

    b) menos de 3 artículos;

    c) más de 3 artículos;

    d) por lo menos uno.

35. En una población el 1 % de las personas padece de daltonismo. ¿Cuál debe ser el tamaño de una muestra aleatoria (con reemplazo), de manera tal que la probabilidad de que la misma contenga por lo menos un daltónico sea mayor o igual a 0.95? Resuelva usando las distribuciones binomial y de Poisson.

36. La probabilidad de que se haga una soldadura defectuosa en una determinada conexión de una tubería es 10^{-4}.

    a) Hallar el valor esperado y la desviación típica de la variable aleatoria X: número de soldaduras defectuosas realizadas, si el número de soldaduras que requiere el tendido de la tubería es 5×10^4.

    b) Calcular la probabilidad de que en dicho trabajo no haya ninguna soldadura defectuosa.

37. Suponga un sistema de comunicación óptica que transmite dos símbolos, **1** y **0**. Como transmisor, se utiliza un láser que se enciende (para enviar un **1**) y se apaga (para enviar un **0**). El número de fotones durante el tiempo que está encendido el láser sigue una distribución de Poisson con media λ. Como receptor, se tiene un dispositivo ideal que puede contar cada fotón que llega. De esta forma, si cuenta al menos un fotón, el receptor decide que se envió un **1**; caso contrario, decide que se envió un **0**. Si se transmiten tantos **1**s como **0**s, determine el valor de λ para que la tasa de error en la decisión del receptor sea menor a 10^{-9}.

38. (Tomado del libro de Devore) Un biólogo recogió 20 especímenes de protistas, 10 de los cuales son Amoeba Proteus y 10 son Chaos Carolinensis. Un asistente toma 15 de los especímenes para un estudio más detallado.

    a) Describa la función de masa de probabilidad del número de Chaos Carolinensis seleccionados.

b) ¿Cuál es la probabilidad de que todos los especímenes de uno de los dos tipos hayan sido seleccionados?

c) ¿Cuál es la probabilidad de que el número de especímenes de Chaos Carolinensis se aleje de su valor medio en no más de un desvío estándar?

39. Repita el ejercicio anterior, pero suponiendo que el biólogo recogió 2×10^6 especímenes de protistas, de los cuales la mitad son Amoeba Proteus y la otra mitad son Chaos Carolinensis.

40. Suponga que el 5 % de los individuos de la población mundial presentan talasemia.

    a) Se toman personas al azar, una tras otra. ¿Cuál es el número medio de personas elegidas antes de que aparezca la primera con talasemia? ¿Cuál es la probabilidad de que las 12 primeras no tengan talasemia?

    b) Se toman personas al azar, una tras otra. ¿Cuál es el número medio de personas elegidas antes de que aparezca la tercera con talasemia? ¿Cuál es la probabilidad de que entre las 36 primeras no haya 3 con talasemia?

    c) ¿Cuál es la probabilidad de que en un grupo de 12 personas tomadas al azar no haya ninguna con talasemia? ¿Cuál es la probabilidad de que no haya 3 personas con talasemia en un grupo de 36?

41. El 20 % de los exámenes que corrige un docente tienen el mismo error (utilizan $\overline{x}_c$ en una prueba de hipótesis para la media con desvío desconocido).

    a) ¿Cuál es la probabilidad de que el docente vea 10 exámenes antes de encontrar el primero con ese error?

    b) Suponga que el docente ya vio 10 exámenes y no encontró el error. ¿Cuál es la probabilidad de que vea 10 exámenes más hasta encontrar ese error por primera vez?

3. Respuestas

2. $C = \frac{1}{6}$

3. a) $\mathrm{E}[X] = 7$ $\mathrm{Var}[X] = 5.833$ $\mathrm{E}[Y] = 4.47$ $\mathrm{Var}[Y] = 1.917$ b) $\mathrm{E}[D] = \frac{19}{9}$

4. Sin reposición: $\mathrm{E}[X] = 0.6$ $\mathrm{Var}[X] = 0.373$, con reposición: $\mathrm{E}[X] = 0.6$ $\mathrm{Var}[X] = 0.48$.

5. a) $\mathrm{P}(X = 0) = 0.84$ $\mathrm{P}(X = 1) = 0.12$ $\mathrm{P}(X = 2) = 0.04$
b) $\mathrm{E}[X] = 0.2$, cada 5 unidades, en promedio se espera una reelaboración
c) $\mathrm{Var}[X] = 0.24$ d) 0.4 % (40 por cada 10000).

6. a) ((0, 0.743759), (1, 0.2375), (2, 0.01875)) b) $\mathrm{E}[X] = 0.275$ $\mathrm{Var}[X] \approx 0.487$.

7. a) $\mathrm{P}(X = 2) = 0.2$ $\mathrm{P}(X > 3) = 0.33$ $\mathrm{P}(2 \leq X \leq 5) = 0.78$ $\mathrm{P}(2 < X < 5) = 0.53$.
c) $\mathrm{E}[X]$ =2.80.

8. a) $\mathrm{P}(X = 0) = 0.0019$ $\mathrm{P}(X = 1) = 0.1962$ $\mathrm{P}(X = 2) = 0.8019$ b) $\mathrm{E}[X] = 1.8$.

9. a) Conviene fabricar 2.

10. a) La primera b) La segunda.

11. 68.

12. b) $\mathrm{E}[X] = 1.25$, $\mathrm{Var}[X] = 0.864$ c) 0.778 d) 0.042.

13. a) {(0, 0.3106), (1, 0.4313), (2, 0.2098), (3, 0.0442), (4, 0.0040), (5, 0.0001)}.

15. La función de probabilidad viene dada por $0.9\ 0.1^{k-1}, k : 1, 2, 3, \ldots$, $\mathrm{E}[X] = \frac{10}{9}$, $\mathrm{Var}[X] = \frac{10}{81}$.

16. (1, $\frac{1}{90}$), (2, $\frac{4}{45}$), (k, $\frac{1}{6}$) para k entre 3 y 6, (7, $\frac{7}{45}$), (8, $\frac{7}{90}$).

17. a) (0 ,$\frac{203}{494}$), (1, $\frac{435}{988}$),(2 , $\frac{135}{988}$) y (3, $\frac{3}{247}$). b) $\mathrm{E}[Y] = -\frac{47}{52}$ = - 0.9038.

18. La probabilidad de que ocurran menos de cinco reacciones negativas consecutivas antes de la primera positiva es 0.9222. La función de probabilidad del número N de reacciones negativas antes de la primera reacción positiva viene dada por $p_N(k) = 0.4\ 0.6^{\ k}$ para $k = 0, 1, \ldots$..

19. La función de probabilidad de N es $p_N(0) = p_1$ y $p_N(k) = p_2^2\, p_1^{k-1}$ para $k = 1, 2, \ldots$..

20. a) El número de intentos puede tomar los valores 1 a 6 con probabilidad constante $\frac{1}{6}$.
b) Esa variable aleatoria tiene distribución geométrica de parámetro $\frac{1}{6}$.

21. a) 0.41 b) 0.97 c) 0.41 d) 0.0016.

22. 0.00086 (1 cada 1157 minutos).

23. a) 0.0298 b1) 0.546 b2) 0.98

24. 0.038 .

25. 0.93.

26. $p_0 + p_0\,p_1 + p_1^2 + p_2\,p_0$, donde $p_0 = (1-p)^m$, $p_1 = m\,p\,(1-p)^{m-1}$ y $p_2 = (m/2)\,(m-1)\,p^2\,(1-p)^{m-2}$.

27. a) 0.93 b) 0.382 c) 0.054.

28. b) 9.99 0.01

29. a) 0.0956 b) 59 c) 0.4473.

30. b1) 0.5, 0.75, 0.5, 0.6875 b2) 1, 1, 0.75, 0.875.

31. Fallan 7 elementos.

32. $n = 230$.

33. a) 0.1468 b) 0.6603.

34. a) 0.0613 b) 0.92 c) 0.019 d) 0.632.

35. $n \geq 299$ usando la distribución binomial y $n \geq 300$ si se usa la distribución de Poisson.

36. a) 5, 2.236 b) 0.007.

38. b) 0.0325 c) 0.6966.

39. b) 0.0 c) 0.6982.

40. a) 19, 0.5404 b) 57, 0.7321 c) 0.5404, 0.7321.

41. a) 0.0215 b) 0.0215.

4. Ejercicios resueltos

Ejercicio 2

Suponga que la demanda diaria de un artículo es una variable aleatoria D con recorrido $R(D) = \{1, 2, 3, 4\}$ y función de probabilidad $\mathrm{P}(r) = C\, 2^r / r!$. Calcule la constante C y calcule el valor esperado de D.

Resolución

La función de probabilidad de una variable aleatoria discreta cumple la condición que la suma de sus valores para todo el recorrido debe ser 1. Si se plantea esa condición para esta variable entonces resulta:

$$C\left(2 + \frac{4}{2} + \frac{8}{6} + \frac{16}{24}\right) = 1. \tag{27}$$

Resolviendo para C resulta $C = \frac{1}{6}$. Con ese valor de C se pueden obtener las probabilidades correspondientes a los 4 valores posibles de D. Se resume la información de la distribución de probabilidades en la siguiente tabla:

Demanda	1	2	3	4
Probabilidad	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$

El valor esperado de D es:

$$\mathrm{E}[D] = 1 \times \frac{1}{3} + 2 \times \frac{1}{3} + 3 \times \frac{2}{9} + 4 \times \frac{1}{9} = \frac{19}{9}. \tag{28}$$

La varianza de D es:

$$\mathrm{Var}[D] = \mathrm{E}\left[D^2\right] - (\mathrm{E}[D])^2 = 1^2 \times \frac{1}{3} + 2^2 \times \frac{1}{3} + 3^2 \times \frac{2}{9} + 4^2 \times \frac{1}{9} - (\frac{19}{9})^2 = \frac{80}{81}. \tag{29}$$

Ejercicio 5

Un fabricante de controladoras de discos rígidos somete cada unidad a una prueba rigurosa. De las controladoras recién ensambladas, el 84 % pasa la prueba sin ninguna modificación. Las que fallan en la prueba inicial son reelaboradas; de éstas, el 75 % pasa una segunda prueba. Aquellas controladoras que fallan en la segunda prueba se rehacen por segunda ocasión y se vuelven a probar; el 90 % de ellas pasan la prueba y el resto se desarman. Se define X como la variable aleatoria que corresponde al número de veces que debe reprocesarse una controladora seleccionada al azar.

a) Especificar el recorrido de X y obtener la distribución de probabilidad.

b) Calcular el valor esperado de X. ¿Cómo se interpreta este número?

c) Calcular la varianza y el desvío estándar de X

d) ¿Cuál es el porcentaje de controladoras que se desarman?

Resolución

Variable Aleatoria

A partir de esta guía, en lo personal sugiero, además de definir eventos y datos como en la primer guía, agregar la definición de las variables aleatorias. En este caso, la variable aleatoria ya está definida:

$$X = \text{Cantidad de veces que se reprocesa una controladora seleccionada al azar} \tag{30}$$

Eventos

Definiremos tres eventos pero basados en la misma idea:

$$A_i = \text{La controladora seleccionada pasa la prueba } i. \quad (1 \leq i \leq 3) \tag{31}$$

Datos

Para los datos hay que tener en cuenta que cada prueba sólo se llevaba a cabo si falló la anterior. Por ejemplo, si la controladora se sometió a la tercera prueba, es porque falló la primera y la segunda.

- $\mathrm{P}\left(A_1\right) = 0.84$
- $\mathrm{P}\left(A_2 \middle| \bar{A}_1\right) = 0.75$
- $\mathrm{P}\left(A_3 \middle| \bar{A}_2 \cap \bar{A}_1\right) = 0.9$

Ítem a

Para este problema hay que estar muy atentos al enunciado, porque se puede confundir la variable X que definimos previamente con:

$$X = \text{Cantidad de pruebas que falla una controladora seleccionada.} \tag{32}$$

Sin embargo, eso puede llevar a pensar que una controladora puede llegar a reprocesarse 3 veces si falla la tercer prueba. Vale aclarar, que si la controladora falla la tercer prueba, no se reprocesa, **se desarma**.

Por lo tanto, la variable aleatoria sigue siendo

$$X = \text{Cantidad de veces que se } \textbf{reprocesa} \text{ una controladora seleccionada al azar} \tag{33}$$

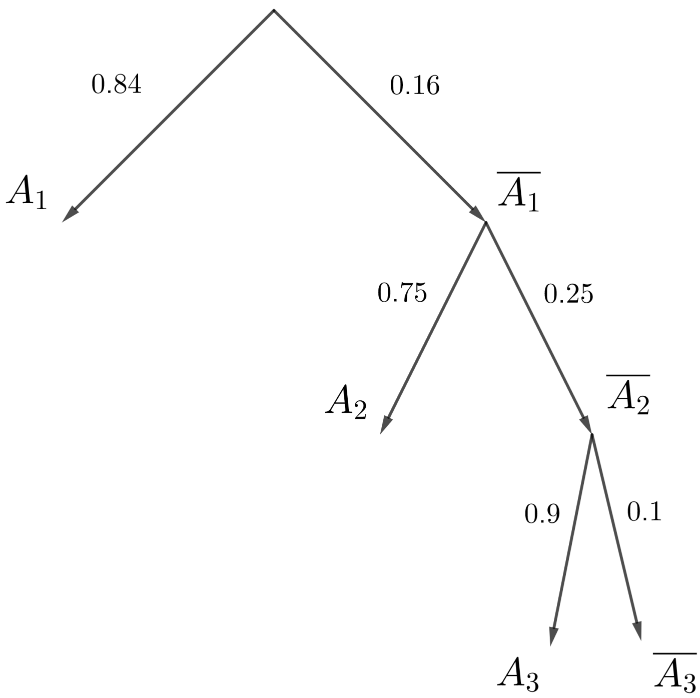
y en consecuencia, su recorrido es $R_X = \{0; 1; 2\}$ ya que puede reprocesarse como máximo 2 veces.

Nos piden la distribución de probabilidad de la variable X. Con este fin, basta conocer las probabilidades de que la variable X tome los valores de su recorrido, es decir:

- $\mathrm{P}(X=0)$
- $\mathrm{P}(X=1)$
- $\mathrm{P}(X=2)$

Notar que si $X=0$, entonces la controladora no debió reprocesarse, eso significa que pasó la primera prueba. Dicho de otro modo, $\mathrm{P}(X=0)=\mathrm{P}(A_1)$. Del mismo modo, si $X=1$ eso significa que se reprocesó una única vez. Es decir, falló la primera prueba pero aprobó la segunda. Concluimos que $\mathrm{P}(X=1)=\mathrm{P}\left(\overline{A_1}\cap A_2\right)$ y por lo tanto, $\mathrm{P}(X=2)=\mathrm{P}\left(\overline{A_1}\cap\overline{A_2}\right)$.

Para calcular estas probabilidades, seguiremos el respectivo diagrama del árbol:



Por lo tanto,

- $\mathrm{P}(X=0)=\mathrm{P}(A_1)=0.84$
- $\mathrm{P}(X=1)=\mathrm{P}\left(A_2\cap\overline{A_1}\right)=\mathrm{P}\left(A_2\middle|\overline{A_1}\right)\cdot\mathrm{P}\left(\overline{A_1}\right)=0.16\cdot 0.75=0.12$
- $\mathrm{P}(X=2)=\mathrm{P}\left(\overline{A_2}\cap\overline{A_1}\right)=\mathrm{P}\left(\overline{A_2}\middle|\overline{A_1}\right)\cdot\mathrm{P}\left(\overline{A_1}\right)=0.16\cdot 0.25=0.04$

Expresado como tabla:

X	0	1	2
p_X	0.84	0.12	0.04

(34)

Notar que todas las probabilidades suman 1.

Ítem b

Ya teniendo el recorrido ($R_X=\{0,1,2\}$), el valor esperado se calcula del siguiente modo:

$$\mathrm{E}[X]=\sum_{i=0}^{2} i\cdot\mathrm{P}(X=i)=0\cdot 0.84+1\cdot 0.12+2\cdot 0.04=0.2 \tag{35}$$

Esto significa que en una producción de controladores, la cantidad promedio de reprocesamientos será aproximadamente 0.2. Es decir, si se producen 100 controladores, la cantidad de reprocesamientos será aproximadamente $20 = 0.2 \cdot 100$.

No confundir la **cantidad de reprocesamientos** con la **cantidad de controladoras reprocesadas**. En este ejemplo puntual, con **20 reprocesamientos**, supongamos que 5 procesadoras fueron reprocesadas más de una vez (es decir, dos veces). Entonces el **número de controladoras reprocesadas fueron 15**.

Ítem c

Nos piden calcular la varianza, que podemos calcular del siguiente modo:

$$\mathrm{Var}\left[X\right] = \underbrace{\mathrm{E}\left[X^2\right]}_{?} - \underbrace{\left(\mathrm{E}\left[X\right]\right)^2}_{0.2^2} \tag{36}$$

Por lo tanto, sólo debemos calcular $\mathrm{E}\left[X^2\right]$:

$$\mathrm{E}\left[X^2\right] = \sum_{i=0}^{2} i^2 \cdot \mathrm{P}\left(X = i\right) = 0^2 \cdot \mathrm{P}\left(X = 0\right) + 1^2 \cdot \mathrm{P}\left(X = 1\right) + 2^2 \cdot \mathrm{P}\left(X = 2\right) = 0.12 + 0.16 = 0.28 \tag{37}$$

Recopilando estos datos, $\mathrm{Var}\left[X\right] = \mathrm{E}\left[X^2\right] - \left[\mathrm{E}\left[X\right]\right]^2 = 0.28 - 0.2^2 = 0.24$.

Para el desvío, debemos calcular la raíz: $\sigma(X) = \sqrt{\mathrm{Var}\left[X\right]} = \sqrt{0.24} \approx 0.4898979$

Ítem d

Para este ítem, consideramos un nuevo evento:

$$D = \text{la controladora extraída es desarmada} \tag{38}$$

La única forma de que la controladora sea desarmada, es que falle las tres pruebas. Por lo tanto, podemos describir al evento D como $D = \overline{A_1} \cap \overline{A_2} \cap \overline{A_3}$.

Es decir,

$$\mathrm{P}\left(D\right) = \mathrm{P}\left(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}\right) = \mathrm{P}\left(\overline{A_3} \middle| \overline{A_2} \cap \overline{A_1}\right) \cdot \mathrm{P}\left(\overline{A_2} \middle| \overline{A_1}\right) \mathrm{P}\left(\overline{A_1}\right) = 0.16 \cdot 0.25 \cdot 0.1 = 0.004 \tag{39}$$

Entonces, se desarman el 0.4 % de las controladoras.

Ejercicio 8

La probabilidad de falla de un cierto tipo de componentes electrónicos es de 0.10. Una compañía produce con ellos dos tipos de circuitos, denominados I y II respectivamente. El circuito tipo I consiste en un paralelo de dos componentes. El circuito tipo II está armado con una serie de dos componentes. De la producción total se elige un circuito de cada tipo. Sea X la variable aleatoria que indica el número de circuitos que funciona cuando se prueban ambos.

a) Definir la función de probabilidad p_X, hallar la función de probabilidad acumulada y representar ambas funciones en forma gráfica.

b) Hallar el valor esperado de X.

Resolución

Variable aleatoria

Nuevamente, aquí está definida la variable aleatoria:

$$X = \text{cantidad de circuitos que funcionan} \tag{40}$$

Sabiendo que hay dos circuitos, podemos concluir que $R_X = \{0, 1, 2\}$.

Eventos

Podemos considerar los siguientes eventos:

- C_i = El componente i funciona. $(1 \leq i \leq 4)$
- F_I = funciona el circuito I
- F_{II} = funciona el circuito II

Datos

Siguiendo la figura 1, se puede concluir que $F_I = C_1 \cup C_2$ (debe funcionar alguno de los dos componentes), mientras que $F_{II} = C_3 \cap C_4$ (deben funcionar ambos componentes).

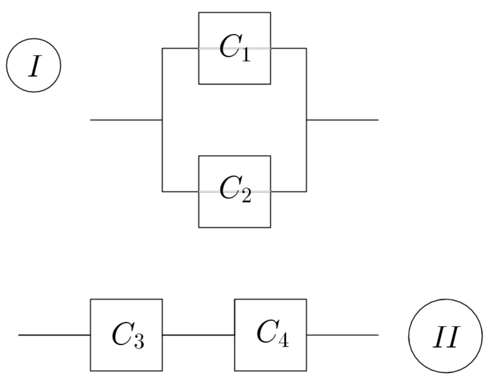


Figura 1: Descripción gráfica de ambos circuitos

Por otro lado, se sabe que los eventos C_i son independientes entre sí y $\mathrm{P}(C_i) = 0.9 \Rightarrow \mathrm{P}\left(\overline{C_i}\right) = 0.1$. Además, como F_I y F_{II} dependen de componentes distintos, también son independientes entre sí.

Ítem a

Nos piden la función de probabilidad puntual $\mathrm{p}_X(x)$, para eso, debemos sacar las probabilidades que involucran cada valor del recorrido de X. Es decir:

- $\mathrm{P}(X = 0)$
- $\mathrm{P}(X = 1)$
- $\mathrm{P}(X = 2)$

Para calcular estas probabilidades, podemos pensar los eventos en función de F_I y F_{II}:

- $\mathrm{P}(X = 0) = \mathrm{P}\left(\overline{F_I} \cap \overline{F_{II}}\right) \overset{IND}{=} \mathrm{P}\left(\overline{F_I}\right) \cdot \mathrm{P}\left(\overline{F_{II}}\right)$
- $\mathrm{P}(X = 1) = \mathrm{P}\left(\overline{F_I} \cap F_{II}\right) + \mathrm{P}\left(F_I \cap \overline{F_{II}}\right) \overset{IND}{=} \mathrm{P}\left(\overline{F_I}\right) \cdot \mathrm{P}(F_{II}) + \mathrm{P}(F_I) \cdot \mathrm{P}\left(\overline{F_{II}}\right)$
- $\mathrm{P}(X = 2) = \mathrm{P}(F_I \cap F_{II}) \overset{IND}{=} \mathrm{P}(F_I) \cdot \mathrm{P}(F_{II})$

Por lo tanto, calculando $\mathrm{P}(F_I)$ y $\mathrm{P}(F_{II})$, se pueden obtener las probabilidades de sus complementos y a partir de ellos construir la función de probabilidad puntual.

$$\begin{aligned}
\mathrm{P}(F_I) &= \mathrm{P}(C_1 \cup C_2) = \mathrm{P}(C_1) + \mathrm{P}(C_2) - \mathrm{P}(C_1 \cap C_2) && (41)\\
&\overset{IND}{=} \mathrm{P}(C_1) + \mathrm{P}(C_2) - \mathrm{P}(C_1) \cdot \mathrm{P}(C_2) && (42)\\
&= 0.9 + 0.9 - 0.9^2 = 0.99 && (43)
\end{aligned}$$

Por otro lado,

$$\mathrm{P}(F_{II}) = \mathrm{P}(C_3 \cap C_4) = \mathrm{P}(C_3) \cdot \mathrm{P}(C_4) = 0.9^2 = 0.81 \tag{44}$$

Además, de estos resultados se concluye que $\mathrm{P}\left(\overline{F_I}\right) = 0.01$ y $\mathrm{P}\left(\overline{F_{II}}\right) = 0.19$.

Notar que la probabilidad de que funcione el circuito I es mayor que la correspondiente al circuito II. Esto tiene sentido ya que para que funcione el circuito II deben funcionar ambos, restringiendo más su funcionamiento.

Por lo tanto,

- $\mathrm{P}(X = 0) = \mathrm{P}\left(\overline{F_I}\right) \cdot \mathrm{P}\left(\overline{F_{II}}\right) = 0.01 \cdot 0.19 = 0.0019$
- $\mathrm{P}(X = 1) = \mathrm{P}\left(\overline{F_I}\right) \cdot \mathrm{P}(F_{II}) + \mathrm{P}(F_I) \cdot \mathrm{P}\left(\overline{F_{II}}\right) = 0.99 \cdot 0.19 + 0.01 \cdot 0.81 = 0.1962$
- $\mathrm{P}(X = 2) = \mathrm{P}(F_I) \cdot \mathrm{P}(F_{II}) = 0.99 \cdot 0.81 = 0.8019$

Expresado como tabla:

x	0	1	2
$\mathrm{p}_X(x)$	0.0019	0.1962	0.8019

(45)

Notar que todas estas probabilidades suman 1, dando consistencia a nuestro resultado.

La función de probabilidad de una variable aleatoria discreta se grafica correspondiendo a cada punto del recorrido un segmento de altura igual a la probabilidad puntual respectiva. Notar que en la figura los segmentos más altos son los correspondientes a $X = 1$ y $X = 2$, la probabilidad $\mathrm{P}(X = 0)$ es tan insignificante respecto de las otras que no se aprecia en el gráfico. Sin embargo, siempre debe ser considerada.

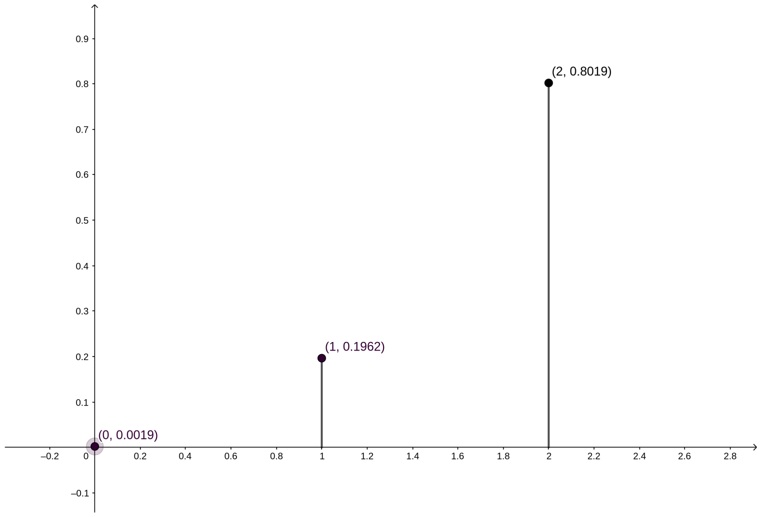


Figura 2: Función de probabilidad puntual

Además, se pide la función de desitribución acumulada. Para **cualquier** variable aleatoria, esta función se define como:

$$\mathrm{F}_X(t) = \mathrm{P}\left(X \leq t\right), \quad \underline{\forall t \in \mathbb{R}} \tag{46}$$

Si bien la la variable X tiene probabilidad cero para todo valor fuera del recorrido, esta función de distribución es válida para todo valor $t \in \mathbb{R}$. Es un error común considerar que sólo hay que definirla en los puntos de probabilidad positiva. Esto se debe a que esta función cambiará su valor únicamente en los valores del recorrido. Por lo tanto, para calcular F_X dividiremos en casos:

- Si $t < 0 \Rightarrow \mathrm{F}_X(t) = \mathrm{P}\left(X \leq t\right) = 0$ (no se acumula ningún punto de probabilidad positiva).
- Si $0 \leq t < 1 \Rightarrow \mathrm{F}_X(t) = \mathrm{P}\left(X \leq t\right) = \mathrm{P}\left(X = 0\right) = 0.0019$ (el único punto de probabilidad positiva es $X = 0$).
- Si $1 \leq t < 2 \Rightarrow \mathrm{F}_X(t) = \mathrm{P}\left(X \leq t\right) = \mathrm{P}\left(X = 0\right) + \mathrm{P}\left(X = 1\right) = 0.1981$ (los únicos valores del recorrido menores o iguales que t son $X = 0$ y $X = 1$).
- Si $t \geq 2 \Rightarrow \mathrm{F}_X(t) = \mathrm{P}\left(X \leq t\right) = \mathrm{P}\left(X = 0\right) + \mathrm{P}\left(X = 1\right) + \mathrm{P}\left(X = 2\right) = 1$ (se acumulan todos los valores del recorrido).

Es decir, la función de distribución se define del siguiente modo:

$$\mathrm{F}_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ 0.0019 & \text{si } 0 \leq t < 1 \\ 0.1981 & \text{si } 1 \leq t < 2 \\ 1 & \text{si } t \geq 2 \end{cases} \tag{47}$$

Por lo tanto, graficar F_X es lo mismo que graficar cualquier función en $\mathbb{R}$. Este gráfico se ve en la Figura 3, notar que los saltos de esta función se dan en los puntos de probabilidad positiva, y la magnitud del salto se corresponde con la probabilidad respectiva.

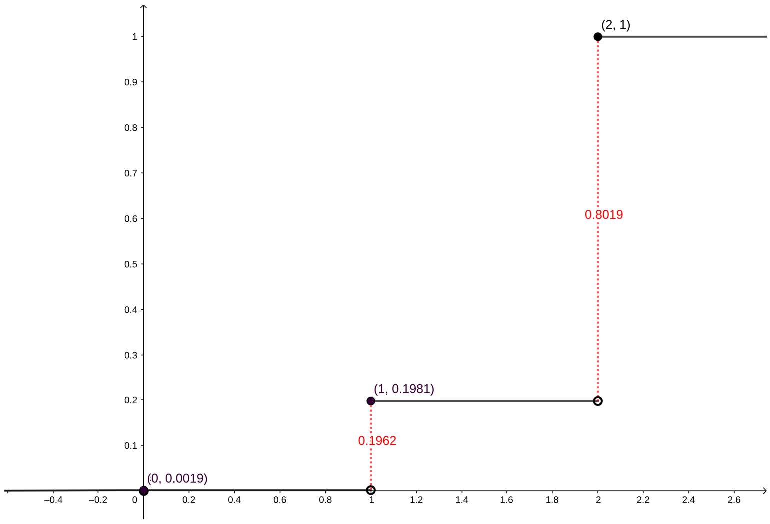


Figura 3: Función de distribución de X

Ítem b

Sabiendo que el recorrido de X es $R_X = \{0, 1, 2\}$, el valor esperado se calcula del siguiente modo:

$$\mathrm{E}\left[X\right] = 0 \cdot \mathrm{P}\left(X = 0\right) + 1 \cdot \mathrm{P}\left(X = 1\right) + 2 \cdot \mathrm{P}\left(X = 2\right) = 0 \cdot 0.0019 + 1 \cdot 0.1962 + 2 \cdot 0.8019 = 1.8 \quad (48)$$

Ejercicio 11

Un vendedor de diarios compra cada periódico a 40 centavos y lo vende a 1 peso y no puede devolver los diarios que no vendió. La demanda diaria es independiente de la del día anterior y tiene la siguiente distribución de probabilidades :

cantidad demandada	63	64	65	66	67	68	69	70
probabilidad	0.01	0.04	0.06	0.08	0.15	0.28	0.22	0.16

¿Cuántos diarios debe adquirir diariamente si desea maximizar la ganancia esperada? (Tenga en cuenta que la insatisfacción de la demanda no está penalizada.)

Resolución

Vamos a resolver este problema de dos maneras distintas, aunque completamente equivalentes.

Forma 1:

Llamemos

$$G_k = \text{Ganancia cuando el vendedor compra } k \text{ diarios } (k \geq 0). \tag{49}$$

Observe que estamos definiendo una variable aleatoria distinta para cada elección de la cantidad de diarios comprados. Esto implica que, para calcular el valor esperado de cada G_k, debemos determinar su recorrido y su distribución de probabilidades. Las Tablas 1-5 muestran los resultados, incluyendo el cálculo del valor esperado. De estas tablas se desprende que el vendedor debe comprar 68 diarios si desea maximizar la ganancia esperada.

Tabla 1: Ganancia G_{70}, G_{69}

G_{70}

x	p_X	$x \cdot \mathrm{p}_X(x)$
35.0	0.01	0.350
36.0	0.04	1.440
37.0	0.06	2.220
38.0	0.08	3.040
39.0	0.15	5.850
40.0	0.28	11.200
41.0	0.22	9.020
42.0	0.16	6.720
SUMA	1.00	39.840

G_{69}

x	$\mathrm{p}_X(x)$	$x \cdot \mathrm{p}_X(x)$
35.4	0.01	0.354
36.4	0.04	1.456
37.4	0.06	2.244
38.4	0.08	3.072
39.4	0.15	5.910
40.4	0.28	11.312
41.4	0.38	15.732
SUMA	1.00	40.080

Forma 2:

En este caso, utilizaremos conocimientos que se corresponden con la guía 3, más específicamente, sobre funciones de variables aleatorias. Nuevamente llamemos

$$X = \text{Número de diarios demandados}, \tag{50}$$
$$G_k = \text{Ganancia cuando el vendedor compra } k \text{ diarios } (k \geq 0). \tag{51}$$

Tabla 2: Ganancia G_{68}, G_{67}

G_{68}

x	p_X	$x \cdot p_X(x)$
35.8	0.01	0.358
36.8	0.04	1.472
37.8	0.06	2.268
38.8	0.08	3.104
39.8	0.15	5.970
40.8	0.66	26.928
SUMA $= E[G_{68}]$		40.100

G_{67}

x	p_X	$x \cdot p_X(x)$
36.2	0.01	0.362
37.2	0.04	1.488
38.2	0.06	2.292
39.2	0.08	3.136
40.2	0.81	32.562
SUMA	1.00	39.840

Tabla 3: Ganancia G_{66}, G_{65}

G_{66}

x	p_X	$x \cdot p_X(x)$
36.6	0.01	0.366
37.6	0.04	1.504
38.6	0.06	2.316
39.6	0.89	35.244
SUMA	1.00	39.430

G_{65}

x	p_X	$x \cdot p_X(x)$
37.0	0.01	0.370
38.0	0.04	1.520
39.0	0.95	37.050
SUMA	1.00	38.94

Tabla 4: Ganancia G_{64}, G_{63}

G_{64}

x	p_X	$x \cdot p_X(x)$
37.4	0.01	0.374
38.4	0.99	38.016
SUMA	1.00	38.390

G_{63}

x	p_X	$x \cdot p_X(x)$
37.80	1.00	37.80
SUMA	1.00	37.80

Tabla 5: Ganancia G_k cuando $k > 70$ o $k < 63$

G_k con $k > 70$

x	p_X	$x \cdot \mathrm{p}_X(x)$
$63 - 0.4k$	0.01	$0.630 - 0.004k$
$64 - 0.4k$	0.04	$2.560 - 0.016k$
$65 - 0.4k$	0.06	$3.900 - 0.024k$
$66 - 0.4k$	0.08	$5.280 - 0.032k$
$67 - 0.4k$	0.15	$10.050 - 0.060k$
$68 - 0.4k$	0.28	$19.040 - 0.112k$
$69 - 0.4k$	0.22	$15.180 - 0.088k$
$70 - 0.4k$	0.16	$11.200 - 0.064k$
SUMA	1.00	$67.840 - 0.400k \leq 39.44$

G_k con $k < 63$

x	p_X	$x \cdot \mathrm{p}_X(x)$
$0.6k$	1.00	$0.6k$
SUMA	$0.6k$	$0.6k \leq 37.2$

El recorrido $\mathcal{R}_X$ y la distribución de probabilidades de X están dados en la tabla del enunciado. G_k es una variable aleatoria cuyo valor depende de X. Es fácil ver que

$$G_k(X) = \begin{cases} 1 \cdot X - 0.40 \cdot k & \text{si } X < k, \\ 1 \cdot k - 0.40 \cdot k & \text{si } X \geq k. \end{cases} \tag{52}$$

Podemos encontrar, entonces, el valor esperado de la ganancia

$$\mathrm{E}\left[G_k\right] = \sum_{i \in \mathcal{R}_X} G_k(i) \mathrm{P}\left(X = i\right). \tag{53}$$

La Tabla 6 muestra los resultados en el caso de utilizar esta expresión para el cálculo del valor esperado de la ganancia. Como se desprende de dicha tabla, el vendedor debe comprar 68 diarios si desea maximizar el la ganancia esperada.

Tabla 6: Ganancia por la venta de diarios como función de la demanda

Prob.	0.01	0.04	0.06	0.08	0.15	0.28	0.22	0.16	
V.A.	**Recorrido**								$E[\cdot]$
X	63	64	65	66	67	68	69	70	67.84
G_k $k < 63$	$0.6k$	$0.6k$	$0.6k$	$0.6k$	$0.6k$	$0.6k$	$0.6k$	$0.6k$	$0.6k \leq 37.20$
G_{63}	37.8	37.8	37.8	37.8	37.8	37.8	37.8	37.8	37.80
G_{64}	37.4	38.4	38.4	38.4	38.4	38.4	38.4	38.4	38.39
G_{65}	37.0	38.0	39.0	39.0	39.0	39.0	39.0	39.0	38.94
G_{66}	36.6	37.6	38.6	39.6	39.6	39.6	39.6	39.6	39.43
G_{67}	36.2	37.2	38.2	39.2	40.2	40.2	40.2	40.2	39.84
G_{68}	35.8	36.8	37.8	38.8	39.8	40.8	40.8	40.8	40.10
G_{69}	35.4	36.4	37.4	38.4	39.4	40.4	41.4	41.4	40.08
G_{70}	35.0	36.0	37.0	38.0	39.0	40.0	41.0	42.0	39.84
G_k $k > 70$	$63 - 0.4k$	$64 - 0.4k$	$65 - 0.4k$	$66 - 0.4k$	$67 - 0.4k$	$68 - 0.4k$	$69 - 0.4k$	$70 - 0.4k$	$67.84 - 0.4k \leq 39.44$

Ejercicio 21

Se determinó, a partir de numerosas experiencias previas, que de cada 5 fusibles que produce una máquina 1 es defectuoso. Calcular la probabilidad de que en una muestra aleatoria (extraída de un lote de gran tamaño) de 4 fusibles se obtengan:

1. uno defectuoso,
2. como máximo dos defectuosos,
3. ninguno defectuoso,
4. los cuatro defectuosos.

Resolución

Con la información disponible se puede asignar la probabilidad $p = 0.2$ de extraer un fusible defectuoso al elegir un fusible al azar. La muestra aleatoria se extrae de un lote de gran tamaño de manera que pueda suponerse que las observaciones muestrales son independientes. Con estas consideraciones se puede suponer que la variable aleatoria discreta X, el número de fusibles defectuosos en la muestra aleatoria de tamaño 4, tiene distribución binomial de parámetros $p = 0.2$ y $n = 4$.

La función de probabilidad de la variable aleatoria discreta X viene dada por

$$p_k = \mathrm{P}\,(X = k) = \binom{4}{k}\, 0.2^k\, 0.8^{4-k}. \tag{54}$$

y variando el valor de k se obtiene la tabla

k	0	1	2	3	4
p_k	0.4096	0.4096	0.1536	0.0256	0.0016

A partir de la tabla se puede responder a lo solicitado:

$\mathrm{P}(X=1)$	$\mathrm{P}(X \leq 2)$	$\mathrm{P}(X=0)$	$\mathrm{P}(X=4)$
0.4096	0.9728	0.4096	0.0016

Ejercicio 24

La máquina A produce diariamente el doble de artículos que la máquina B; el 4 % de los artículos producidos por la máquina A tiende a ser defectuoso, mientras que para la máquina B el porcentaje de defectuosos es del 2 %. Se combina la producción diaria de ambas máquinas y se toma una muestra aleatoria de 10 artículos.

a) ¿Cuál es la probabilidad de que la muestra contenga exactamente 2 artículos defectuosos?

b) Determinar el valor esperado y el desvío estándar del número de artículos defectuosos en esa muestra aleatoria de 10 artículos.

Resolución

Variable aleatoria

La variable aleatoria es la siguiente:

$$X = \text{Cantidad de artículos defectuosos entre los 10 artículos extraídos.}$$

Eventos

Dado que un artículo puede provenir de la máquina A o de la máquina B, y que puede ser defectuoso o no, consideraremos los siguientes eventos:

- A= un artículo extraído proviene de la máquina A.
- B= un artículo extraído proviene de la máquina B.
- D = Un artículo defectuoso es extraído

Datos

Como dice que la máquina A produce el doble de artículos que la máquina B, en cada extracción la probabilidad de extraer un artículo de la máquina A será el doble de la probabilidad de extraer un artículo de la máquina B. Por lo tanto, $\mathrm{P}(A) = 2\mathrm{P}(B)$, como además $\mathrm{P}(A) + \mathrm{P}(B) = 1$, se tiene que $\mathrm{P}(A) = \frac{2}{3}$ y $\mathrm{P}(B) = \frac{1}{3}$.

Por otro lado, tenemos como dato que $\mathrm{P}\big(D|\, A\big) = 0.04$ y $\mathrm{P}\big(D|\, B\big) = 0.02$.

Ítem a

La consigna pide hallar $\mathrm{P}(X = 2)$. Para eso, nos conviene calcular la distribución de X.

El ejercicio supone que la muestra es suficientemente grande como para que extraer un defectuoso, no altera las probabilidades de extraer otro defectuoso en el siguiente paso. Por lo tanto, se puede asumir que la variable X consiste en contar la cantidad de defectuosos en 10 experimentos aleatorios independientes. La independencia surge de asumir que la probabilidad de extraer un defectuoso se mantiene constante. Es decir, la variable X se considera una variable aleatoria de distribución binomial:

$$X \sim \mathrm{Bi}(10, p)$$

Como sabemos que son 10 experimentos, el primer parámetro de la distribución fue sencillo. Ahora resta determinar p. Al contar la cantidad de defectuosos en la muestra, p equivale a la probabilidad de extraer un defectuoso (es decir, $\mathrm{P}(D)$).

Los datos probabilísticos de D los tenemos a partir de los eventos A y B. Por lo tanto, utilizando el teorema de probabilidad total:

$$\mathrm{P}(D) = \mathrm{P}\big(D|\, A\big) \cdot \mathrm{P}(A) + \mathrm{P}\big(D|\, B\big) \cdot \mathrm{P}(B) = 0.04 \cdot \frac{2}{3} + 0.02 \cdot \frac{1}{3} \approx 0.0333$$

Por último, si $p = 0.0333$, $\mathrm{P}\left(X = 2\right) = \binom{10}{2} p^2(1-p)^8 \approx 0.03812273$

Ítem b

Una gran ventaja de las variables aleatorias "con nombre", es que, una vez conocidos sus parámetros, su valor esperado, varianza y desvío siguen fórmulas conocidas. De este modo, se evitan los cálculos de la guía anterior que con recorridos grandes pueden resultar largos y tediosos.

En este caso, $X \sim \mathrm{Bi}(10; 0.03333)$, dando lugar a las siguientes fórmulas:

- $\mathrm{E}\left[X\right] = n \cdot p = 10 \cdot 0.0333 = 0.3333$
- $\sigma(X) = \sqrt{n \cdot p \cdot (1-p)} \approx 0.5676462$

Ejercicio 28

Sea p la probabilidad de que cualquier símbolo particular de un código se transmita erróneamente a través de un sistema de comunicaciones. Suponga que en diferentes símbolos ocurren errores independientemente uno de otro. Suponga también que con probabilidad p_2 un símbolo erróneo se corrige al recibirse. Sea X_n el número de símbolos correctos recibidos en un bloque de mensaje formado por n símbolos (una vez que el proceso de corrección haya terminado).

a) Demostrar que X_n es una variable aleatoria discreta con distribución binomial.

b) Determinar el valor esperado y la varianza de X_n si $n = 10$, $p = 0.02$ y $p_2 = 0.95$.

Resolución

Variable aleatoria

La variable aleatoria es la siguiente:

$$X_n = \text{Cantidad de dígitos correctos recibidos.}$$

Eventos

Consideramos los siguientes eventos:

- A_i = el i-ésimo símbolo se transmite correctamente en primera instancia.
- B_i = el i-ésimo símbolo se corrige.
- C_i = el dígito se recibe correctamente.

Datos

Tenemos los siguientes datos:

- $\mathrm{P}(A_i) = 1 - p \Rightarrow \mathrm{P}\left(\overline{A_i}\right) = p$
- $\mathrm{P}\left(B_i \middle| \overline{A_i}\right) = p_2$

Además, los eventos A_i son independientes entre sí, los eventos B_i son independientes entre sí y los eventos C_i son independientes entre sí.

Observación: Esto dice, por ejemplo, que A_i es independiente de A_j si $i \neq j$, pero A_i y B_i **no son independientes.** Es decir, la independencia sólo se da en eventos relativos a distintos dígitos.

Ítem a

Para demostrar que X_n es binomial, basta ver que la variable cuenta la cantidad de dígitos correctos en n transmisiones independientes. Es decir, $X_n \sim \mathrm{Bi}(n, \theta)$. Resta determinar cuál es la probabilidad de que un símbolo se transmita correctamente (dado por $\theta = \mathrm{P}(C_i)$) para terminar de caracterizar la distribución.

El dígito se transmite de forma correcta en dos casos. Si es inicialmente enviado correctamente (A_i) o si es transmitido erróneamente y luego corregido $(\overline{A_i} \cap B_i)$. Por lo tanto, $\theta = \mathrm{P}(C_i)$ será:

$$\theta = \mathrm{P}(C_i) = \mathrm{P}(A_i) + \mathrm{P}\left(B_i \middle| \overline{A_i}\right) \cdot \mathrm{P}\left(\overline{A_i}\right) = 1 - p + p_2 \cdot p$$

Como los errores ocurren de forma independiente y con la misma probabilidad, se obtiene que:

$$\mathrm{P}\left(X_n = k\right) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$$

y por lo tanto, $X_n \sim Bi(n, \theta)$.

Ítem b

Como nos dicen los valores para $p = 0.02$ y $p_2 = 0.95$, tenemos que:

$$\theta = 1 - p + p_2 \cdot p1 - 0.02 + 0.95 \cdot 0.02 = 0.999$$

Por lo tanto, como $X_n \sim \mathrm{Bi}(n, \theta)$:

- $E(X_n) = n \cdot \theta = 10 \cdot 0.999 = 9.99$
- $V(X_n) = n \cdot \theta \cdot (1-\theta) = 10 \cdot 0.999 \cdot 0.001 = 0.00999$

Ejercicio 29

Se supone que un proceso de fabricación de fusibles produce un 1 % de fusibles defectuosos. El proceso se controla periódicamente examinando 10 fusibles, y, si alguno falla, se detiene el proceso para revisarlo.

a) ¿Cuál es la probabilidad de detener el proceso con el sistema de control implementado?

b) ¿Cúantos fusibles deberán probarse (en vez de 10) si se desea que la probabilidad de detener el proceso para revisarlo sea 0.95 cuando esté produciendo 5 % de fusibles defectuosos (que es un porcentaje de defectos muy por encima del aceptable)?

c) Con el tamaño de muestra obtenido en el punto (b), ¿cuál es la probabilidad de detener el proceso cuando esté produciendo al 1 % de fusibles defectuosos?

Resolución

Variable aleatoria

Consideramos

$$X_{10} = \text{Cantidad de fusibles defectuosos en una muestra de 10 fusibles.}$$

Eventos

Definimos el siguiente evento:

$$D = \text{En una extracción cualquiera, el fusible es defectuoso}$$

Datos

Tenemos como dato que $\mathrm{P}(D) = 0.01$. Además, considerando que las extracciones son independientes por la gran cantidad de fusibles desde la cual se extrae, y que además siempre tienen la misma probabilidad de ser defectuosos, podemos suponer que $X_{10} \sim \text{Binomial}(10, 0.01)$.

Ítem a

El proceso se detiene si se encuentra al menos un fusible defectuoso en la muestra de 10, es decir $X_{10} \geq 1$. Para calcular esta proababilidad, se pueden sumar todos los valores del recorrido de X_{10} mayores o iguales que 1. Como $R_{X_{10}}$ consiste de todos los enteros entre 0 y 10, esto equivaldría a sumar la probabilidad de todos los números entre 1 y 10. Es decir:

$$\mathrm{P}(X_{10} \geq 1) = \sum_{k=1}^{10} \mathrm{P}(X_{10} = k)$$

Sin embargo, es más fácil calcular el complemento, que depende únicamente de el término $X_{10} = 0$:

$$\mathrm{P}(X_{10} \geq 1) = 1 - \mathrm{P}(X_{10} < 1) = 1 - \mathrm{P}(X_{10} = 0) = 1 - \binom{10}{0} 0.01^0 \cdot 0.99^{10} = 1 - 0.9043821 = 0.09561792$$

Como $X_{10} \sim \text{Binomial}(10, 0.01)$, se puede utilizar su fórmula para la probabilidad puntual:

$$\mathrm{P}(X_{10} \geq 1) = 1 - \mathrm{P}(X_{10} = 0) = 1 - \binom{10}{0} 0.01^0 \cdot 0.99^{10} = 1 - 0.9043821 = 0.09561792$$

Por lo tanto, la probabilidad de detener el proceso es de 0.09561792.

Ítem b

En este ejercicio, el porcentaje de defectuosos es del 5 % y el tamaño de muestra no es necesariamente 10. Sin embargo, podemos considerar un tamaño de muestra n y la siguiente variable:

$$X_n = \text{Cantidad de fusibles defectuosos en una muestra de n fusibles.}$$

Con el mismo criterio que en el ítem anterior, podemos asumir que $X_n \sim \text{Bi}(n; 0.05)$. En este contexto, se pide determinar n de forma que

$$\text{P}\left(X_n \geq 1\right) = 0.95.$$

Vale aclarar que n debe ser entero, por lo que puede ser que no haya un valor entero de n que cumpla la ecuación. En estos casos, en vez de una igualdad se plantea una desigualdad, pero ¿cómo elegirla?.

Como la proporción de defectuosos (5 %) es mayor a la deseada, lo intuitivo es pedir una mayor probabilidad de detener el proceso. Es decir, pediremos:

$$\text{P}\left(X_n \geq 1\right) \geq 0.95.$$

Como $\text{P}\left(X_n \geq 1\right) = 1 - \text{P}\left(X_n = 0\right) = 1 - 0.95^n$, entonces la inecuación deviene en pedir:

$$1 - 0.95^n \geq 0.95 \Leftrightarrow 0.05 \geq 0.95^n$$

Aplicando logaritmos a ambos lados y usando sus propiedades:

$$\ln(0.05) \geq \ln(0.95^n) \Leftrightarrow \ln(0.05) \geq n \cdot \ln(0.95)$$

Notar que $\ln(0.95) < 0$, por lo que al despejar, se invierte el signo de la desigualdad:

$$\frac{\ln(0.05)}{\ln(0.95)} \leq n \Leftrightarrow 58.40397 \leq n$$

Como n debe ser entero, consideramos que la mínima cantidad de extracciones debe ser $n = 59$ para detener el proceso con una probabilidad mayor o igual que 0.95.

Ítem c

Considerando $n = 59$ y $p = 0.01$, se tiene que:

$$\text{P}\left(X_n \geq 1\right) = 1 - \text{P}\left(X_n = 0\right) = 1 - 0.99^{59} = 0.4473165$$

Es importante saber interpretar los resultados en esta materia. Notar que en este caso, se pensó en incrementar el tamaño de muestra para detener con probabilidad alta el proceso cuando la proporción de defectuosos es inaceptable. Sin embargo, esta decisión tiene sus implicancias. Viendo el resultado, tratando de descartar proporciones grandes (5 %) de defectuosos, se detiene el proceso mucho más (44.73 % vs. 9.56 %) en el caso de tener 1 % de defectuosos. Esto puede ser indeseable ya que interrumpe constantemente la producción. Estos análisis de costo-beneficio son muy usuales en esta materia.

Ejercicio 30

Una partícula esta sometida a la acción de impulsos aleatorios, pudiéndose mover en la dirección del eje x. Supongamos conocida la probabilidad de un salto unitario hacia la derecha es p y la probabilidad de un salto unitario hacia la izquierda $q = 1 - p$. Este proceso aleatorio se denomina un *Random Walk* o caminata aleatoria unidimensional.

a) Obtener una expresión para la probabilidad que la partícula alcance cierta abscisa x al cabo de n saltos. Suponga que la posición inicial es 0.

   Sugerencia: Si la partícula se encuentra en la posición x tras n saltos de los cuales r fueron a la derecha y l a la izquierda entonces se tiene: $r + l = n$ y $r - l = x$.

b) Sea X_n la posición de la partícula al cabo de n saltos. Si $p = 0.5$ y $X_0 = 0$ calcular las siguientes probabilidades:

   a) $\mathrm{P}(X_n \geq 0)$ para n tomando los valores de 1 a 4.

   b) $\mathrm{P}(|X_n| \leq 2)$ para n tomando los valores de 1 a 4.

Resolución

Variable aleatoria

La variable de interés depende de la cantidad de pasos. La definiremos como:

$$X_n = \text{Posición de la partícula luego de n saltos}$$

Es decir, si en dos saltos, ambos fueron a la derecha, se tiene que $X_2 = 2$, mientras que si en tres saltos hubo 2 saltos a la derecha y uno a la izquierda, $X_3 = 1$.

Además podemos considerar el siguiente conjunto de variables:

$$Y_n = \begin{cases} 1 & \text{si el n-ésimo salto fue a la derecha} \\ -1 & \text{si el n-ésimo salto fue a la izquierda} \end{cases}$$

Esto permite individualizar los saltos. Por ejemplo, el caso $X_3 = 1$ se puede obtener a partir de varias combinaciones:

- $Y_1 = 1,\ Y_2 = 1,\ Y_3 = -1 \Rightarrow X_3 = 1$ (dos saltos consecutivos a la derecha y uno a la izquierda)
- $Y_1 = 1,\ Y_2 = -1,\ Y_3 = 1 \Rightarrow X_3 = 1$ (un salto a la derecha, uno a la izquierda y otro a la derecha)
- $Y_1 = -1,\ Y_2 = 1,\ Y_3 = 1 \Rightarrow X_3 = 1$ (uno salto a la izquierda y dos consecutivos a la derecha)

Datos

No tenemos muchos datos sobre X_n, pero sí sobre Y_n:

- $R_{Y_n} = \{-1; 1\}$
- $\mathrm{P}(Y_n = -1) = 1 - p$
- $\mathrm{P}(Y_n = 1) = p$

Además, como los saltos son independientes entre sí, las variables Y_n son independientes.

Ítem a

Por empezar, observemos el comportamiento de X_n para los primeros valores de $n = 1, 2, 3$. En los siguientes gráficos, los círculos verdes representan puntos de partida, en rojo los saltos a la izquierda, en azul los saltos a la derecha.

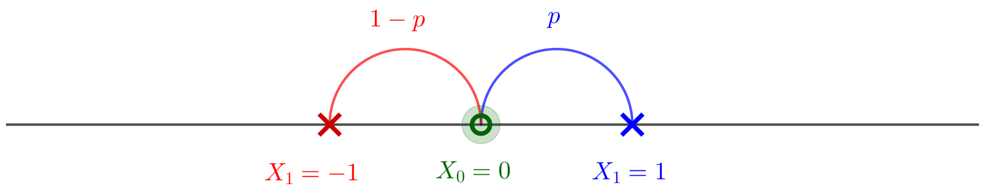


Figura 4: Análisis de los resultados de la variable X_1, es decir, el primer salto.

En este caso, partiendo de $X_0 = 0$, es simple ver que $R_{X_1} = \{-1; 1\}$, $\mathrm{P}(X_1 = 1) = p$ y $\mathrm{P}(X_1 = -1) = 1 - p$.

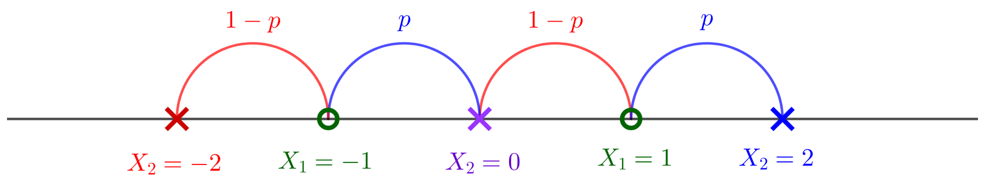


Figura 5: Análisis de la variable X_2 luego de dos saltos.

Aquí se agrega un color violeta (con nostalgias de la educación primaria, combinación de azul y rojo) en el caso de $X_2 = 0$ ya que se puede acceder a partir tanto de $X_1 = 1$ como de $X_1 = -1$. Podemos ver además que el recorrido es $R_{X_2} = \{-2; 0; 2\}$. Por lo que debemos calcular las probabilidades puntuales de estos valores, valiéndonos del teorema de probabilidad puntual:

$$\mathrm{P}(X_2 = -2) = \mathrm{P}\big(X_2 = -2 |\, X_1 = -1\big) \cdot \mathrm{P}(X_1 = -1) + \mathrm{P}\big(X_2 = -2 |\, X_1 = 1\big) \cdot \mathrm{P}(X_1 = 1)$$

Notemos que para pasar de $X_1 = -1$ a $X_2 = -2$, el segundo salto debe ser a la izquierda ($Y_2 = -1$). Además, como corresponden a saltos distintos, $X_1 = -1$ e $Y_2 = -1$ son independientes. Por último, como en un salto no es posible llegar de $X_1 = 1$ hasta $X_2 = -2$, podemos afirmar que $\mathrm{P}\big(X_2 = -2 |\, X_1 = 1\big) = 0$. Con toda esta información, llegamos a lo siguiente:

$$\mathrm{P}(X_2 = -2) = \underbrace{\mathrm{P}\big(X_2 = -2 |\, X_1 = -1\big)}_{\mathrm{P}\left(Y_2 = -1 | X_1 = -1\right) = \mathrm{P}(Y_2 = -1) = 1 - p} \cdot \underbrace{\mathrm{P}(X_1 = -1)}_{1-p} + \underbrace{\mathrm{P}\big(X_2 = -2 |\, X_1 = 1\big)}_{0} \cdot \underbrace{\mathrm{P}(X_1 = 1)}_{p} = (1-p)^2$$

Este resultado obedece a una intuición: para llegar a la posición -2 en dos saltos, deben haber sido dos saltos a la izquierda. Y como esos dos saltos son independientes, se multiplican sus probabilidades.

Del mismo modo, se calculan las otras probabilidades:

$$\mathrm{P}\left(X_2=0\right)=\underbrace{\mathrm{P}\big(X_2=0|\,X_1=-1\big)}_{\mathrm{P}\big(Y_2=1|X_1=-1\big)=\mathrm{P}(Y_2=1)=p}\cdot\underbrace{\mathrm{P}\left(X_1=-1\right)}_{1-p}+\underbrace{\mathrm{P}\big(X_2=0|\,X_1=1\big)}_{\mathrm{P}\big(Y_2=-1|X_1=1\big)=\mathrm{P}(Y_2=-1)=1-p}\cdot\underbrace{\mathrm{P}\left(X_1=1\right)}_{p}=2p(1-p)$$

$$\mathrm{P}\left(X_2=2\right)=\underbrace{\mathrm{P}\big(X_2=2|\,X_1=-1\big)}_{0}\cdot\underbrace{\mathrm{P}\left(X_1=-1\right)}_{1-p}+\underbrace{\mathrm{P}\big(X_2=2|\,X_1=1\big)}_{\mathrm{P}\big(Y_2=1|X_1=1\big)=\mathrm{P}(Y_2=1)=p}\cdot\underbrace{\mathrm{P}\left(X_1=1\right)}_{p}=p^2$$

Por útimo, visualizaremos lo que sucede con X_3:

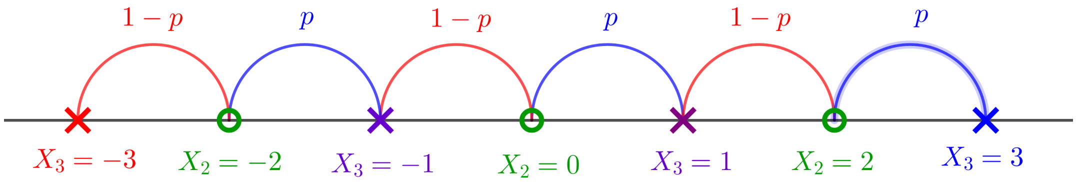


Figura 6: Análisis de la variable X_3 luego de tres saltos.

Viendo este gráfico, el recorrido es $R_{X_3}=\{-3;-1;1;3\}$. Calculamos sus probabilidades puntuales, a partir de la variable X_2, así los cálculos se reducen a analizar únicamente el último salto:

$$\mathrm{P}\left(X_3=-3\right)=\underbrace{\mathrm{P}\big(X_3=-3|\,X_2=-2\big)}_{\mathrm{P}(Y_3=-1)=1-p}\cdot\underbrace{\mathrm{P}\left(X_2=-2\right)}_{(1-p)^2}+\underbrace{\mathrm{P}\big(X_3=-3|\,X_2=0\big)}_{0}\cdot\mathrm{P}\left(X_2=0\right)+\underbrace{\mathrm{P}\big(X_3=-3|\,X_2=2\big)}_{0}\cdot\mathrm{P}\left(X_2=2\right)=(1-p)^3$$

$$\mathrm{P}\left(X_3=-1\right)=\underbrace{\mathrm{P}\big(X_3=-1|\,X_2=-2\big)}_{\mathrm{P}(Y_3=1)=p}\cdot\underbrace{\mathrm{P}\left(X_2=-2\right)}_{(1-p)^2}+\underbrace{\mathrm{P}\big(X_3=-1|\,X_2=0\big)}_{\mathrm{P}(Y_3=-1)=1-p}\cdot\underbrace{\mathrm{P}\left(X_2=0\right)}_{2p(1-p)}+\underbrace{\mathrm{P}\big(X_3=-1|\,X_2=2\big)}_{0}\cdot\mathrm{P}\left(X_2=2\right)=3p(1-p)^2$$

$$\mathrm{P}\left(X_3=1\right)=\underbrace{\mathrm{P}\big(X_3=1|\,X_2=-2\big)}_{0}\cdot\mathrm{P}\left(X_2=-2\right)+\underbrace{\mathrm{P}\big(X_3=1|\,X_2=0\big)}_{\mathrm{P}(Y_3=1)=p}\cdot\underbrace{\mathrm{P}\left(X_2=0\right)}_{2p(1-p)}+\underbrace{X_3=1|X_2=2\big)}_{\mathrm{P}(Y_3=-1)=1-p}\cdot\underbrace{\mathrm{P}\left(X_2=2\right)}_{p^2}=3p^2(1-p)$$

$$\mathrm{P}\left(X_3=3\right)=\underbrace{\mathrm{P}\big(X_3=3|\,X_2=-2\big)}_{0}\cdot\mathrm{P}\left(X_2=-2\right)+\underbrace{\mathrm{P}\big(X_3=3|\,X_2=0\big)}_{0}\cdot\mathrm{P}\left(X_2=0\right)+\underbrace{\mathrm{P}\big(X_3=3|\,X_2=2\big)}_{\mathrm{P}(Y_3=1)=p}\cdot\underbrace{\mathrm{P}\left(X_2=2\right)}_{p^2}=p^3$$

Conclusiones

Observemos que siempre el valor máximo de X_n es n y el mínimo es $-n$, pero que no siempre alcanza todos los valores intermedios. Por ejemplo, X_2 no puede tomar el valor 1 ni -1. De hecho, cuando n es par, X_n toma los valores pares entre $-n$ y n, y cuando n es impar, X_n toma los valores impares entre $-n$ y n. Es decir, formalizando un poco esta conclusión:

- Si $n=2k$ es par, $R_{X_n}=\{2m:-k\leq m\leq k\}$
- Si $n=2k-1$ es impar, $R_{X_n}=\{2m-1:-k+1\leq m\leq k\}$

Notar que se cumplen los valores mínimos y máximos. Si $n=2k$ es par, el mínimo valor de m es $-k$, entonces, el mínimo valor del recorrido es $-2k=-n$. Del mismo modo, $m\leq k$, entonces $2m\leq 2k=n$. Si $n=2k-1$ es impar, el mínimo valor de m es $-k+1$, entonces, el mínimo valor del recorrido es $2(-k+1)-1=-2k+1=-n$. Del mismo modo, $m\leq k$, entonces $2m-1\leq 2k-1=n$.

Concluimos que la cantidad de saltos y los valores del recorrido comparten la paridad. Por lo que será nula la probabilidad de que X_n tome un valor x impar si n es par y viceversa.

Nuevas variables aleatorias

Como vimos en la definición de las variables aleatorias, todas las formas de llegar a la posición 1 en tres saltos ($X_3 = 1$) se componen de dos saltos a la derecha y uno a la izquierda. Es decir, dada una cantidad de saltos n, alcanzar la posición x en esa cantidad de saltos, define una cantidad de saltos a la derecha, y por lo tanto, una cantidad de saltos a la izquierda.

Consideramos las variables

- R_n = cantidad de saltos hacia la derecha en n pasos
- L_n = cantidad de saltos hacia la izquierda en n pasos

Notemos que, al haber n saltos totales, $n = R_n + L_n$ y que la posición viene dada por $X_n = R_n - L_n$, ya que se suman la cantidad de saltos a la derecha y se restan los saltos a la derecha.

Concentrémosnos en R_n. Su recorrido es $R_{R_n} = \{0, 1, \cdots, n\}$ y cada salto puede ser a la derecha con la misma probabilidad p. Además, como estos saltos ocurren de forma independiente, podemos decir que $R_n \sim \text{Binomial}(n, p)$. Por lo tanto:

$$\mathrm{P}\left(R_n = r\right) = \binom{n}{r} p^r \cdot (1-p)^{n-r}$$

Notemos que el número combinatorio $\binom{n}{r}$ obedece a la cantidad de combinaciones posibles para ubicar r saltos a la derecha entre n saltos totales.

Sin embargo, la consigna pide otra cosa, $\mathrm{P}\left(X_n = x\right)$, sin especificar r. Entonces, para cada x, debemos determinar la cantidad de pasos realizados hacia la derecha (dados por r).

Como $n = r + l$ y $x = r - l$, sumando miembro a miembro, $n + x = 2r$ y por lo tanto $\frac{n+x}{2} = r$.

Vale aclarar que $\frac{n+x}{2}$ es entero porque si n es par, x también lo es (al menos, para que tenga sentido pensar su probabilidad). Y si n es impar, x también lo es. Así que $n + x$ siempre será par.

Por lo tanto:

$$\mathrm{P}\left(X_n = x\right) = \mathrm{P}\left(R_n = \frac{n+x}{2}\right) = \binom{n}{\frac{n+x}{2}} p^{\frac{n+x}{2}} \cdot (1-p)^{n-\frac{n+x}{2}} = \binom{n}{\frac{n+x}{2}} p^{\frac{n+x}{2}} \cdot (1-p)^{\frac{n-x}{2}}$$

Ítem b1

Notemos que si $p = \frac{1}{2}$, $p = 1 - p$. Además $n - \frac{n+x}{2} = \frac{n-x}{2}$ y por la propiedad de los números combinatorios $\left(\binom{n}{k} = \binom{n}{n-k}\right)$, se obtiene:

$$\mathrm{P}\left(X_n = x\right) = \binom{n}{\frac{n+x}{2}} \left(\frac{1}{2}\right)^{\frac{n+x}{2}} \cdot \left(\frac{1}{2}\right)^{\frac{n-x}{2}} = \binom{n}{\frac{n+x}{2}} \left(\frac{1}{2}\right)^{n} = \binom{n}{\frac{n-x}{2}} \left(\frac{1}{2}\right)^{n} = \mathrm{P}\left(X_n = -x\right)$$

Podemos concluir entonces que la distribución es simétrica respecto del cero. Por lo tanto, $\mathrm{P}\left(X_n \geq k\right) = \mathrm{P}\left(X_n \leq -k\right)$ para todo $k \geq 0$. Por lo tanto, podemos concluir lo siguiente:

$$\mathrm{P}\left(X_n = 0\right) + \mathrm{P}\left(X_n > 0\right) + \mathrm{P}\left(X_n < 0\right) = 1 \Leftrightarrow \mathrm{P}\left(X_n = 0\right) + 2 \cdot \mathrm{P}\left(X_n > 0\right) = 1 \Rightarrow \mathrm{P}\left(X_n > 0\right) = \frac{1 - \mathrm{P}\left(X_n = 0\right)}{2}$$

Entonces, se puede calcular solamente $\mathrm{P}\left(X_n = 0\right)$ y deducir el valor de $\mathrm{P}\left(X_n > 0\right)$.

- $n = 1 : \mathrm{P}\left(X_1 = 0\right) = 0 \Rightarrow \mathrm{P}\left(X_1 > 0\right) = \frac{1-0}{2} = \frac{1}{2}$ y $\mathrm{P}\left(X_1 \geq 0\right) = \frac{1}{2}$
- $n = 2 : \mathrm{P}\left(X_2 = 0\right) = 2 \cdot \frac{1}{4} \Rightarrow \mathrm{P}\left(X_2 > 0\right) = \frac{1 - \frac{1}{2}}{2} = \frac{1}{4}$ y $\mathrm{P}\left(X_2 \geq 0\right) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$
- $n = 3 : \mathrm{P}\left(X_3 = 0\right) = 0 \Rightarrow \mathrm{P}\left(X_3 > 0\right) = \frac{1-0}{2} = \frac{1}{2}$ y $\mathrm{P}\left(X_3 \geq 0\right) = \frac{1}{2}$
- $n = 4 : \mathrm{P}\left(X_4 = 0\right) = 6 \cdot \frac{1}{16} \Rightarrow \mathrm{P}\left(X_4 > 0\right) = \frac{1 - \frac{3}{8}}{2} = \frac{5}{16}$ y $\mathrm{P}\left(X_4 \geq 0\right) = \frac{3}{8} + \frac{5}{16} = \frac{11}{16}$

Ítem b2

Para $n = 1, 2$, $\mathrm{P}(|X_n| \leq 2) = 1$ ya que su recorrido se encuentra entre -2 y 2.

Para $n = 3, 4$, los valores del recorrido de X_n que no se encuentran entre -2 y 2 son el máximo y el mínimo, por lo tanto, conviene restar las probabilidades de esos valores a 1 para obtener la probabilidad deseada, ya que consiste de menor cantidad de teérminos.

- $n = 3 : \mathrm{P}(|X_3| \leq 2) = 1 - \mathrm{P}(|X_3| > 2) = 1 - \mathrm{P}(X_3 = 3) - \mathrm{P}(X_3 = -3) = 1 - 2\mathrm{P}(X_3 = 3) = 1 - 2 \cdot \left(\frac{1}{2}\right)^3 = \frac{3}{4}$
- $n = 4 : \mathrm{P}(|X_4| \leq 2) = 1 - \mathrm{P}(|X_4| > 2) = 1 - \mathrm{P}(X_4 = 4) - \mathrm{P}(X_4 = -4) = 1 - 2\mathrm{P}(X_4 = 4) = 1 - 2 \cdot \left(\frac{1}{2}\right)^4 = \frac{7}{8}$

Ejercicio 32

Una batería de artillería dispara a un blanco y se sabe que la probabilidad de acertar es $p = 0.01$. ¿Cuántos disparos tendrá que realizar para tener probabilidad mayor o igual que 90 % de dar en el blanco por lo menos una vez? Resolver usando las distribuciones binomial y de Poisson. Generar una tabla con la razón de las respuestas obtenidas con una y otra distribución para valores de p desde 0.01 hasta 0.1 en pasos de 0.005.

Resolución

Llamemos X_n a la cantidad de veces que la batería da en el blanco en n tiros. Si cada tiro es independiente de los demás, $X_n \sim \text{Binomial}(n, p)$. Luego,

$$P(X_n \geq 1) = 1 - P(X_n < 1) = 1 - P(X_n = 0). \tag{55}$$

Como se requiere que esta probabilidad sea mayor que 0.9, necesitamos encontrar n tal que

$$1 - P(X_n = 0) \geq 0.9 \Leftrightarrow P(X_n = 0) \leq 0.1. \tag{56}$$

Usando la distribución binomial,

$$P(X_n = 0) = \binom{n}{0} p^0 (1-p)^{n-0} = (1-p)^n \leq 0.1 \Leftrightarrow n \log_{10}(1-p) \leq \log_{10}(0.1) = -1 \Leftrightarrow \tag{57}$$

$$\Leftrightarrow n \geq -\frac{1}{\log_{10}(1-p)}, \tag{58}$$

donde hemos usado el hecho que la función logaritmo es creciente y que $\log(x) < 0$ para $x \in (0, 1)$.

En ciertas ocasiones, se puede usar la distribución Poisson para aproximar la binomial, usando como parámetro $\lambda = np$. Haciendo esto, tenemos

$$P(X_n = 0) \approx \frac{(np)^0}{0!} e^{-np} = e^{-np}, \tag{59}$$

$$e^{-np} \leq 0.1 \Leftrightarrow -np \leq \ln(0.1) \Leftrightarrow n \geq -\frac{\ln(0.1)}{p}. \tag{60}$$

La Tabla 7 muestra los resultados de calcular el mínimo número de tiros necesario usando la distribución binomial y la aproximación Poisson. Como puede observarse, las diferencias son pequeñas, de a lo sumo 2.

Tabla 7: Mínimo número de tiros de acuerdo a la distribución binomial y la aproximación Poisson

p	Binomial $n^{(b)}_{\text{mín}} = \left\lceil -\frac{1}{\log_{10}(1-p)} \right\rceil$	Aproximación Poisson $n^{(P)}_{\text{mín}} = \left\lceil -\frac{\ln(0.1)}{p} \right\rceil$	Diferencia $\left\| n^{(b)}_{\text{mín}} - n^{(P)}_{\text{mín}} \right\|$
0.010	230	231	1
0.015	153	154	1
0.020	114	116	2
0.025	91	93	2
0.030	76	77	1
0.035	65	66	1
0.040	57	58	1
0.045	51	52	1
0.050	45	47	2
0.055	41	42	1
0.060	38	39	1
0.065	35	36	1
0.070	32	33	1
0.075	30	31	1
0.080	28	29	1
0.085	26	28	2
0.090	25	26	1
0.095	24	25	1
0.100	22	24	2