Departamento de Ingeniería Civil Industrial
Probabilidades - IN3141
Profesor: Charles Thraves - Natalia Trigo
Aux's: Camila Carrasco - Monserrat Marchant - Leonardo Meneses - Sebastián Rojas

Auxiliar #5

Variable aleatoria discreta II

Resumen

Distribuciones conocidas:

- Bernoulli $X \sim Ber(p)$, $R_X = \{0, 1\}$ tal que

$$p_X(x) = \begin{cases} p \text{ si } x = 1 \\ 1 - p \text{ si } x = 0 \end{cases}$$

Interpretación: X modela el resultado de un experimento dicotómico, de probabilidad p de resultar exitoso.

$$\mathbb{E}[X] = p \qquad \text{Var}(X) = p(1-p)$$

- Binomial $X \sim Bin(n, p)$, $R_X = \{0. \cdots, n\}$ tal que

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Observación: Si $\{X_i\}_{i=1}^n$ es una familia de v.as. independiente tal que $X_i \sim Ber(p)$ $\forall i \in \{1, \cdots, n\}$, entonces $\sum_{i=1}^n X_i \sim Bin(n, p)$
Interpretación: X modela el número de éxitos en n realizaciones de un experimento dicotómico de probabilidad favorable p.

$$\mathbb{E}[X] = np \qquad \text{Var}(X) = np(1-p)$$

- Hipergeométrica $X \sim \text{Hipergeom}(N, K, n)$, $R_X = \{\max(0, n + K - N), \cdots, \min(K, n)\}$ tal que

$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Interpretación: Dado un conjunto de tamaño N, con un subconjunto especial de tamaño K, se nos permite sacar n elementos sin repetir. X es la cantidad de elementos especiales extraídos.

$$\mathbb{E}[X] = n\frac{K}{N} \qquad \text{Var}(X) = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{N-1}$$

- Geométrica $X \sim \text{Geometrica}(p)$, $R_X = \{1, 2, 3, \cdots\}$ tal que

$$p_X(k) = (1-p)^{k-1} \cdot p \ \forall k \in R_X$$

Interpretación: Dado un experimento dicotómico de probabilidad favorable p, X es la primera vez que tal experimento resulta ser exitoso.

$$\mathbb{E}[X] = \frac{1}{p} \qquad \text{Var}(X) = \frac{1-p}{p^2}$$

- Binomial Negativa $X \sim \text{Binomial Negativa}(m, p)$, $R_X = \{m, m+1, m+2, \cdots\}$ es tal que

$$p_X(k) = \binom{k-1}{m-1} p^m (1-p)^{k-m}$$

Interpretación: Dado un experimento dicotómico de probabilidad favorable p, X es la cantidad de intentos para haber obtenido m éxitos.

$$\mathbb{E}[X] = \frac{m}{p} \qquad \text{Var}(X) = \frac{m(1-p)}{p^2}$$

- Poisson $X \sim \text{Poisson}(\lambda)$, $R_X = \{0, 1, \cdots\}$ tal que

$$p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!} \ \forall k \in R_X$$

Interpretación: Representa el número de veces que ocurre un evento, dado que ocurre λ veces por unidad de tiempo.

$$\mathbb{E}[X] = \lambda \qquad \text{Var}(X) = \lambda$$

- Teorema límite de Poisson Para p_n secuencia de números en $(0, 1)$ tal que $np_n \to \lambda$ cuando $n \to \infty$, se tiene que

$$\lim_{n \to \infty} \binom{n}{k} p_n^k (1 - p_n)^{n-k} = e^{-\lambda}\frac{\lambda^k}{k!}$$

En particular, para un experimento binomial $X \sim Bin(n, p)$ tal que $p = \frac{\lambda}{n}$, para n suficientemente grande se asemeja a una v.a. $Y \sim Poisson(\lambda)$ *Interpretación:* Sirve para modelar eventos de probabilidad muy baja de ocurrir, pero que tienen muchas instancias posibles donde podrían ocurrir.

Pregunta 1

Dos amigos, Antonia y Benito, tienen una importante reunión en el otro extremo de la ciudad. La ciudad es convenientemente modelada por una grilla de 7×7. Específicamente, Antonia y Benito se encuentran en el punto con coordenadas $(0,0)$, mientras que la reunión se realizará en el punto con coordenadas $(7,7)$. Para trasladarse al lugar de la reunión, en cada esquina (par de coordenadas cuya abscisa y ordenada son valores enteros) lanzan una moneda, la cual tiene probabilidad p de salir cara, y $1-p$ de salir sello. Cada vez que la moneda sale cara, avanzan por un arco hacia la derecha, es decir, avanzan hacia la derecha en una unidad. Si por el contrario, la moneda sale sello, entonces avanzan por el arco hacia arriba. Asuma que en caso de llegar a los puntos con abscisa u ordenada (pero no ambas) igual a 7, entonces los siguientes tramos son recorridos en línea recta en la dirección que apunta hacia el punto de coordenadas $(7,7)$.

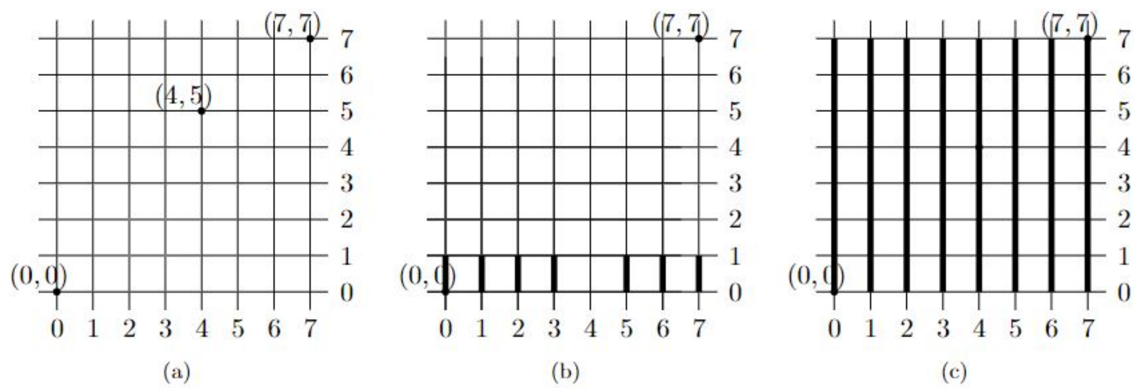

Figure 1: Mapas de la ciudad para partes 1, 2 y 3, respectivamente.

1. Calcule la probabilidad de pasar por el punto con coordenadas $(x,y)=(4,5)$ en el trayecto hacia la reunión.
   Hint: note que solo importa mirar los primeros $x+y$ movimientos en la grilla. (Fig. 1 (a))

   Solución

   Supongamos que moverse a la derecha es un éxito. Así, en un camino de 9 pasos, necesitaremos 4 éxitos y 5 fracasos. La distribución que representa contar los éxitos en un número de intentos fijos es la binomial: más precisamente, si X es el número de pasos a la derecha en 9 movimientos, entonces $X \sim \text{Bin}(9,p)$ y

   $$\mathbb{P}(\text{pasar por } (4,5)) = \mathbb{P}(X=4) = \binom{9}{4} \cdot p^4(1-p)^{9-4}$$

2. Suponga que en las calles marcadas con linea gruesa (Fig. 1 (b)) hay alto tráfico. Calcule la probabilidad de que Antonia y Benito **no** hayan pasado por zonas de alto tráfico en su trayecto hacia la reunión.

   Solución

   Ahora, necesitamos que el trayecto pase por los puntos $(4,0)$ y $(4,1)$.

   - Para el tramo $(0,0) \to (4,0)$ necesitamos 4 pasos a la derecha seguidos: con probabilidad p^4 de ocurrir (suponiendo independencia de lanzamientos de las monedas).
   - Para el tramo $(4,0) \to (4,1)$ necesitamos 1 paso hacia arriba, con probabilidad $1-p$ de ocurrir.

- Para los tramos siguientes hasta la oficina, sabemos que (eventualmente) se llega siempre a la oficina, por lo que tienen probabilidad 1 de ocurrir.

Luego, la probabilidad de no haber pasado por zonas de alto tráfico es $p^4 \cdot (1-p)$.

3. Ahora suponga que existe alto tráfico en todos los arcos en sentido vertical (Fig. 1 (c)). Se sabe que Antonia manejó 6 arcos y Benito el resto, pero no se sabe cuáles manejaron cada uno. Se sabe además que Benito se enoja cuando maneja 3 o más arcos con tráfico. ¿Cuál es la probabilidad de que Benito haya llegado enojado a la reunión?
Hint: note que para responder la pregunta, no importa el trayecto específico que hicieron para llegar desde el punto de origen al punto destino.

Solución

Primero notamos que para llegar al punto $(7,7)$ hay que recorrer 14 arcos en total, de los cuales 7 son verticales y 7 son horizontales. Como se sabe que Antonia maneja 6 arcos, entonces Benito necesariamente maneja los 8 arcos restantes.

Se define X como la variable aleatoria que cuenta la cantidad de arcos verticales que maneja Benito. Debido a que se cuenta la cantidad de arcos recorridos dentro del subgrupo de arcos verticales, que son en total 7, X sigue una distribución hipergeométrica con parámetros $N = 14$, $K = 7$ y $n = 8$, es decir, $X \sim HG(14, 7, 8)$.

Así, la probabilidad de que Benito se enoje corresponde a manejar 3 o más arcos verticales:

$$\mathbb{P}(B \text{ se enoje}) = \mathbb{P}(X \geq 3) = \sum_{k=3}^{7} \mathbb{P}(X = k) = \sum_{k=3}^{7} \frac{\binom{7}{k}\binom{14-7}{8-k}}{\binom{14}{8}}$$

En donde se utiliza la fórmula de probabilidad acumulada para la distribución hipergeométrica.

Pregunta 2

Gato Bros es un popular juego de plataformas en que el personaje principal es un felino, que debe superar n niveles para convertirse en un héroe. Este juego no tiene un límite de vidas, pero es bastante difícil, por lo que podría tomarle muchos intentos pasar cada nivel. Cada vez que usted pierde en un nivel o que usted supera un nivel, se considera como un intento. Todos estos intentos son independientes entre si. En particular, para cada nivel i usted tiene una probabilidad p_i de superarlo cada vez que lo intenta.

a) Sea X_i : el número de intentos para superar la etapa i. ¿Cuál es la distribución de X_i? Justifique su respuesta

Solución

X_i es una variable aleatoria con distribución Geométrica de parámetro p_i, pues podemos considerar superar el nivel como el éxito, y cada intento en que no lo supero como un fracaso.

$$P(X_i = x) = p_i(1-p_i)^{x-1} \qquad \text{si } x \in \mathbb{N}/0$$

$$P(X_i = x) = 0 \qquad \text{si } x \notin \mathbb{N}/0$$

b) Sea X el número de intentos totales para superar los n niveles, calcule $E(X)$.

Solución

Notemos que podemos escribir X como $\sum_{i=1}^{n} X_i$. Luego usando al linealidad de la esperanza.

$$\begin{aligned}
E(X) &= E(\sum_{i=1}^{n} X_i) \\
&= \sum_{i=1}^{n} E(X_i) \qquad -> \quad \text{Linealidad de la esperanza} \\
&= \sum_{i=1}^{n} (\frac{1}{p_i}) \quad -> \qquad \text{Esperanza de una geométrica de parámetro } p_i
\end{aligned}$$

Para las próximas preguntas asumiremos que $p_i = p \quad \forall i \in \{1, 2, 3..., n\}$.

c) Sea X el número de intentos para superar el juego completo. ¿Cuál es la distribución de X? Justifique su respuesta.

Solución

Similar a la pregunta anterior, nos damos cuenta que si tenemos x intentos, esto significa que son n intentos exitosos asociados a cada uno de los niveles, y $x - n$ intentos fallidos repartidos dentro de los n niveles. Acá hay que darse cuenta que la suma de los fracasos en cada nivel debe ser igual a $x - n$.

Así, notamos que X distribuye $BinomialNegativa(n, p)$, porque se tiene una cantidad de n éxitos fija. Expresamos su PMF:

$$\begin{aligned}
P(X = x) &= \binom{x-1}{n-1} p^n (1-p)^{x-n} \\
&= \binom{x-1}{n-1} p^n (1-p)^{x-n} \qquad x \in \{n, n+1, n+2, n+3...\}
\end{aligned}$$

Considere que ahora recibe una recompensa de R monedas cuando supera el nivel i en menos de h intentos. Además, debe pagar un costo de c monedas en cada nivel i que juega (independiente del número de intentos que le tome superarlo).

d) Sea Y el saldo final en monedas que usted tiene al superar los n niveles. Calcule $\mathbb{E}[Y]$ y $Var(Y)$. Note que Y puede tomar valores negativos.
(Hint: Notar que el número de monedas a obtener puede escribirse como $c_1 \cdot Z - c_2$, donde c_1, c_2 son constantes, y Z una VA discreta con distribución conocida.)

Solución

Primero, notemos que en cada nivel i la probabilidad de recibir la recompensa es igual $P(X_i < h)$, recordando que X_i distribuye geométrica de parámetro p (considerando que ahora $p_i = p$), llegamos a que

$$P(X_i < h) = \sum_{k=1}^{h-1} p(1-p)^{k-1} = \sum_{k=0}^{h-2} p(1-p)^k$$

En este caso, consideremos $q = P(X_i < h)$ que sería la probabilidad de éxito asociada a conseguir la recompensa en el nivel i.
Como cada nivel es independiente y tiene la misma probabilidad p, podemos considerar la variable Z: cantidad de veces que consigo la recompensa. Notamos que esta variable Z distribuye $Binom(n, q)$. Además, como siempre pagaremos el mismo costo c en cada nivel, sabemos que el costo total será nc.

Finalmente, podemos escribir la esperanza como:

$$\begin{aligned} E(Y) &= E(R \cdot Z - nc) \\ &= R \cdot E(Z) - nc \quad -> \quad \text{Linealidad de la esperanza} \\ &= Rnq - nc \quad -> \quad \text{Esperanza de una distribución binomial} \\ &= n(Rq - c) \end{aligned}$$

Ahora para la varianza usamos la misma idea:

$$\begin{aligned} V(Y) &= V(R \cdot Z - nc) \\ &= V(R \cdot Z) \quad -> \quad \text{Constante que se suma a la VA no se considera en la varianza} \\ &= R^2 \cdot V(Z) \quad -> \quad \text{Constante que se multiplica a la VA sale al cuadrado} \\ &= R^2 nq(1-q) \quad -> \quad \text{Varianza de una distribución binomial} \end{aligned}$$

Pregunta 3

El número de correos electrónicos que recibe Sara en un día de trabajo (de lunes a viernes) se puede modelar con una distribución de Poisson con un promedio de $\frac{1}{6}$ correos electrónicos por minuto. Mientras que el número de correos electrónicos que recibe los fines de semana (sábado y domingo) sigue una distribución de Poisson con un promedio de $\frac{1}{30}$ correos electrónicos por minuto.

a) ¿Cuál es la probabilidad de que no reciba correos electrónicos en un intervalo de 4 horas en un domingo?

Solución

Se tiene que Sara recibe correos el día domingo con una distribución *Poisson* de tasa promedio $\frac{1}{30}$ correos por minuto. Sea X la cantidad de correos que recibe Sara en un intervalo de 4 horas el domingo, debemos convertir el tasa de correos electrónicos por minuto a correos electrónicos por 4 horas. $X \sim Poisson(\lambda)$, con $\lambda = 4 \cdot 60 \cdot \frac{1}{30} = 8$ correos en 4 horas.

Luego, la probabilidad a calcular resulta:

$$\mathbb{P}(X = 0) = \frac{e^{-8} \cdot 8^0}{0!} = e^{-8}$$

b) Se selecciona al azar un día de la semana (con igual probabilidad para todos los días), y se escoge de manera aleatoria un intervalo de una hora dentro de ese día. Si Sara no recibió ningún correo electrónico en ese intervalo. ¿Cuál es la probabilidad de que el día elegido sea un día de trabajo?

Solución

Partimos por definir los siguientes eventos:

E: Sara no recibe ningún correo en el intervalo de una hora seleccionado.

A_t: Se escoge un día de trabajo.

A_f: Se escoge un día de fin de semana.

Y las siguientes variables aleatorias:
X_t = correos que Sara recibe en un intervalo de una hora un día de trabajo.
Dónde X_t distribuye *Poisson* con tasa de 10 correos por hora (resultado de convertir la tasa de $\frac{1}{6}$ correos por minuto).
X_f=correos que Sara recibe en un intervalo de una hora en un día de fin de semana.
Dónde X_f distribuye *Poisson* con tasa de 2 correos por hora.

Luego, la probabilidad que se pide por enunciado es $\mathbb{P}(A_t|E)$, y utilizamos Teorema de Bayes para calcularla:

$$\mathbb{P}(A_t|E) = \frac{\mathbb{P}(E|A_t) \cdot \mathbb{P}(A_t)}{\mathbb{P}(E)}$$

En la expresión, $\mathbb{P}(E|A_t)$ representa a la probabilidad de que se reciban 0 correos un día de trabajo, así:

$$\mathbb{P}(E|A_t) = \mathbb{P}(X_t = 0) = \frac{e^{-10} \cdot 10^0}{0!}$$

Por otro lado, aplicando probabilidades totales para el término $\mathbb{P}(E)$:

$$\begin{aligned}\mathbb{P}(E) &= \mathbb{P}(E|A_t) \cdot \mathbb{P}(A_t) + \mathbb{P}(E|A_f) \cdot \mathbb{P}(A_f) \\ &= \frac{e^{-10} \cdot 10^0}{0!} \cdot \frac{5}{7} + \frac{e^{-2} \cdot 2^0}{0!} \cdot \frac{2}{7} \\ &= \frac{5}{7} \cdot e^{-10} + \frac{2}{7} \cdot e^{-2}\end{aligned}$$

Luego, la expresión resultante es:

$$\mathbb{P}(A_t|E) = \frac{e^{-10} \cdot \frac{5}{7}}{\frac{5}{7} \cdot e^{-10} + \frac{2}{7} \cdot e^{-2}} = \frac{5e^{-10}}{5 \cdot e^{-10} + 2 \cdot e^{-2}}$$