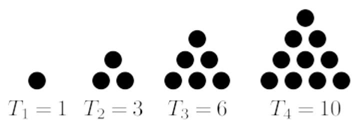
PROBABILIDADES Y ESTADÍSTICA (C)

PRÁCTICA 2

1 Preliminares

1. En las antigua Grecia, un método que se adoptaba para estudiar números naturales era representarlos mediante ciertos esquemas geométricos. Así, se acuñó la idea de "número triangular" que es cualquier número que pueda representarse mediante un triángulo con tantos puntos como unidades tenga el número, como indica la figura:



   Si llamamos T_n al n-ésimo número triangular:

   (a) Demostrar que $T_n + T_{n-1} = n^2$. *Sugerencia:* hagan dibujos.

   (b) Demostrar que $2T_n = n(n+1)$

   (c) Deducir que la suma de los primeros n números naturales es $\frac{n(n+1)}{2}$ (es decir, que $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$)

   (d) (*elija su propia aventura*) Así como se definen los números triangulares, pueden definirse los números cuadrangulares, pentagonales... ¿qué regularidades pueden encontrar?

2. (a) En una hoja, dibujen un cuadrado. Ahora, divídanlo en cuatro cuadrados iguales, y pinten uno de ellos. Seleccionen el cuadrado que está en diagonal a ese, divídanlo en cuatro partes, y repitan el proceso. Argumenten, sin mucho rigor, por qué al continuar indefinidamente este proceso, quedará pintada la tercera parte del cuadrado original.

   (b) Supongamos que tenemos r, K dos números positivos. Definimos, para cada $n \in \mathbb{N}$ (incluyendo $n = 0$), el número $a_n = Kr^n$.

   (c) Mostrar que, si r es mayor que 1, entonces se verifica que si $n \leq m$, entonces $a_n \leq a_m$. ¿Qué ocurre si r es menor que 1?

   (d) Demostrar que $S_n - rS_n = K - Kr^n$, donde S_n es la suma de los primeros n valores de a_n (es decir, $S_n = \sum_{i=0}^{n} a_i$)

   (e) Concluir que $\sum_{i=0}^{n} Kr^i = K\frac{1-r^n}{1-r}$

   (f) ¿Qué ocurre si tomamos $n \to +\infty$ en el ítem anterior?. Interpretar el enunciado del ítem a) en estos términos.

   (g) Calcular $\sum_{i=1}^{\infty} Kr^i$.

3. Sabiendo que, para cada $x \in \mathbb{R}$, $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$:

   (a) Demostrar que $\sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} = 1$

   (b) Calcular $\sum_{k=0}^{\infty} \frac{k.e^{-\lambda}\lambda^k}{k!}$

2 Variables aleatorias discretas

1. De un lote que contiene 15 artículos, de los cuales 4 son defectuosos, se eligen 3 artículos al azar con reposición. Si llamamos X al número de artículos defectuosos entre los seleccionados,

   a) Hallar la función de probabilidad puntual asociada a X y graficarla usando R.

   b) ¿Cuál es la probabilidad de que al menos 2 artículos sean defectuosos?

   c) Hallar la función de distribución acumulada de X y graficarla usando R.

   d) Estimar mediante una simulación las probabilidades calculadas en el item a).

2. Sea X una variable aleatoria con función de distribución

$$F_X(x) = \begin{cases} 0 & \text{si } x < 1 \\ 0.3 & \text{si } 1 \leq x < 3 \\ 0.4 & \text{si } 3 \leq x < 6 \\ 0.6 & \text{si } 6 \leq x < 12 \\ 1 & \text{si } 12 \leq x \end{cases}$$

   a) Hallar la función de probabilidad puntual de X.

   b) Calcular de dos maneras: utilizando la función de distribución y utilizando la función de probabilidad puntual, las siguientes probabilidades:

$$P(3 < X \leq 6) \qquad P(3 \leq X \leq 6) \qquad P(X \geq 4) \qquad P(X \geq 6)$$

   c) Utilizando el comando `sample`, generar 5 realizaciones de esta variable aleatoria en R.

   d) Mediante una simulación, estimar las probabilidades del item b).

3. Definir una función en R que, dados dos vectores $x = (x_1, \ldots, x_n)$ y $p = (p_1, \ldots, p_n)$ y un número real t, calcule $F_X(t)$, es decir, la probabilidad de que de la variable aleatoria discreta que toma valores $x_1, \ldots, x_n$ con probabilidades $p_1, \ldots, p_n$ respectivamente, sea menor o igual que t. Probarla con la variable aleatoria del ejercicio 2.

4. Si X es una v.a. discreta que toma sólo valores enteros, probar que para todo $k \in Z$:

$$p_X(k) = F_X(k) - F_X(k-1)$$

5. Calcular la esperanza de la variable aleatoria definida en el ejercicio 1 utilizando la definición y estimarla usando una simulación. Comparar los resultados.

6. Definir una función en R que, dados dos vectores $x = (x_1, \ldots, x_n)$ y $p = (p_1, \ldots, p_n)$ calcule la esperanza de la variable aleatoria discreta que toma valores $x_1, \ldots, x_n$ con probabilidades $p_1, \ldots, p_n$ respectivamente. Probar que funciona para la variable aleatoria X del ejercicio 1.

7. Definir una función en R que, dados dos vectores $x = (x_1, \ldots, x_n)$ y $p = (p_1, \ldots, p_n)$ calcule la varianza de la variable aleatoria discreta que toma valores $x_1, \ldots, x_n$ con probabilidades $p_1, \ldots, p_n$ respectivamente. Probar que funciona para la variable aleatoria X del ejercicio 1.

8. Sea X una v.a. con distribución Bernoulli de parámetro p.

    a) Calcular $E(X^k)$ para $k \in \mathbb{N}$
    b) Mostrar que $V(X) = p(1-p)$.

9. El 70% de las consultas de un sistema interactivo de computación requiere de acceso a bases de datos. Un sistema recibe 25 consultas independientes unas de otras,

    a) ¿cuál es la probabilidad de que:
        i. exactamente 20 consultas requieran acceso a una base de datos?
        ii. el número de consultas que requieran acceso a una base de datos esté entre 20 y 24 inclusive?
    b) Calcular el valor esperado y la varianza del número de consultas que requieren acceso a una base de datos.

10. Se tienen dos dados, uno equilibrado y el otro cargado en el cual los números 1 y 2 tienen probabilidad 1/3 y el resto 1/12. Se elige un dado al azar y se lo arroja tres veces (independientemente). Sea X el número de veces que sale 1 ó 2.

    a) ¿Cuál es la distribución de X condicional a que se eligió el dado cargado?
    b) Hallar una expresión general para $p_X(k)$.

11. Para verificar si se cumplen las normas establecidas para arrojar residuos al río Reconquista, un inspector visita al azar 10 de las 50 industrias establecidas a orillas de dicho río.

    a) Si en realidad 35 industrias no cumplen con alguna de las normas, ¿cuál es la distribución del número de industrias visitadas que están en infracción? Calcular la probabilidad de que 6 de las industrias visitadas estén en infracción.
    b) Si hay 500 industrias de las cuales 350 están en infracción, aproximar la distribución de (a) por una más simple. Calcular nuevamente la probabilidad de que 6 de las industrias visitadas estén en infracción.
    c) Sea X el número de fábricas que están en infracción entre las 10 visitadas. Calcular $E(X)$ y $V(X)$ para las distribuciones exacta (a) y aproximada (b).

12. Una rueda de ruleta está dividida en 38 secciones, de las cuales 18 son rojas, 18 son negras y las 2 restantes son verdes. Sea X el número necesario de juegos hasta obtener una sección verde en jugadas independientes.

    a) ¿Cuál es la probabilidad de que sean necesarias al menos 4 jugadas?
    b) Hallar la función de distribución acumulada de la v.a. X.
    c) Si fueron necesarias 7 o más jugadas, ¿cuál es la probabilidad de que se necesiten al menos 10 jugadas? Comparar con (a).
    d) ¿Cuál es la probabilidad de que sea necesario un número impar de jugadas?
    e) Hallar $E(X)$ y $V(X)$.

13. Si en el ejercicio anterior se define Y: número de juegos hasta obtener exactamente tres secciones verdes,

    a) ¿qué distribución tiene la v.a. Y?
    b) ¿cuál es la probabilidad de que se requieran exactamente 5 jugadas?
    c) hallar $E(Y)$ y $V(Y)$.

14. Con el fin de encontrar una palabra clave, un motor de búsqueda de internet explora una secuencia de sitios de la WEB en orden aleatorio. Al iniciar la búsqueda, el motor elige, al azar y con igual probabilidad, una entre dos secuencias posibles de sitios. Se sabe que el 10% de los sitios de la primera secuencia contienen esta palabra clave, mientras que sólo el 5% de los sitios de la segunda contienen dicha palabra.

    a) Si la búsqueda termina ni bien se encuentra un sitio que contenga la palabra clave, ¿cuál es la probabilidad de que más de 5 sitios deban ser explorados?
    b) Si se sabe que el motor de búsqueda encontró la palabra clave en la sexta visita ¿cuál es la probabilidad de que la haya encontrado en la segunda secuencia?
    c) Si la búsqueda termina cuando se encuentran 2 sitios que contenga la palabra clave ¿cuál es la probabilidad de que deban explorarse exactamente 10 sitios?

15. Un minorista ha verificado que la demanda semanal de cajones de cierto producto es una v.a. con distribución de Poisson de parámetro $\lambda = 2$. Completa su existencia los lunes por la mañana de manera de tener 4 cajones al principio de la semana. Al efectuar un análisis de la actividad de su negocio, se le plantean las siguientes preguntas:

    a) ¿Cuál es la probabilidad de vender todo su stock durante la semana?
    b) ¿Cuál es la probabilidad de que sea incapaz de cumplir con un pedido por lo menos?
    c) ¿Cuál es la distribución del número de cajones vendidos en una semana?
    d) ¿Con cuántos cajones deberí a iniciar la semana a fin de que la probabilidad de cumplir con todos sus pedidos fuese mayor o igual que 0.99?

16. Un bibliotecario ubica 1000 libros en un cierto día. Si la probabilidad de que un libro cualquiera sea mal ubicado es 0.001 y los libros se ubican en forma independiente, ¿cuál es la distribución aproximada del número de libros mal ubicados en ese día?

    Utilizando esta distribución, calcular la probabilidad de que

    a) por lo menos un libro sea mal ubicado ese día.
    b) exactamente 3 libros sean mal ubicados ese día. Comparar con el valor exacto.

17. En un concurso de pesca cada pescador paga 100$ por participar. La cantidad de peces obtenida por cada pescador durante el desarrollo del concurso es una v.a. con distribución de Poisson de parámetro $\lambda = 4$. Hay un premio de 50$ por pieza. Cada pescador tiene permitido cobrar a lo sumo 6 piezas (es decir, aunque pesque más de 6 cobrará sólo por 6).

a) Calcular la función de probabilidad puntual de la ganancia neta de un pescador.

b) ¿Cuánto dinero espera ganar cada participante?

18. Sea X una variable aleatoria binomial con $n = 5$ y $p = 1/3$.

    (a) Graficar la función de probabilidad puntual p_X y la función de distribución acumulada F_X, basándose en los comandos `dbinom` y `pbinom` respectivamente.

    (b) Graficar la función de distribución acumulada de X utilizando la función implementada en el ejercicio 3. Comparar con el gráfico obtenido en (a).

    (c) Utilizando el comando `rbinom` generar 1000 realizaciones de la variable aleatoria X y calcular la proporción de veces que X toma los valores 0, 1, 2, 3, 4 y 5. Comparar con los valores que toma la fpp.

19. Las tareas llegan a una cola de un sistema de computación con un solo servidor de acuerdo con un proceso de Poisson de parámetro $\lambda = 5$ tareas por minuto. Llamemos a dicho proceso X_t a) ¿Cuál es la probabilidad de que en 30 segundos lleguen menos de 5 tareas? b) Sea X_1 la cantidad de tareas recibidas en un minuto. Calcular:

$$\begin{array}{ll} P\left(X_1 \leq 6\right) & P\left(3 \leq X_1 \leq 6\right) \\ P\left(X_1 = 6\right) & P\left(3 < X_1 < 6\right) \\ P\left(X_1 \geq 5\right) & P\left(3 \leq X_1 \leq 6 \mid X_1 \geq 4\right) \end{array}$$

    Calcularlas exactamente utilizando los comandos dpois o ppois y luego estimarlas mediante una simulación (utilizar el comando rpois). c) ¿Cuál es el número esperado de tareas que se reciben en media hora? d) Estimar la esperanza calculada en el item anterior usando R.

20. El número de veces que una red de computadoras se bloquea sigue un proceso de Poisson de parámetro igual a 2 bloqueos por semana. Hallar la probabilidad de que a) en 2 semanas no se bloquee. b) en un periodo de 4 semanas, haya exactamente 1 semana en la que no se bloquea. (Sugerencia: Notar que se está preguntando por cantidad de semanas, no de bloqueos.)