Centro de Investigación en Matemáticas (CIMAT). Unidad Monterrey

Inferencia estadística (Tarea 2)

Ana Beatriz Rodríguez Mendoza

11 de septiembre de 2018

1. Ejercicio 1

Una pareja decide tener hijos hasta el nacimiento de la primer niña. Calcule la probabilidad de que tengan más de 4 hijos. Suponga que las probabilidades de tener niño o niña son iguales. ¿Cuál es el tamaño esperado de la familia?

Éxito será cuando salga una niña.
$x =$número de hijos que tienen.
$P\left(\text{exito}\right) = 0.5$.

$$\begin{aligned} P\left(x > 4\right) &= 1 - \left[P\left(x = 4\right) + P\left(x = 3\right) + P\left(x = 2\right) + P\left(x = 1\right)\right] \\ &= 1 - \left[\left(0.5\right)^3\left(0.5\right) + \left(0.5\right)^2\left(0.5\right) + \left(0.5\right)^1\left(0.5\right) + \left(0.5\right)^0\left(0.5\right)\right] \\ &= 1 - \left[0.0625 + 0.125 + 0.25 + 0.5\right] \\ &= 1 - 0.9375 \\ &= 0.0625 \end{aligned}$$

La probabilidad de que tengan más de 4 hijos es 0.06.
$E\left(x\right) = \frac{1}{0.5} = 2$
El tamaño esperado de la familia es 4, 2 hijos y 2 padres.

2. Ejercicio 2

Cuando una máquina no se ajusta adecuadamente tiene una probabilidad 0.15 de producir un artículo defectuoso. Diariamente, la máquina trabaja hasta que se producen 3 artículos defectuosos. Se detiene la máquina y se revisa para ajustarla. ¿Cuál es la probabilidad de que una máquina mal ajustada produzca 5 o más artículos antes de que sea detenida? ¿Cuál es el número promedio de artículos que la máquina producirá antes de ser detenida?

a) Éxito cuando un artículo sea defectuoso.
$x =$ número de artículos que se producen hasta encontrar 3 éxitos.
$P\left(\text{éxito}\right) = 0.15$ constante.
$r = 3$ éxitos se detiene la máquina.

$$
\begin{aligned}
P\left(x \geq 5\right) &= 1 - \left[P\left(x = 3\right) + P\left(x = 4\right)\right] \\
&= 1 - \left[\binom{3-1}{3-1}\left(1 - 0.15\right)^{3-3}\left(0.15\right)^{3} + \binom{4-1}{3-1}\left(1 - 0.15\right)^{4-3}\left(0.15\right)^{3}\right] \\
&= 1 - \left[\binom{2}{2} \cdot 1 \cdot \left(0.15\right)^{3} + \binom{3}{2}\left(0.85\right)\left(0.15\right)^{3}\right] \\
&= 1 - \left[1 \cdot \left(0.003375\right) + 3 \cdot \left(0.85\right)\left(0.003375\right)\right] \\
&= 1 - \left[0.003375 + 0.0086\right] \\
&= 1 - \left[0.01198\right] \\
&= 0.988
\end{aligned}
$$

La probabilidad de que produzca 5 o más artículos antes de que sea detenida es 0.99.

b)

$$E\left(x\right) = \frac{3}{0.15} = 20$$

El número promedio de artículos que se producen antes de que la máquina sea detenida son 20.

3. Ejercicio 3

Los empleados de una compañía de aislantes son sometidos a pruebas para detectar residuos de asbesto en sus pulmones. Se le ha pedido a la compañía que envíe a tres empleados cuyas pruebas resulten positivas, a un centro médico para realizarles más análisis. Si se sospecha que el 40 % de los empleados tienen residuos de asbesto en sus pulmones, encuentre la probabilidad de que deban de ser analizados 10 trabajadores para poder encontrar a 3 con resultado positivo.

éxito igual a que el empleado tenga residuos de asbesto.
$x =$ número de empleados analizados.
$r = 3$ éxitos.
$P\,(\text{éxito}) = 0.4$.

$$
\begin{aligned}
P\,(x = 10) &= \binom{10-1}{3-1} (1-0.4)^{10-3} (0.4)^3 \\
&= \binom{9}{2} (0.6)^7 (0.064) \\
&= 36\,(0.028)\,(0.064) \\
&= 0.06
\end{aligned}
$$

La probabilidad de que deban ser analizados 10 trabajadores es 0.06.

4. Ejercicio 4

Para el siguiente ejercicio es necesario usar ***R****.*

a) *Considere una moneda desequilibrada que tiene probabilidad p de obtener águila. Usando el comando* ***sample****, escriba una función que simule N veces lanzamientos de esta moneda hasta obtener águila. La función deberá recibir como parámetros a la probabilidad p de obtener águila y al número N de veces que se repite el experimento; y tendrá que regresar un vector de longitud N que contenga el número de lanzamientos hasta obtener un águila en cada uno de los N experimentos.*

   Solución en el archivo de R.

b) *Usando la función anterior simule $N = 10^4$ veces una variable aleatoria* $\text{Geom}\,(p)$ *para $p = 0.5, 0.1, 0.01$. Grafique las frecuencias normalizadas en color azul. Sobre esta última figura empalme en rojo la gráfica de la función de masa correspondiente. ¿Qué observa?*

   Solución en el archivo de R.

c) *Repita el inciso anterior para $N = 10^6$. Además calcule el promedio y la desviación estándar de las simulaciones que realizó ¿Qué observa?*

Solución en el archivo de R.

5. Ejercicio 5

*Usando las ideas del inciso anterior escriba una función en **R** que simule N veces los lanzamientos de una moneda hasta obtener r águilas. La función deberá recibir como parámetros la probabilidad p de obtener águila, al número r de águilas a observar antes de detener el experimento y al número N de veces que se repite el experimento; y tendrá que regresar un vector de longitud N que contenga el número de lanzamientos hasta obtener las r águilas en cada uno de los N experimentos. Grafique las frecuencias normalizadas de los experimentos para $N = 10^6, p = 0.2, 0.1$ y $r = 2, 7$ y compárelos contra la función de masa de la distribución más adecuada para modelar este tipo de experimentos.*

Solución en el archivo de R.

6. Ejercicio 6

Considera X una v.a. con función de distribución F y una función de densidad f, y sea A un intervalo de la línea real $\mathbb{R}$. Definimos la función indicadora $1_A(x)$:

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{en otro caso} \end{cases}$$

Sea $Y = 1_A(x)$. Encuentre una expresión para la distribución acumulada de Y.

$$Y = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases} \tag{6.1}$$

Se conoce F_x, f_x, sin perdida de generalidad $A := [a, b]$ con $a \leq b \in \mathbb{R}$, como Y es una variable aleatoria discreta $F_y = \sum f_y$ con $f_y = P(Y = y) = P(1_A(x) = y)$, entonces f_y está dividido en los siguientes casos

$$\begin{aligned} P(Y = 1) &= P(x \in A) \\ &= P(a \leq x \leq b) \\ &= P(x \leq b) - P(x < a) \\ &= F_x(b) - F_x(a) \end{aligned}$$

$$
\begin{aligned}
P\left(Y=0\right) &= P\left(x \notin A\right) \\
&= P\left(x < a\right) + P\left(x > b\right) \\
&= F_x\left(a\right) + 1 - F_x\left(b\right) \\
&= 1 - P\left(Y=1\right)
\end{aligned}
$$

con la información que se tiene se observa que

$$Y \sim \text{Bernulli}\left(F_x\left(b\right) - F_x\left(a\right)\right)$$

por lo tanto

$$
F_Y\left(Y\right) = \begin{cases} 0 & \text{si } y \leq 0 \\ F_x\left(b\right) - F_x\left(a\right) & \text{si } 0 \leq y \leq 1 \\ 1 & \text{si } 1 \leq y \end{cases}
$$

7. Ejercicio 7

Entre las más famosas de todas las lluvias de meteoros están las Perseidas, que ocurren cada año a principios de agosto. En algunas áreas, la frecuencia de Perseidas visibles promedian ses por cada cuarto de hora. El modelo de probabilidad que describe a Y, el número de meteoros que una persona ve en un cuarto de hora, tiene la función de probabilidad

$$f_Y\left(y\right) = \tfrac{e^{-6}6^y}{y!} \quad y = 0, 1, \ldots$$

Encuentre la probabilidad de que una persona vea en un cuarto de hora determinado al menos la mitad de los meteoros que esperaría ver.

$y =$ número de meteoros que una persona ve en un cuarto de hora.
$\lambda = 6/\frac{1}{4}$hr con $f_Y\left(y\right) = \frac{e^{-6}6^y}{y!}$ $y = 0, 1, 2, \ldots$

Los meteoros que se esperan ver son 6 en un cuarto de hora ya que $E\left(x\right) = \lambda$.

$$
\begin{aligned}
P\left(x \geq \frac{\lambda}{2}\right) &= P\left(x \geq 3\right) \\
&= 1 \cdot \left[\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!}\right] \\
&= 1 - \left[0.00248 + 0.0149 + 0.0446\right] \\
&= 1 - \left[0.062\right] \\
&= 0.938
\end{aligned}
$$

La probabilidad de que una persona vea la mitad de meteoros esperados es 0.93 en un cuarto de hora.

8. Ejercicio 8

Las calificaciones de un estudiante de primer semestre en un examen de química se describen por la densidad de probabilidad

$$f_y\left(y\right) = 6y\left(1-y\right) \quad 0 \leq y \leq 1.$$

donde y representa la proporción de preguntas que el estudiante contesta correctamente. Cualquier calificación menor a 0.4 es reprobatoria. Responda lo siguiente

$$f_y = 6y\left(1-y\right) \quad 0 \leq y \leq 1 \quad f'y = 0 \Rightarrow f_y \in [0, 0.5]$$

$y :=$ proporción de preguntas que el estudiante responde correctamente.
$f_y :=$ calificación de un estudiante.
$f_y < 0.4 \Rightarrow$ calificación reprobatoria.

Se verifica

$$\begin{aligned}
\int_0^1 6y\left(1-y\right) \mathrm{d}y &= \int_0^1 \left[6y - 6y^2\right] \mathrm{d}y \\
&= \left[3y^2 - 2y^3\right] |_0^1 \\
&= \left(3\left(1\right)^2 - 2\left(1\right)^2\right) - 0 \\
&= 1
\end{aligned}$$

a) *¿Cuál es la probabilidad de que un estudiante repruebe?*

$$\begin{aligned}
P\left(f_y < 0.4\right) &= \int_0^{0.4} 6y\left(1-y\right) \mathrm{d}y \\
&= \int_0^{0.4} \left[6y - 6y^2\right] \mathrm{d}y \\
&= \left[3y^2 - 2y^3\right]_0^{0.4} \\
&= 3\left(0.4\right)^2 - 2\left(0.4\right)^3 - 0 \\
&= 0.352
\end{aligned}$$

La probabilidad de que un estudiante repruebe es de 0.35.

b) *Si 6 estudiantes toman el examen ¿Cuál es la probabilidad de exactamente 2 reprueben?*

$n = 6$ con $P = 0.35, x =$ número de estudiantes que reprueban.

$$
\begin{aligned}
P\left(x=2\right) &= \binom{6}{2}\left(0.35\right)^2\left(1-0.35\right)^{6-2} \\
&= 15\left(0.1225\right)\left(0.178\right) \\
&= 0.328
\end{aligned}
$$

La probabilidad de que dos reprueben es 0.32.

9. Ejercicio 9

Escriba una función en ***R*** *que simule una aproximación al proceso Poisson a partir de las 5 hipótesis que usamos en clase para construir tal proceso. Usando esta función, simule tres trayectorias de un proceso Poisson* $\lambda = 2$ *sobre el intervalo* $[0, 10]$ *y grafíquelas. Además simule* 10^4 *veces un proceso de Poisson* N *con* $\lambda = \frac{1}{2}$ *y hasta el tiempo* $t = 1$*. Haga un histograma de* $N\left(1\right)$ *en su simulación anterior y compare contra la distribución de Poisson correspondiente.*
Considere el intervalo $[0, T]$ *y un número real positivo* dt *que sea mucho más pequeño que la longitud de* $[0, T]$ *y que divida dicha longitud, digamos* $\frac{T}{dt} = 1000$*. Divida el intervalo* $[0, T]$ *en intervalos de longitud* dt *que tengan la forma* $\left(k \cdot dt, (k+1) \cdot dt\right], k = 0, 1, 2, \ldots, \left(\frac{T}{dt} - 1\right)$*. Para cada uno de estos intervalos simule una v.a Bernoulli*$\left(\lambda \cdot dt + 10^{-6}\right)$ *y guarde su resultado en un vector del tamaño adecuado.*

Solución en el archivo de R.