Práctica N°2

Ivan Fernando Mujica Mamani

Análisis Estadístico I

Maestría en Ciencia de Datos v.2

Octubre – 2020

1. Ejercicio 27

En una encuesta de pretemporada de futbol americano de la NCAA 2001 se pregunto: "¿Este año habrá un equipo del Big Ten o del Pac'10 en el juego del Rose Bowl?" De los 13429 interrogados, 2961 dijeron que habria uno del Big Ten. 4494 señalaron que habria uno del Pac-10 y 6823 expresaron que niel Big Ten ni el Pac'10 tendria un equipo dentro el Rose Bowl.

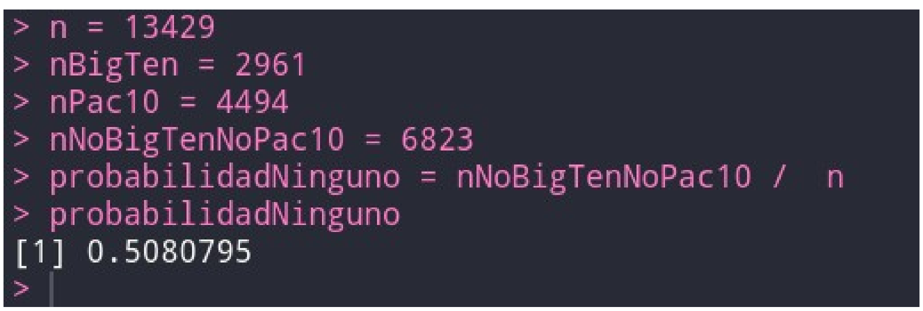
Respuestas

A. *¿Cuál es la probabilidad de que el interrogado responda que ni el Big Ten ni el Pac-10 tendrán un equipo en el Rose Bowl?*

Aplicando la teoría básica de la probabilidad, tenemos:

$$p(Ninguno) = \frac{casos\,favorables}{total\,de\,casos} = \frac{6823}{13429} = 0.5081$$

Aplicación en R



```
> n = 13429
> nBigTen = 2961
> nPac10 = 4494
> nNoBigTenNoPac10 = 6823
> probabilidadNinguno = nNoBigTenNoPac10 /  n
> probabilidadNinguno
[1] 0.5080795
>
```

B. *¿De que afirme que el Big Ten o el Pac-10 tendrán un equipo en el campeonato Rose Bowl?*

Aplicando la regla del complemento:

$$P(Cualquiera) = (1) - P(Ninguno) = (1) - (0.5081) = 0.4919$$

Aplicación en R:

```
> # B
> probabilidadCualquiera = 1 - probabilidadNinguno
> probabilidadCualquiera
[1] 0.4919205
>
```

C. *Halle la probabilidad de que la respuesta sea que tanto el Big Ten como el Pac 10 tendrán un equipo en el Rose Bowl.*

Aplicando la fórmula

$P(A \lor B)=P(A)+P(B)-P(A \land B)$

$P(Ambos)=P(Big\,10)+P(Pac\,10)-P(Cualquiera)$

$P(Ambos)=\frac{2961}{13429}+\frac{4494}{13429}-0.4919=0.0632$

Aplicacion en R

```
> probabilidadBigTen = 2961 / 13429
> probabilidadPac10 = 4494 / 13429
> probabilidadAmbos = probabilidadBigTen + probabilidadPac10 - probabilidadCualquiera
> probabilidadAmbos
[1] 0.06322139
```

Existe una probabilidad de 6.32% de que la respuesta sea ambos equipos.

2. Ejercicio 28

En una encuesta aplicada a los suscriptores de una revista se encontró que en los ultimos 12 meses 45.8% habián rentado un automovil por razones de trabajo, 54% por razones personales y 30% por razones de trabajo y personales.

Dado que:

$P(Trabajo)=45.8\%=0.458$

$P(Personal)=54\%=0.54$

$P(Trabajo\ y\ Personal)=30\%=0.30$

Respuestas

A. *¿Cuál es la probabilidad de que un suscriptor hay rentado un automovil en los ultimos 12 meses por razones de trabajo o por razones personales?*

Aplicando la formula:

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

$$P(A \vee B) = P(A) + P(B) - P(A \wedge B)$$

$$P(Trabajo \vee Personal) = 0.548 + 0.54 - 0.30 = 0.698$$

Aplicación en R:

```
> probabilidadTrabajo = 0.458
> probabilidadPersonal = 0.54
> probabilidadAmbos = 0.30
> #A
> probabilidadCualquiera  = probabilidadTrabajo + probabilidadPersonal - probabilidadAmbos
> probabilidadCualquiera
[1] 0.698
>
```

Existe una probabilidad del 69.8% de que se alquile un vehiculo ya sea por trabajo o motivos personales.

B *¿Cuál es la probabilidad de que un suscriptor no haya rentado un automóvil en los últimos 12 meses ni por razones de trabajo ni por razones personales?*

Aplicando la fórmula de complemento:

$$P(Ninguno) = (1) - P(Cualquiera) = (1) - (0.698) = 0.302$$

```
> # B
> probabilidadNinguno = 1 - probabilidadCualquiera
> probabilidadNinguno
[1] 0.302
>
```

Existe una probabilidad del 30.2% de que se alquile un vehiculo ni por trabajo ni por motivos personales.

3. Ejercicio Probabilidad Condicional

Suponga dos eventos, A y B, que son mutuamente excluyentes. Admita, además, que P(A) = 0.30 y P(B) = 0.40

A. Obtenga P(A intersección B)

A y B son mutuamente excluyentes, por lo tanto sabemos que:

$P(A \wedge B)=0$

B. Calcule P(A | B)

$P(Ambos)=\frac{2961}{13429}+\frac{4494}{13429}-0.4919=0.0632$

Aplicación en RÑ

```
> # Ejercicio Probabilidad Condicional
> probablidadB = 0.40
> probabilidadAUnionB = 0
> probabilidadADadoB = probabilidadAUnionB / probablidadB
> probabilidadADadoB
[1] 0
>
```

C. *Un estudiante argumenta que los conceptos de eventos mutuamente excluyentes y eventos independientes son en realidad lo mismo y que si los eventos son mutuamente excluyentes deben ser tambien independientes. Esta usted de acuerdo ? Use la información de probabilidades para justificar su respuesta.*

No. Porque si dos eventos son mutuamente excluyentes, entonces no ocurren al mismo tiempo, por lo tanto no pueden ser independientes.

4. Ejercicio de probabilidad posterior con el Teorema de Bayes

Un Banco local revisa su política de tarjetas de crédito con objeto de retirar algunas de ellas. En el pasado aproximandamente 5% de los tarjeta habientes incumplieron, *dejando al banco sin posibilidad de cobrar el saldo pendiente. De manera que el director estableció una probabilidad previa de 0.05 de que un tarjetahabiente no cumpla. El banco encontró tambien que la probabilidad previa de 0.05 de que un tarjetahabiente no cumpla. El banco encontró también que la probabilidad e que un cliente que es cumplido no haga un pago mensual es de 0.20. Por supuesto que la probabilidad de no hacer un pago mensual entre los que incumplen es 1.*

dado que :

P(incumple) = 5% = 0.05

P(falle/no incumple) = 0.20

P(falle/incumple) = 1

A. dado que el cliente no hizo el pago de uno o mas meses, calcule la probabilidad posterior de que el cliente no cumpla.

según el teorema de Bayess:

$$P(Incumple/falle)=\frac{P(Incumple)*P(falle/Incumple)}{P(Incumple)*P(falle/Incumple)+P(Cumple)*P(falle/cumple)}$$

$$P(Incumple/falle)=\frac{0.05*1}{0.05*1+(1-0.05)*0.20}=0.2083$$

Aplicación en R:

```
> # Ejercicio probailidad posterior con el teorema de bayess
> probIncumple = 0.05
> probFalleDadoNoIncumple = 0.20
> probFalleDadoIncumple = 1
> probFalleDadoIncumple = (probIncumple * probFalleDadoIncumple) / ( probIncumple * probFalleDadoIncumple
 + (1-probIncumple) * probFalleDadoNoIncumple )
> probFalleDadoIncumple
[1] 0.2083333
>
```

Existe una probabilidad de 20.83% de que un cliente Incumpla dado que falla un mes.

B. *El banco deseara retirar sus tarjetas si la probabilidad de que un cliente no cumpla es mayor que 0.20. ¿Debe retirar el bando una tarjeta si el cliente no hace el pago mensual?*

Si, por que la probabilidad de que incumpla dado que falle es mayor a 0.20.

5. Ejercicio De Distribución Discreta Binomial

Una encuesta de Harris Interactive para InterContinental Hoteld and Resorts preguntó: “Cuando viaja al extranjero, ¿suele aventurarse usted solo para conocer la cultura o prefiere permanecer con el grupo de su tour y apegarse al itinerario?” Se encontró que 23% prefiere permanecer con el grupo de su tour (USA Today, 21 de enero de 2004).

Respuestas

dado que:

$P = 0.23$

$n = 6$

A.¿Cuál es la probabilidad de que en una muestra de seis viajeros, dos prefieran permanecer con su grupo?

$$f(x)=P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}$$

dado que la combinatoria de x tomados de n es:

$$\binom{n}{x}=\frac{n!}{x!(n-x)!}$$

$$\binom{6}{2}=\frac{6!}{2!(4-2)!}=15$$

$$f(2)=P(X=2)=\binom{6}{2}0{,}23^2(1-0.23)^{6-2}=0.278939$$

Aplicación en R

```
> # Ejercicio distribuciones
> # Binomial
> # A
> p=0.23
> n=6
> dbinom(2,n,p)
[1] 0.2789394
> 
```

Existe una probabilidad de 27.89% de que 2 viajeros se queden con el grupo.

B. *¿De que en una muestra de 10 viajeros, ninguno prefiera permanecer con su grupo?*

N = 10

$$\binom{10}{0}=\frac{10!}{2!(10-0)!}=1$$

$$f(0)=P(X=0)=\binom{10}{0} * 0{,}23^{0} * (1-0.23)^{10-0}=0.0733$$

Aplicacion en R

```
> # B
> p = 0.23
> n=10
> dbinom(0, n, p)
[1] 0.0732668
>
```

6. Ejercicio De Distribución De Poison

Durante el periodo en que una universidad recibe inscripciones por teléfono, llegan llamadas a una velocidad de una cada dos minutos.

Dado que la distribucion de Poison esta definida por:

$$f(x)=P(X=x)=\frac{e^{-\mu}\mu^{x}}{x!}$$

donde e es 2.71828

A.*¿Cuál es el número esperado de llamadas en una hora?*

La media esperada es la media de dos minutos por el numero de periodos en una hora, en este caso 30

$$\mu=1x30=30$$

B.*¿Cuál es la probabilidad de que haya tres llamadas en cinco minutos?*

La media esperada es la media de dos minutos por el numero de periodos en 5 minutos.

$$\mu=1x2.5=2.5$$

k = 3

$$f(3)=P(X=3)=\frac{e^{-2.5}2.5^{3}}{3!}=0.2138$$

Aplicacion en R

```
> # Poison
> media = 2.5
> k = 3
> dpois(k, media)
[1] 0.213763
>
```

Existe una probabilidad de 21.37% de que existan 3 llamadas en 5 minutos.

C.*¿De que no haya llamadas en un lapso de cinco minutos?*

$\mu=1x\,2.5=2.5$

k = 0

$$f(0)=P(X=0)=\frac{e^{-2.5}2.5^{0}}{0!}=0.0821$$

Aplicación en R:

```
> # C
> media = 2.5
> k = 0
> dpois(k, media)
[1] 0.082085
>
```

Existe una probabilidad de 8.2% de que no existan llamadas en un lapso de 5 minutos.

7. Ejercicio De Distribución Hipergeométrica

En una encuesta realizada por Gallup Organization, se les preguntó a los interrogados, “Cuál esel deporte que prefieres ver”. Futbol y básquetbol ocuparon el primero y segundo lugar de preferencia (www.gallup.com, 3 de enero de 2004). Si en un grupo de 10 individuos, siete prefieren futbol y tres prefieren básquetbol. Se toma una muestra aleatoria de tres de estas personas.

Dado que la distribucion hipergeometrica esta dada por:

$$f(x)=P(X=x)=\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

La media y la varianza de una distribución hipergeométrica son las siguientes.

Dado que N = Tamaño de la población = 10

n = muestra aleatoria = 3

r = Numero de éxitos = 7

A.¿Cuál es la probabilidad de que exactamente dos prefieren el futbol?

$$P(X=2)=\frac{\binom{7}{2}\binom{10-7}{3-2}}{\binom{10}{3}}=\frac{21}{40}=0.525$$

Lo que significa que existe una probabilidad del 52.2% de que dos personas prefieran el futbol.

Aplicación en R

```
> # Hipergeometrica
> # A
> n = 3
> r = 7
> N = 10
> dhyper(2,r,N -r,n)
[1] 0.525
>
```

B.*¿De que la mayoría (ya sean dos o tres) prefiere el futbol?*

Para la respuesta necesitamos X >= 2, es decir X =3, porque esta dada por la suma de las probabilidades mayores a 2.

$$P(X=3)=\frac{\binom{7}{3}\binom{10-7}{3-3}}{\binom{10}{3}}=\frac{7}{24}=0.2917$$

$$P(X\geq 2)=f(2)+f(3)=0.525+0.2917=0.0.8167$$

Aplicación en R

```
> #B
> f2 = dhyper(2,r,N -r,n)
> #B
> f2 = dhyper(2,r,N -r,n)
> f3 = dhyper(3,r,N -r,n)
> f2 + f3
[1] 0.8166667
>
```

Lo que significa que existe una probabilidad de un 81.66% de que dos o mas prefieran futbol.

8. Ejercicio De Distribución Uniforme

En las botellas de un detergente líquido se indica que el contenido es de 12 onzas por botella. En la operación de producción se llenan las botellas uniformemente de acuerdo con la siguiente función de densidad de probabilidad.

$$f(x)=\begin{cases}\frac{1}{b-a}\ para\, a\leq x\leq b\\ 0\ para\, x\leq a\, o\, x\leq b\end{cases}$$

$$f(x)=\begin{cases}8\ para\ 11.975\leq x\leq 12.100\\ 0\ en\, otro\, caso\end{cases}$$

A. *¿Cuál es la probabilidad de que el contenido de una botella esté entre 12 y 12.05 onzas?*

a = 11.975
b = 12.100

$$P(x)=P(12\leq x\leq 12.05)=\int_{12}^{12.5} f(x)\,dx=8(12.05)-8(12)=0.4$$

Lo que significa de que existe una probabilidad de 40% de que una botella este entre 12 y 12.05 onzas.
Aplicacion en R:

```
> #Uniforme
> # A
> a = 11.975
> b = 12.100
> min = 12
> max=12.05
> fx = 8
> prob1205 = punif(12.05,min=a, max=b, lower.tail=TRUE)
> prob12 = punif(12,min=a, max=b, lower.tail=TRUE)
> prob1205 - prob12
[1] 0.4
>
```

B.¿De que el contenido de una botella sea 12.02 onzas o más?

$$P(x>12.02)=\int_{12.02}^{12.100} f(x)dx=8(12.10)-8(12.02)=0.064$$

Aplicacion en R

```
> # B
> prob121 = punif(12.100,min=a, max=b, lower.tail=TRUE)
> prob1202 = punif(12.02,min=a, max=b, lower.tail=TRUE)
> prob121 - prob1202
[1] 0.64
```

Se interpreta que existe una probabilidad del 64% de que el contenido de la botella sea mayor a 12.02 onzas.

9. Ejercicio De Distribución Exponencial

El periodo de vida de un interruptor eléctrico tiene una distribución exponencial con un promedio de fallo de 2 años ¿Cuál es la probabilidad de que falle después de tres años?

La distribución exponencial es una distribución de probabilidad continua con un parámetro $\lambda>0$ cuya función de densidad es:

$$f(x)=\begin{cases} \lambda e^{-\lambda x}, \wedge\ para\, x \geq 0 \\ 0, \wedge\ en\, caso\, contrario \end{cases}$$

El valor esperado y la varianza de una variable aleatoria X con distribución exponencial son:

$E(X)\frac{1}{\lambda}=,V(X)=\frac{1}{\lambda^2}$

***X*: Tiempo de fallo de un interruptor eléctrico**

Nos dicen que la distribución es exponencial cuya media es 2, por lo que $\lambda\frac{1}{2}$, y tenemos una distribución exponencial $X\exp\left(\frac{1}{2}\right)$

$P(X>3)=1-P(X\leq 3)=1-\left(e^{\frac{-1}{2}3}\right)=1-1+e^{-3/2}=e^{-3/2}=0{,}2231$

Es decir, que la probabilidad de que el interruptor falle después de tres años es del 22.13%.

Algoritmo en R:

```
> # Exponencial
> x=3
> rate = 1/2
> pexp(3, rate=rate, lower.tail=F)
[1] 0.2231302
```

10. Ejercicio De Distribución Normal

El precio promedio de las acciones que pertenecen a S&P500 es de $30 y la desviación estándares $8.20 (BusinessWeek, Special Annual Issue, primavera de 2003). Suponga que los precios de las acciones están distribuidos normalmente.

$\mu=media=30$

$\sigma=desviación\,estandar=8.20$

A. ¿Cuál es la probabilidad de que el precio de las acciones de una empresa sea por lo menos de $40?

calculamos z para media 40.

$$Z=\frac{X-\mu}{\sigma}=\frac{40-30}{8.20}=1.22$$

Segun la tabla de la distribucion normal se tiene:

$$P(z>1.22)=1-P(X<1.22)=1-08888=0.1112$$

Aplicacion en R:

```
> # Normal
> #A
> media = 30
> desviacion = 8.20
> 1- pnorm(40, media, desviacion)
[1] 0.1113249
>
```

B. ¿De que el precio de las acciones de una empresa no sea mayor a $20?

$$Z=\frac{X-\mu}{\sigma}=\frac{20-30}{8.20}=-1.22$$

Segun tablas se tiene:

$$P(X<-1.22)=0.1112$$

C. *¿De cuánto deben ser los precios de las acciones de una empresa para que esté entre las 10% mejores?*

Calculamos z para 1 – 0.1 al ser que se necesita el 90 %

P(Z<0.90) = 1.28

$$X=\mu+Z * \sigma=30+1.28 * 8.2=40.496$$

Los precios deben ser 40.496 para que la empresa se situe entre los 10 primeros.