

Taller Probabilidad [1era Parte - G4 y G5]

Bioestadística 1 - UNAL Medellín

Edimer David Jaramillo

18 de marzo de 2018

Contents

Nota	1
Primera parte del taller	1
Ejemplos	1
Fecha y modo de entrega	4

Nota

Nota: el taller consta de dos partes:

- Probabilidades obtenidas a partir de los resultados de la encuesta [**primera parte**].
- Ejercicios de probabilidad desarrollados en clase [**segunda parte**].

Primera parte del taller

Basado en las consultas formuladas (o ideas nuevas) en el primer taller acerca de la encuesta y teniendo en cuenta las notas de clase de probabilidad, cada grupo deberá obtener la probabilidad de 5 eventos o sucesos cualquiera respecto a los resultados de la encuesta.

Ejemplos

1.1 Ejemplo: si se elige al azar un estudiante del curso de Bioestadística 1, teniendo en cuenta a ambos grupos y que haya contestado la encuesta (56 estudiante de 60 respondieron la encuesta), ¿cuál es la probabilidad de elegir un estudiante que tenga Smartphone y haya tenido contacto con algún lenguaje de programación?

Solución: para solucionar el problema planteado, se recurre a obtener las frecuencias absolutas de cada uno de los eventos, teniendo como eventos los dos siguientes:

- **Evento A:** que sí tenga Smartphone
- **Evento B:** que sí haya tenido contacto con algún lenguaje de programación

1. Se obtienen las frecuencias absolutas para dichos eventos:

```
encuesta <- read.csv("Encuesta.csv")
EventoA <- encuesta$Smartphone
EventoB <- encuesta$LenguajeP
table(EventoA, EventoB)
```

```
##          EventoB
## EventoA No Si
##      No  3  0
##      Si 36 17
```

2. $P(A \cap B)$: los que sí tienen celular y sí han tenido contacto con algún lenguaje de programación:

$$P(A \cap B) = \frac{17}{56} = 0.3035$$

Nota: a través de la propiedad de la intersección se llega al mismo resultado. Con la siguiente fórmula:

$$P(A \cap B) = P(A|B)P(B)$$

donde,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Ejemplo 1.2: si se elige al azar un estudiante del curso de Bioestadística 1, teniendo en cuenta a ambos grupos y que haya contestado la encuesta (56 estudiante de 60 respondieron la encuesta), ¿cuál es la probabilidad de elegir un estudiante que se demore menos de 30 minutos desde su casa a la universidad dado que se moviliza en carro o moto?

Solución: se definen los eventos de importancia:

- **Evento A:** que se demore menos de 30 minutos en llegar a la universidad
- **Evento B:** que se movilice en carro o moto

El evento A se puede clasificar como “Sí” y “No”, obteniendo lo siguiente:

- Los que sí se demoran menos de 30 minutos en llegar a la universidad, lo conforman aquellos que respondieron A o B a dicha pregunta. Por tanto, aquellos que respondieron C o D son los que no se demoran menos de 30 minutos en llegar a la universidad. Las frecuencias absolutas se obtienen de la siguiente manera:

```
encuesta <- read.csv("Encuesta.csv")
EventoA <- encuesta$Casa_U
table(EventoA)
```

```
## EventoA
##  A  B  C  D
##  5 33 13  5
```

De 56 personas, 38 se demoran menos de 30 minutos en llegar a la universidad (5 + 33), el resto demoran más de dicho tiempo.

El evento B también se puede clasificar como “Sí” y “No”, obteniendo lo siguiente:

- Los que sí se movilizan en carro o moto lo conforman aquellos que respondieron la opción D en dicha pregunta. Por tanto, aquellos que respondieron A, B o C son los que no se movilizan en carro o moto. Las frecuencias absolutas se obtienen de la siguiente manera:

```
EventoB <- encuesta$Transporte
table(EventoB)
```

```
## EventoB
##  A  B  C  D
##  6 11 31  8
```

De 56 personas, 8 utilizan carro o moto como medio de transporte para dirigirse a la universidad.

- Ahora es necesario obtener las frecuencias absolutas de los eventos “cruzados”:

```
table(encuesta$Casa_U, encuesta$Transporte)
```

```
##
##      A  B  C  D
##  A  0  2  0  3
##  B  5  9 15  4
##  C  1  0 11  1
##  D  0  0  5  0
```

Con esta tabla se obtiene lo siguiente:

- Se sabe que 38 personas son las que demoran menos de 30 minutos en llegar a la universidad, es decir, las que respondieron A o B, que en la tabla corresponde a las dos primeras filas ($0 + 2 + 0 + 3 + 5 + 9 + 15 + 4$). De esas 38 personas, 7 se movilizan en carro o moto ($3 + 4$), es decir, las que respondieron la opción D, que en la tabla corresponde a la columna 4; las 31 personas restantes, aunque demoran menos de 30 minutos en llegar a la universidad, no se movilizan en carro.
- Como 38 personas demoran menos de 30 minutos, 18 restantes demorarán más de ese tiempo, es decir, aquellos que respondieron la opción C o D, que en la tabla corresponde a las filas 3 y 4 ($1 + 0 + 11 + 1 + 0 + 0 + 5 + 0$). De esas 18 personas que se tardan más de 30 minutos en llegar a la universidad, 1 se moviliza en carro o moto (opción D) y 17 en otro medio de transporte.
- Con la información anterior es posible construir una tabla como la siguiente:

```
tabla_frecuencias <- data.frame(EventoA = c(7, 31),
                                EventoB = c(1, 17))

rownames(tabla_frecuencias) <- c("SiB", "NoB")
colnames(tabla_frecuencias) <- c("SiA", "NoA")
```

	SiA	NoA
SiB	7	1
NoB	31	17

Con la tabla anterior es posible responder al planteamiento inicial a través de la probabilidad condicional, que sugiere lo siguiente:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

donde,

$$P(A \cap B) = \frac{7}{56} = 0.125$$

y

$$P(B) = \frac{8}{56} = 0.1428$$

entonces:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.125}{0.1428} = 0.8753$$

Fecha y modo de entrega

- Hasta el viernes **23 de marzo de 2018** hasta las **6:00 pm** (aula 50-230).
- Esta primera parte deberá ser entregada a mano, con las probabilidades de interés desarrolladas.
- La segunda parte del taller se desarrollará en clase los días **22 y 23 de marzo de 2018**, para los grupos 5 y 4, respectivamente.