Esperanza y varianza v.a. discretas

28 de abril de 2020

Ejemplo 1: Televisión

Una empresa proveedora de servicio de televisión por cable tiene 20000 clientes en una zona de la ciudad de Buenos Aires y alrededores. Cada cliente puede optar por contratar de 1 a 5 paquetes de señales (el abono básico consiste en un solo paquete y cada uno de los otros paquetes incluye grupos de señales temáticas o premium)

La empresa quiere poder predecir cuántos paquetes contratará cada cliente con el menor error posible.

La distribución del número de paquetes contratados por los clientes es:

x	1	2	3	4	5
número de clientes	7500	5500	3500	2000	1500
proporción de clientes	0.375	0.275	0.175	0.100	0.075

¿Cuántos paquetes tienen contratados, en promedio, los clientes de esta zona?

Ejemplo 1: Televisión

Cantidad promedio de paquetes por cliente:

$$\frac{1 \cdot 7500 + 2 \cdot 5500 + 3 \cdot 3500 + 4 \cdot 2000 + 5 \cdot 1500}{20000} = \frac{44500}{20000} = 2.225,$$

que es lo mismo que

$$1 \cdot 0.375 + 2 \cdot 0.275 + 3 \cdot 0.175 + 4 \cdot 0.10 + 5 \cdot 0.075 = 2.225$$

Sea X el número de paquetes contratados por un cliente elegido al azar.

$$R_X = \{1, 2, 3, 4, 5\}$$

x	1	2	3	4	5
$p_X(x)$	0.375	0.275	0.175	0.100	0.075

La cantidad promedio de paquetes por cliente es:

$$1 \cdot p_X(1) + 2 \cdot p_X(2) + 3 \cdot p_X(3) + 4 \cdot p_X(4) + 5 \cdot p_X(5) = 2.225$$

Definición de esperanza de una v.a. discreta

Sea X una v.a. discreta que toma valores en R_X con función de probabilidad puntual $p_X(x)$, la esperanza o valor esperado de X se define como

$$E(X) = \sum_{x \in R_X} x p_X(x)$$

Ejemplo: Dadoss

Se tiran dos Dadoss equilibrados

$X =$ cantidad de resultados pares

$$E(X) = \sum_{x \in R_X} x p_X(x)$$

+

$$\begin{aligned} &= 0 \cdot p_X(0) + 1 \cdot p_X(1) + 2 \cdot p_X(2) \\ &= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} \\ &= 1. \end{aligned}$$

Ejemplo: alumnos del turno tarde

Se elige un alumno del turno tarde al azar.

$$V = \begin{cases} 1 \text{ si el alumno tiene un equipo de fútbol favorito} \\ 0 \text{ si no} \end{cases}$$

```
encuesta <- read.csv("~/Dropbox/proba_c_2020/clases_practicas_tarde/encuesta.csv")
mean(encuesta$Equipo.de.futbol.favorito != "No tengo equipo favorito")
```

```
## [1] 0.5588235
```

```
round(mean(encuesta$Equipo.de.futbol.favorito != "No tengo equipo favorito"),2)
```

```
## [1] 0.56
```

$$E(V) = 0 \cdot (1 - 0.56) + 1 \cdot 0.56 = 0.56$$

La esperanza como centro de masa

$E(X)$ es el centro de gravedad de la función de probabilidad puntual. Es decir que si imaginamos que sobre cada valor posible de X, x_i , colocamos una masa $p_X(x_i)$, el punto de equilibrio del sistema es $E(X)$. En este sentido, podemos decir que $E(X)$ es una medida del "centro" de la distribución.

Ejemplo: Dados.

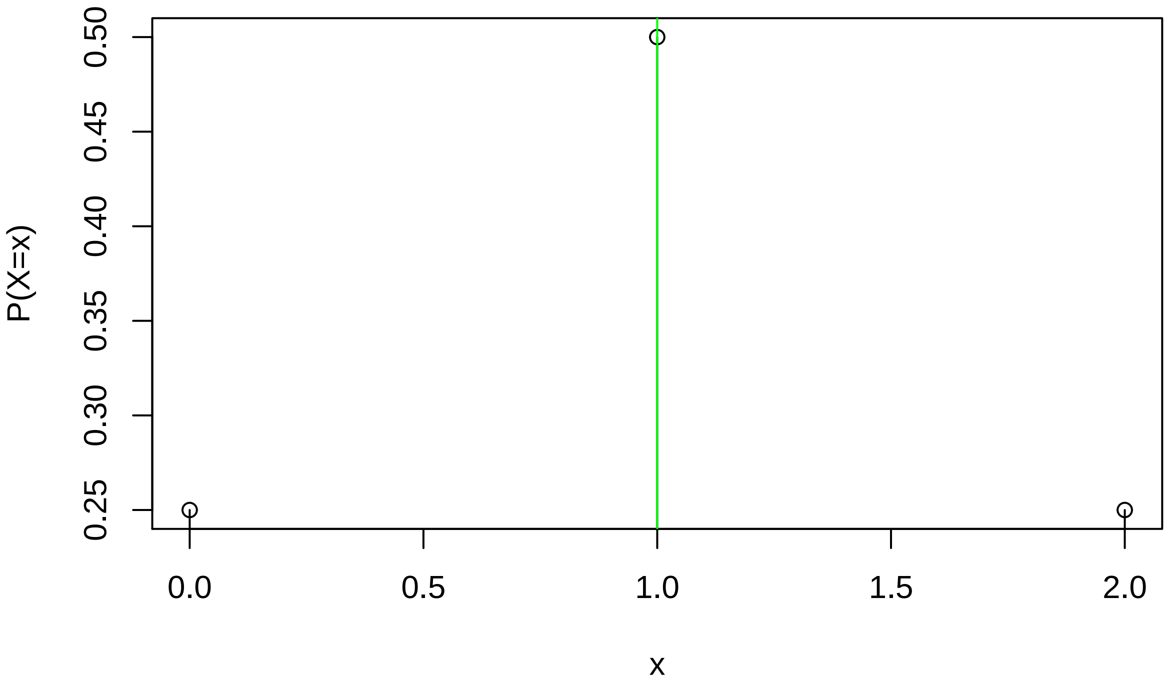
```
rx <- 0:2
px <- c(9/36, 18/36, 9/36)
fpp <- cbind(rx, px)
colnames(fpp) <- c("x","P(X=x)")
fpp
```

```
##      x P(X=x)
## [1,] 0   0.25
## [2,] 1   0.50
## [3,] 2   0.25
```

Ejemplo: Dados.

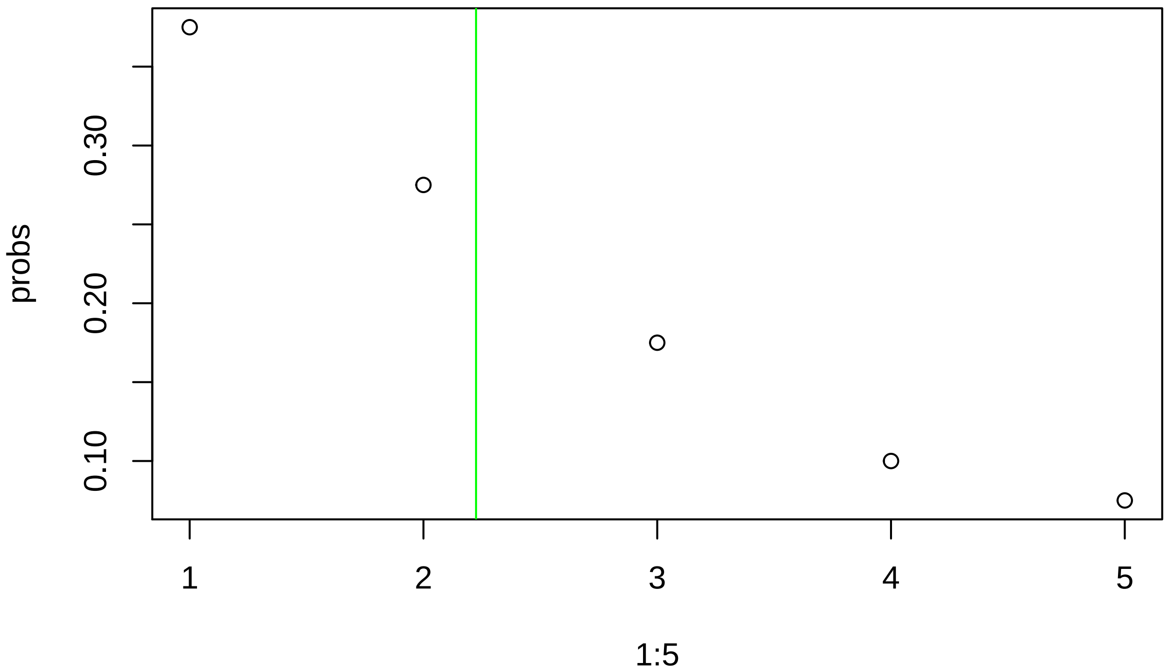
```
plot(rx, px)
lines(c(rx[1],rx[1]), c(0,px[1]))
lines(c(rx[2],rx[2]), c(0,px[2]))
lines(c(rx[3],rx[3]), c(0,px[3]))
```

Ejemplo: Dados.



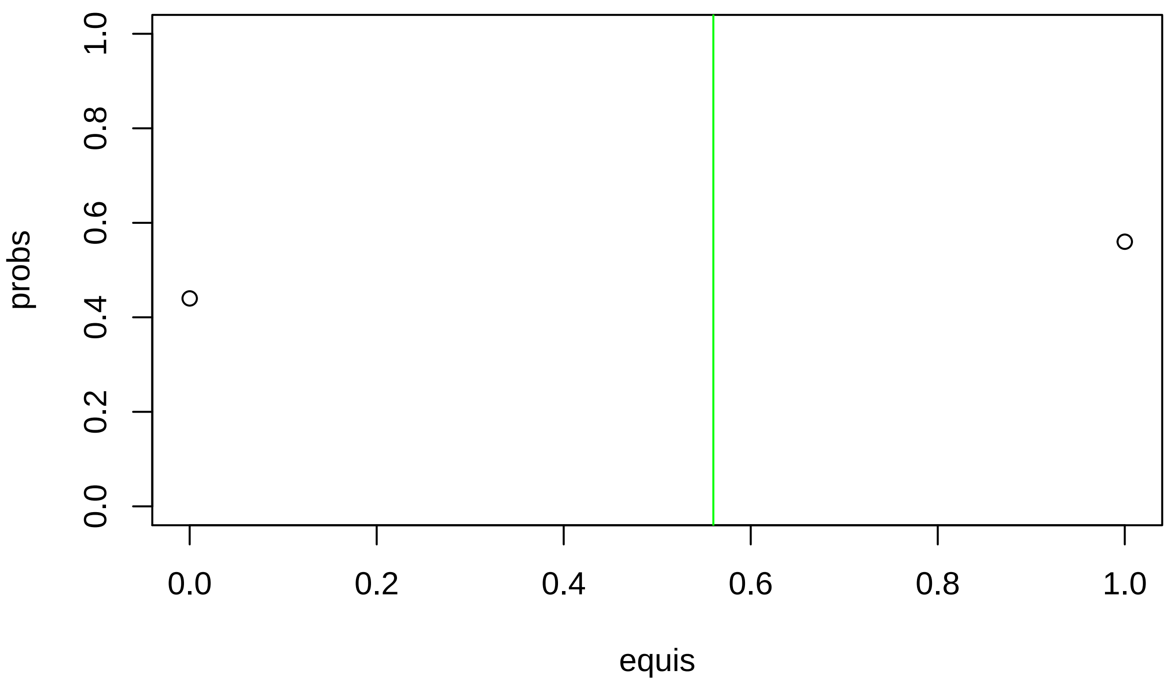
Ejemplo: Televisión

```
equis <- 1:5
probs <-  c(0.375, 0.275, 0.175,  0.1, 0.075)
plot(1:5, probs)
abline(v=2.225, col="green")
```



Ejemplo: alumnos del turno tarde

```
equis <- 0:1
probs <-  c(0.44,0.56)
plot(equis, probs, ylim= c(0,1))
abline(v=0.56, col="green")
```



La esperanza como límite de las frecuencias relativas.

La "ley de los grandes números" dice que si se repite indefinidamente un experimento aleatorio y se observa cuanto vale X en cada repetición, el promedio de los resultados obtenidos tiende a estabilizarse en un número que es $E(X)$, si es que ésta existe.

Veremos cómo formalizar este resultado más adelante.

Ejemplo: Dados

```
sample(1:6, 2, replace = TRUE)
```

```
## [1] 4 4
```

```
resultados <- replicate(10000, sample(1:6, 2, replace = TRUE))
```

Para ver si un número es par o no busco el resto de dividir por 2:

```
4%%2
```

```
## [1] 0
```

```
101%%2
```

```
## [1] 1
```

Ejemplo: Dadoss.

Estimación de la esperanza de X= cantidad de resultados pares

```
sample(1:6, 2, replace = TRUE)
```

```
## [1] 3 2
```

```
resultados <- replicate(1000, sample(1:6, 2, replace = TRUE))
resultados[,1:10]
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
## [1,]    4    6    2    4    1    2    6    2    1     6
## [2,]    6    2    1    1    6    5    6    6    6     6
```

```
cant_pares <- colSums(1-resultados%%2)
mean(cant_pares)
```

```
## [1] 1.018
```

Ejemplo: Televisión

Supongamos que el costo del servicio Y es función del número de paquetes contratado, según la siguiente fórmula:

$$Y = 30(X+1)$$

Hallar $E(Y)$

Como $R_X = \{1,2,3,4,5\}$, $R_Y = \{60,90,120,150,180\}$

Ejemplo: Televisión

Se tiene

x	1	2	3	4	5
$p_x(x)$	0.375	0.275	0.175	0.100	0.075

y

y	60	90	120	150	180
$p_Y(y)$	0.375	0.275	0.175	0.100	0.075

$$E(Y) = \sum_{y \in R_Y} y p_Y(y) = \sum_{y \in R_Y} 30(x+1) p_X(x)$$

$$= 60 \cdot 0.375 + 90 \cdot 0.275 + 120 \cdot 0.175 + 150 \cdot 0.10 + 180 \cdot 0.075 = 96.75$$

Esperanza de una función de una variable aleatoria

Proposición: Si la v.a. X tiene función de probabilidad puntual $p_X(x)$ entonces la esperanza de cualquier función real $h(X)$, está dada por

$$E(h(X)) = \sum_{x \in R_X} h(x) p_X(x)$$

Propiedades

1) Linealidad de la esperanza: Si $a, b \in \mathbb{R}$ son constantes, entonces

$$E(aX+b) = aE(X) + b$$

Dem:

$$E(h(X)) = E(aX+b) = \sum_{x \in R_x} (ax+b) p_X(x) = a \sum_{x \in R_x} x p_x(x) + b \sum_{x \in R_x} p_x(x) = aE(X) + b$$

2) Si X es una v.a. tal que $P(X=c) = 1$, entonces $E(X) = c$.

Dem:

$$E(X) = cp_X(c) = c$$

Ejemplo: Televisión (de nuevo)

Supongamos que el costo del servicio Y es función del número de paquetes contratado, según la siguiente fórmula:

$$Y = 30(X + 1)$$

Hallar $E(Y)$

$$E(Y) = E(30X + 30) = 30 \cdot 2.25 + 30 = 96.75$$

La esperanza como mejor predictor

Y es el ingreso de la empresa de televisión por cliente.

Cuándo la empresa necesita usar un número para aproximar, o predecir Y, utilizará su esperanza.

¿Cuál es el error que cometemos al hacer esta aproximación?

La diferencia entre la variable aleatoria y su esperanza:

$$Y - E(Y)$$

La diferencia entre la variable aleatoria y su esperanza al cuadrado es una manera de medir la distancia entre la variable aleatoria y su esperanza:

$$(Y - E(Y))^2$$

Pero $(Y - E(Y))^2$ no es un número, es una nueva v.a. Queremos un número que nos resuma esa diferencia

$$E\left((Y - E(Y))^2\right)$$

Ese número se llama varianza.

Definición de varianza

La varianza de una variable aleatoria X es

$$V(X) = E\left((X - E(X))^2\right)$$

La varianza de una variable aleatoria mide cuan dispersa esta ésta alrededor de su esperanza.

La esperanza como mejor predictor

Proposición La esperanza es la constante que mejor aproxima o predice a una variable aleatoria, en el sentido de que minimiza la función

$$t \mapsto E\left((X-t)^2\right)$$

Demostración

$$\begin{aligned} f(t) &= E\left((X-t)^2\right) \\ f'(t) &= -2E\left((X-t)\right) = -2(E\left(X\right)-t) \\ f'(t) &= 0 \Leftrightarrow t = E(X) \end{aligned}$$

Además $f''(t) = 2 > 0$. Por lo tanto, en $t = E(X)$ se alcanza el mínimo. $\square$

Observación

El error(cuadrático) que se comete al hacer esta predicción es la varianza

Ejemplo: Televisión

1) Hallar la varianza dela cantidad de paquetes que contrata un cliente de la empresa de televisión por cable, en la zona estudiada. Interpretar

$$\begin{aligned} V(X) &= E\left((X-E(X))^2\right) = E\left((X-2.225)^2\right) \\ &= \sum_{x \in R_X} (x-2.225)^2 p_X(x) \end{aligned}$$

Contestar por chat por mensaje privado.

2) Hacer el ejercicio llamado Varianza en el campus virtual

Fórmula útil para el cálculo de la varianza

$$V(X) = E\left(X^2\right) - (E(X))^2$$

Demostración Llamemos μ_X a $E(X)$

$$\begin{aligned} V(X) &= E\left((X-\mu_X)^2\right) = \sum_{x \in R_X} \left(x-\mu_X\right)^2 p_X(x) = \sum_{x \in R_X} \left(x^2 - 2\mu_X x + \mu_X^2\right) p_X(x) = \\ &= \sum_{X = R_X} x^2 p_X(x) - 2\mu_X \sum_{x \in R_X} x p_X(x) + \mu_X^2 \sum_{x = R_X} p_X(x) = E\left(X^2\right) - 2\mu_X E(X) + \mu_X^2 = \\ &= E\left(X^2\right) - 2\mu_X^2 + \mu_X^2 = E\left(X^2\right) - \mu_X^2 = E\left(X^2\right) - (E(X))^2 \end{aligned}$$

Definición de desvío estandar

$$DS(X) = \sqrt{V(X)}$$

La ventaja del desvío estandar es que está en las mismas unidades que la variable X.

Propiedades de la varianza y del desvío standard

1) $V(X) \geq 0$ y $DS(X) \geq 0$.

2) $V(aX+b) = a^2 V(X)$ y $DS(ax+b) = |a|\sigma_x$.

Demostración:

$$\begin{aligned} V(aX+b) &= \sum_{x \in R_x} (ax+b-E(aX+b))^2 p_X(x) = \sum_{x \in R_x} (ax+b-aE(X)-b)^2 p_X(x) = \\ &= \sum_{x \in R_x} (ax-aE(X))^2 p_x(x) = a^2 \sum_{x \in R_x} (x-E(X))^2 p_x(x) = a^2 V(X) \end{aligned}$$

3) Si X es una v.a. tal que $P(X=c)=1$, entonces $V(X) = DS(X) = 0$.

Demostración: Se tiene $R_X = \{x\}$, $E(X) = c$ y $E(X^2) = c^2$.

Por lo tanto, $V(X) = E\left(X^2\right) - (E(X))^2 = c^2 - c^2 = 0$