Variables aleatorias discretas.

Definición de variable aleatoria discreta

- Definición: Una v.a. se llama discreta si toma un número finito o infinito numerable de valores.
- Notación: Indicaremos con R_X el rango de la v.a. X, es decir el conjunto de valores posibles de la v.a. X.

Función de probabilidad puntual

- Definición: La función de probabilidad puntual o de masa de la v.a. discreta X, se define para todo x como

$$p_X(x) = P(X = x) = P(\{w \in \Omega / X(w) = x\})$$

- Propiedades

$$p_X(x) \geq 0 \quad \forall x$$
$$\sum_{x \in R_x} p_X(x) = 1$$

Ejemplo: Dados

Hallar la función de probabilidad puntual de la v.a.

X : número de resultados pares al arrojar dos veces un dado equilibrado.

$$R_X = \{0, 1, 2\}$$

- $p_X(0) = P(X = 0) = P\{\text{ambos resultados son impares}\} = 9/36$

- $p_X(1) = P(X = 1) =$
  $P\{$el primero es par y el segundo impar o el primero es impar y el segundo
  $P\{$el primero es par y el segundo impar$\} +$
  $P\{$el primero es impar y el segundo par$\} = 2 * 9/36$

- $p_X(2) = P(X = 2) = P\{\text{ambos resultados son pares}\} = 9/36$

Observación

Si X es una v.a. discreta con función de probabilidad puntual $p_X(x)$, su fda es

$$F_X(x) = \sum_{t \leq x, t \in R_x} p_X(t)$$

Ejemplo: Dado.

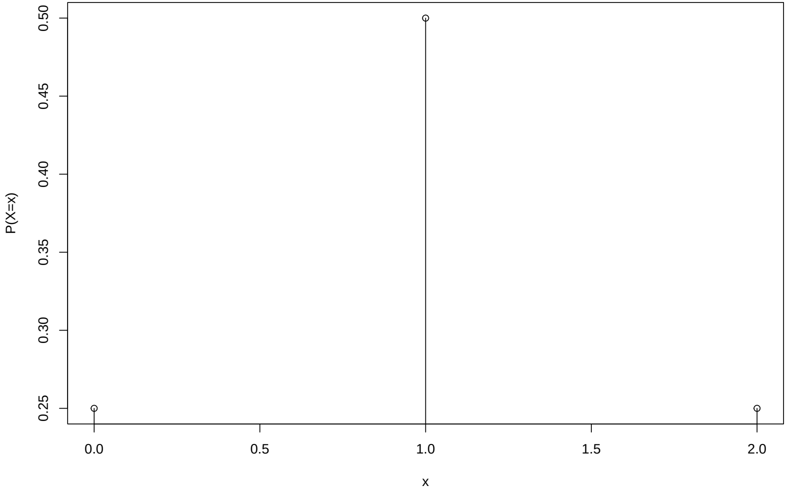
```r
rx <- 0:2
px <- c(9/36, 18/36, 9/36)
fpp <- cbind(rx, px)
colnames(fpp) <- c("x","P(X=x)")
fpp
```

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

Ejemplo: Dado. Gráfico de la función de probabilidad puntual

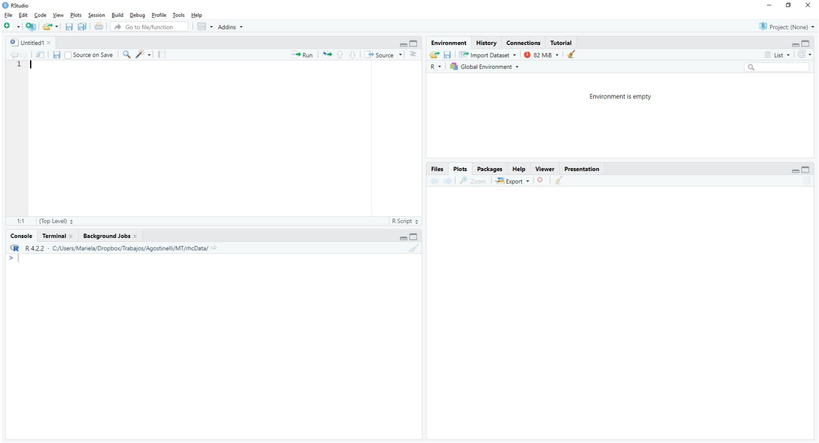
```
plot(rx, px)
lines(c(rx[1],rx[1]), c(0,px[1]))
lines(c(rx[2],rx[2]), c(0,px[2]))
lines(c(rx[3],rx[3]), c(0,px[3]))
```

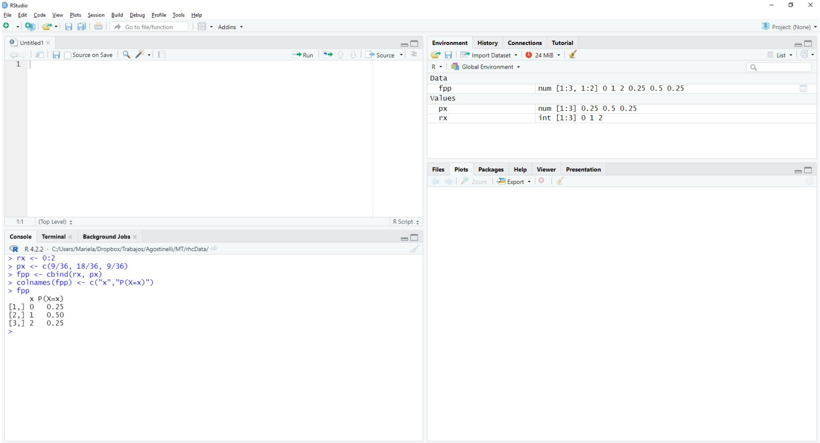
Ejemplo: Dado. Función de probabilidad puntual

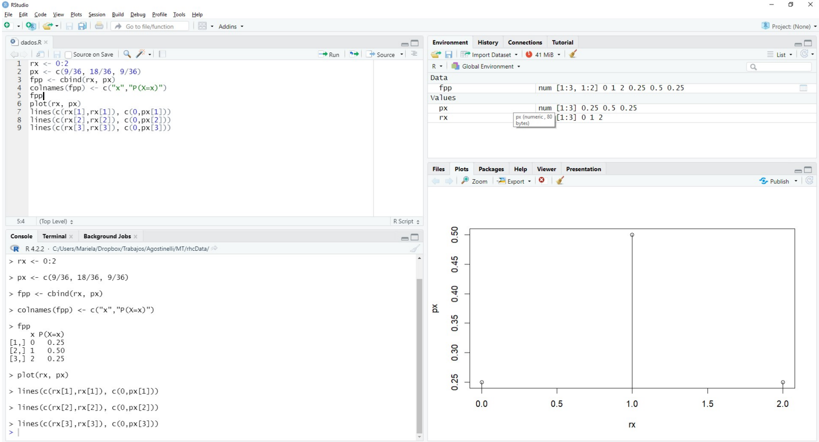


R

Bajar e instalar R y Rstudio




RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Untitled1
Source on Save
Run
Source
1:1
(Top Level)
R Script
Console
Terminal
Background Jobs
R 4.2.2 · C:/Users/Mariela/Dropbox/Trabajos/Agostinelli/MT/rhcData/
> 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
>
Environment
History
Connections
Tutorial
Import Dataset
24 MiB
List
Global Environment
Data
fpp
num [1:3, 1:2] 0 1 2 0.25 0.5 0.25
Values
px
num [1:3] 0.25 0.5 0.25
rx
int [1:3] 0 1 2
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
dados.R
Source on Save
Run
Source
rx <- 0:2
px <- c(9/36, 18/36, 9/36)
fpp <- cbind(rx, px)
colnames(fpp) <- c("x","P(X=x)")
fpp
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]))
5:4
(Top Level)
R Script
Console
Terminal
Background Jobs
R 4.2.2 · C:/Users/Mariela/Dropbox/Trabajos/Agostinelli/MT/rhcData/
> 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
> 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]))
>
Environment
History
Connections
Tutorial
Import Dataset
41 MiB
List
Global Environment
Data
fpp
num [1:3, 1:2] 0 1 2 0.25 0.5 0.25
Values
px
num [1:3] 0.25 0.5 0.25
rx
px (numeric, 80 bytes)
[1:3] 0 1 2
Files
Plots
Packages
Help
Viewer
Presentation
Zoom
Export
Publish
px
0.25
0.30
0.35
0.40
0.45
0.50
0.0
0.5
1.0
1.5
2.0
rx

Ejemplo: Dado

$X =$ cantidad de resultados pares

$$F_X(x) = \sum_{t \leq x, t \in R_x} p_X(t)$$

Hagamos una mini tabla de valores:

- Si $x = 0$, $F_X(x) = F_X(0) = p_X(0) = 9/36$
- Si $x = 1$,

$$F_X(x) = F_X(1) = p_X(0) + p_X(1) = 9/36 + 18/36 = 27/36$$

- Si $x = 2$,

$$F_X(x) = F_X(2) = p_X(0)+p_X(1)+p_X(2) = 9/36+18/36+9/36 = 1$$

Tomemos algunos valores intermedios:

- Si $x = -0.5$, $F_X(x) = F_X(-0.5) = 0$
- Si $x = 0.5$, $F_X(x) = F_X(0.5) = p_X(0) = 9/36$
- Si $x = 1.5$, $F_X(x) = F_X(1.5) = p_X(0) + p_X(1) = 27/36$
- Si $x = 2.5$, $F_X(x) = F_X(2.5) = p_X(0) + p_X(1) + p_X(2) = 1$

Ejemplo: Dado

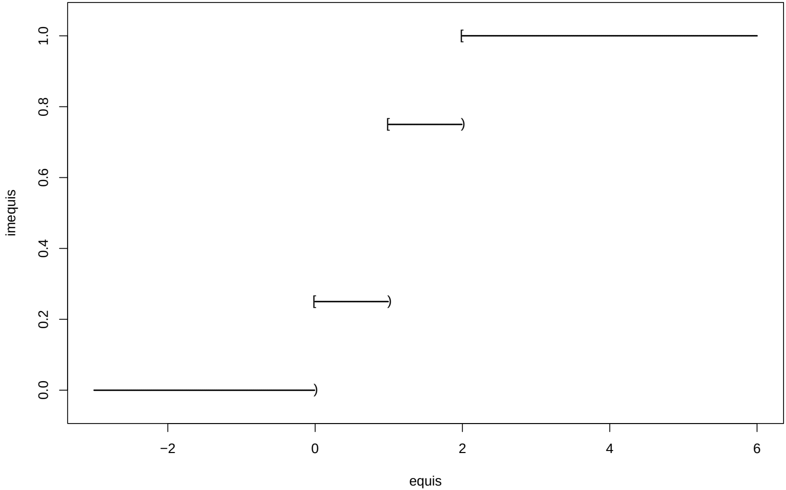
$X =$ cantidad de resultados pares.

$$F_X(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{9}{36} & \text{si } 0 \leq x < 1 \\ \frac{27}{36} & \text{si } 1 \leq x < 2 \\ 1 & \text{si } x \geq 2 \end{cases}$$

Ejemplo: Dado. Gráfico de F_X

```r
F_X <-  function(x){
  if(x<0) ans <- 0
  if(x>=0&x<1) ans <- 9/36
  if(x>=1 & x<2) ans<- 27/36
  if(x>=2) ans <- 1
  ans
  }
equis <- seq(-3,6, length = 1000)
imequis <- sapply(equis, F_X)
plot(equis, imequis, ylim = c(-0.05,1.05), cex=0.1)
text(-0.01,9/36, "[")
text(0.01,0/36, ")")
text(0.99,27/36, "[")
text(1.01,9/36, ")")
text(1.99,1, "[")
text(2.01,27/36, ")")
```

Ejemplo: Dados. Gráfico de F_X



Propiedad importante

Si X es una v.a. discreta con función de probabilidad puntual $p_X(x)$, en cada punto x, el valor del salto es la probabilidad puntual, es decir

$$p_X(x) = F_X(x) - F_X(x^-)$$

Dem:

$$p_X(x) = P(X = x) = P(X \leq x) - P(X < x) = F_X(x) - F_X\left(x^-\right)$$

Así, de la fda se puede deducir la fpp

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
proporición 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 2: Dados

Se tiran dos dados equilibrados

$X =$ cantidad de resultados pares

$E(X) = \sum_{x \in R_X} x p_X(x) = 0 \cdot p_X(0) + 1 \cdot p_X(1) + 2 \cdot p_X(2)$

$= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4}$

$= 1.$

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 2: Dado.

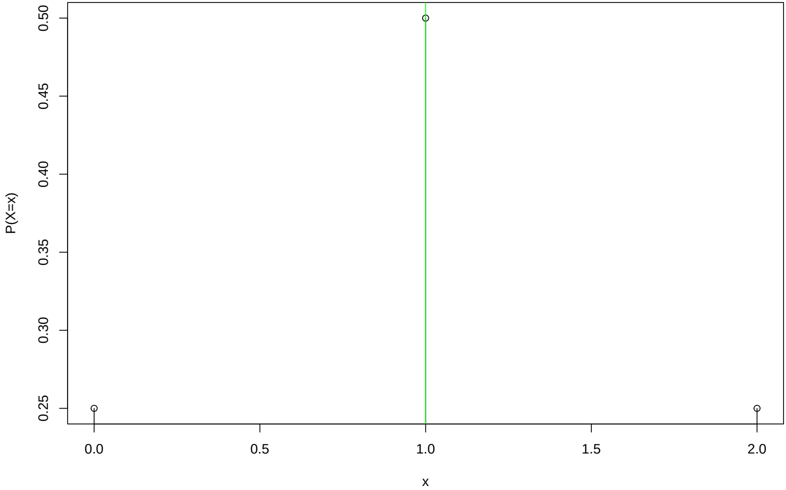
```r
rx <- 0:2
px <- c(9/36, 18/36, 9/36)
fpp <- cbind(rx, px)
colnames(fpp) <- c("x","P(X=x)")
fpp
```

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

Ejemplo 2: Dado.

```
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]))
```

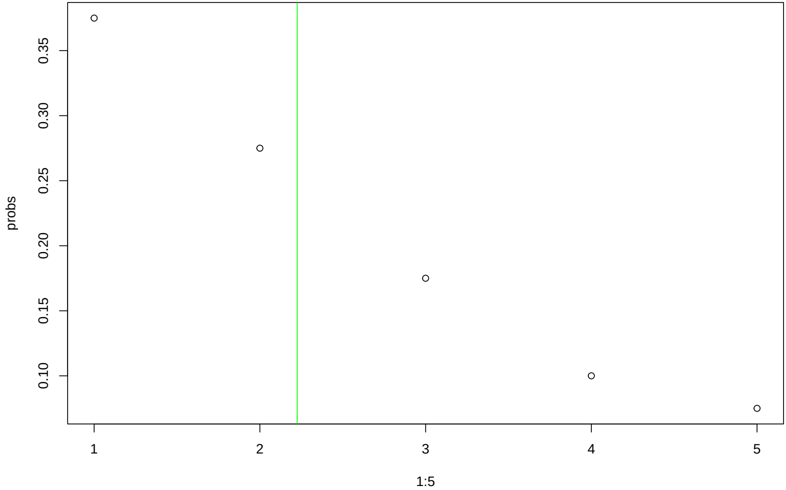
Ejemplo 2: Dado.



Ejemplo 1 : Televisión

```r
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 1 : Televisión



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 2: Dado

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

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

```r
resultados <- replicate(1000,
                        sample(1:6, 2, replace = TRUE))
cant_pares <- colSums(resultados%%2==0)
# %%2 es el resto de dividir por dos
mean(cant_pares)
```

```
## [1] 0.974
```

Ejemplo 1: 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 1: 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

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

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

Propiedades

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

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

Dem:

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

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 $E(Y)$

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

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

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

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?

$$Y - E(Y)$$

Esa diferencia 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 $f(t) = E\left((X-t)^2\right)$.

Demostración (haciendo algunos supuestos para simplificar)

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

$$f'(t) = -2E\left((X-t)\right) = -2(E\left(X\right) - t)$$

$$f'(t) = 0 \Leftrightarrow t = E(X)$$

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

Observación: Una medida del error(cuadrático) que se comete al hacer esta predicción es la varianza

Ejemplo 1: Televisión

1) Predecir la cantidad de paquetes que contrata un cliente de la empresa de televisión por cable, en la zona estudiada.
Predicción $E(X) = 2.225$

Error (cuadrático) esperado de la predicción:

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

```r
equis <- 1:5
probs <-  c(0.375, 0.275, 0.175,  0.1, 0.075)
sum((equis-2.225)^2*probs)
```

```
## [1] 1.574375
```

2) Predecir el ingreso que tendrá la empresa de televisión por cable por un nuevo cliente. ¿Cuál es el error esperado de esa predicción?

Predicción: $E(Y) = 96.75$

Error (cuadrático) esperado de la predicción: $V(Y)$

¿Cuánto da?

```
equis <- c(60 , 90 , 120 , 150 , 180)
probs <-  c(0.375, 0.275, 0.175,  0.1, 0.075)
esp <- sum(equis*probs)
sum((equis-esp)^2*probs)
```

```
## [1] 1416.938
```

Fórmula útil para el cálculo de la varianza

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

*Demostración*Llamemos μ_X a $E(X)$

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

Definición de desvío estandar

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

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