1. Una urna contiene n tarjetas numeradas del 1 al n. Un experimento aleatorio consiste en seleccionar al azar, una a una y con reemplazo, hasta que se obtiene una tarjeta que ya se seleccionó con anterioridad. Sea X la variable aleatoria que indica el número de tarjetas que son seleccionadas en este proceso. Calcula $P(X = k)$ para $k = 1, 2, 3, \ldots$

Sol.

Tenemos 2 casos con $K \in \mathbb{N}$, $K \leq n+1$ ó $K > n+1$. Si $K \leq n+1$, como extraemos K tarjetas siendo exactamente la K-ésima alguna de las anteriores, esto es, las $K-1$ tarjetas anteriores fueron todas diferentes, las cuales pueden salir de $\frac{n!}{(n-(K-1))!}$ formas diferentes, y la K-ésima puede ser elegida de $K-1$ formas, por tanto, en total serían $\frac{n!(K-1)}{(n-(K-1))!}$ formas en las que podemos obtener K cartas. Como el número total de casos en el que podemos sacar las K-cartas es n^K, entonces:

$$P(X=K) = \frac{\frac{n!(K-1)}{(n-(K-1))!}}{n^K} = \frac{n!(K-1)}{(n-(K-1))!\, n^K}$$

Si $K > n+1$, entonces $P(X=K)=0$, pues cuando $K > n+1$, en las primeras $n+1$ cartas, se debió de repetir alguna, por tanto, jamás se pueden extraer más de $n+1$ cartas. Por tanto $P(X=K)=0$.

$$\forall K \in \mathbb{N},\ P(X=K) := \begin{cases} \frac{n!(K-1)}{(n-(K-1))!\, n^K} & \text{si } K \leq n+1. \\ 0 & \text{si } K > n+1. \end{cases}$$

4. Sea X una variable aleatoria continua. Expresa $P(|X-a|>b)$ y $P(a \leq X < b)$ en términos de la función de distribución de X.

Sol.

La función de distribución de X está dada por:

$$F_X(x) = P(X \leqslant x)$$

Como X es v.a continua, entonces f_X es continua, así $P(X \leqslant x) = P(X < x)$. Por tanto:

$$\begin{aligned} P(a \leqslant X < b) &= P(a < X \leqslant b) \\ &= P(a < X \leqslant b) + P(X \leqslant a) - P(X \leqslant a) \\ &= P(X \leqslant b) - P(X \leqslant a) \\ &= F(b) - F(a) \end{aligned}$$

claro, en el caso que $a<b$. Si $b \leqslant a$, entonces $P(a \leqslant X < b) = 0$, pues X no puede ser $<b$, y a la vez ser $\geqslant a$, ya que $b \leqslant a$. Para $P(|X-a|>b)$, tenemos 2 casos:

a) $b \leqslant 0$. Si $b \leqslant 0$, entonces como $|X-a| \geqslant 0$ siempre, $P(|X-a|>b) = 1$.

b) $b > 0$. Como $|X-a|>b$, entonces $X-a>b$ ó $X-a<-b$, entonces $X > a+b$ ó $X<a-b$. Además $-b<b$, por lo cual $a-b<a+b$, entonces los eventos $[X>a+b]$, y $[X<a-b]$ son ajenos (ya que X no puede ser $>a+b$ y, a la vez $<a-b$). Por tanto:

$$\begin{aligned} P(|X-a|>b) &= P([X>a+b] \cup [X<a-b]) \\ &= P(X>a+b) + P(X<a-b) \\ &= 1 - P(X \leqslant a+b) + P(X \leqslant a-b) \\ &= 1 - F_X(a+b) + F_X(a-b) \end{aligned}$$

7. Sea X una variable aleatoria con valores en $x_1, x_2, x_3, \ldots$. Se repite n veces el experimento aleatorio asociado con X. Sea n_i el número de veces que ocurre el evento $[X = x_i]$; y sean $\overline{X}_n$ y S_n el promedio y la desviación estándar, respectivamente, de los valores observados en las n repeticiones del experimento.

La definición frecuencial de la probabilidad establece que $\frac{n_k}{n}$ tiende a estabilizarse alrededor de $P(X = k)$. Así para n suficientemente grande

$$\overline{X}_n = \sum_{i=1}^{\infty} x_i \cdot \frac{n_i}{n} \approx \sum_{i=1}^{\infty} x_i f_X(x_i) = E(X)$$
$$S_n^2 = \sum_{i=1}^{\infty} (x_i - \overline{X}_n)^2 \cdot \frac{n_i}{n} \approx \sum_{i=1}^{\infty} (x_i - E[X])^2 f_X(x_i) = Var(X)$$

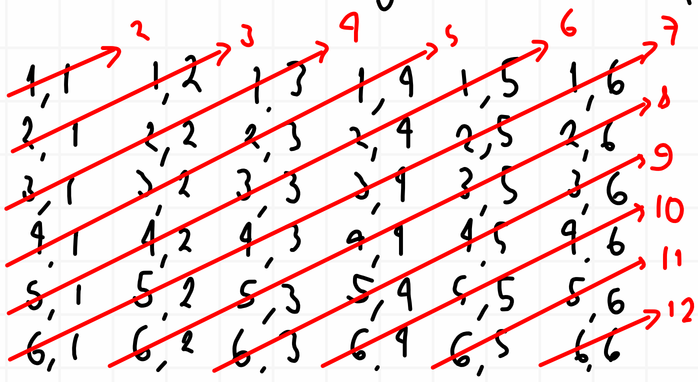
Realiza en R una simulación del lanzamiento de un par de dados, en la que X indique el valor de la suma de los puntos. Calcula, con lapiz y papel, la esperanza y la varianza de X y calcula en R los valores de X_n y de S_n; verifica si se cumple que $\overline{X}_n \underset{n\to\infty}{\longrightarrow} E(X)$ y $S_n^2 \underset{n\to\infty}{\longrightarrow} Var(X)$. Según los resultados obtenidos en tu simulación, comenta a partir de qué valor de n la aproximación puede considerarse como buena.

Sol.

La función de distribución de densidad, f_X está dada por:

$$f_X(x) = P(X = x), \ \forall \ x \in \mathbb{R}$$

Como X solo toma valores en $\{2, 3, \ldots, 12\}$, basta con calcular la probabilidad para estos valores. Del diagrama de la izquierda, vemos que:



$$P(X=2) = \frac{1}{36} = P(X=12),$$
$$P(X=3) = \frac{2}{36} = P(X=11),$$
$$P(X=4) = \frac{3}{36} = P(X=10),$$

$$P(X=5) = \frac{4}{36} = P(X=9), \quad P(X=6) = \frac{5}{36} = P(X=8), \quad \text{y}$$
$$P(X=7) = \frac{6}{36}$$

Por tanto, la esperanza será:

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = 7$$

$$\therefore E(X) = 7$$

Y, la varianza:

$$Var(X) = (2-7)^2 \cdot \frac{1}{36} + (3-7)^2 \cdot \frac{2}{36} + (4-7)^2 \cdot \frac{3}{36} + (5-7)^2 \cdot \frac{4}{36} + (6-7)^2 \cdot \frac{5}{36} + (7-7)^2 \cdot \frac{6}{36} +$$

$$(8-7)^2 \cdot \frac{5}{36} + (9-7)^2 \cdot \frac{4}{36} + (10-7)^2 \cdot \frac{3}{36} + (11-7)^2 \cdot \frac{2}{36} + (12-7)^2 \cdot \frac{1}{36}$$

$$= 5.83$$

$$\therefore Var(\bar{X}) = 5.83$$

Se realizó la simulación en R:

```r
n <- 1000000

dado_1 <- sample(1:6, n, replace = TRUE)
dado_2 <- sample(1:6, n, replace = TRUE)

suma <- dado_1 + dado_2

num <- c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)

for(i in 1:11){
  num[i] <- sum(suma == i+1)
}

x_prom <- 0
x_desv <- 0

for(i in 1:11){
  x_prom <- x_prom + ((i+1)*(num[i]/n))
}

for(i in 1:11){
  x_desv <- x_desv + (((i+1) - x_prom)^2)*(num[i]/n)
}

x_prom
x_desv
```

Donde se obtuvieron los sig. valores de $E(\bar{X})$ y $Var(\bar{X})$:

```
> x_prom
[1] 6.998004
> x_desv
[1] 5.826602
```

; $n = 1,000,000$

Vemos que, para $n = 10, 100, 1000, 10\,000$, se tienen:

```
> x_prom
[1] 6.2

> x_desv
[1] 3.76
```

$n = 10$

```
[1] 6.77

> x_desv
[1] 5.5771
```

$n = 100$

```
> x_prom
[1] 6.918

> x_desv
[1] 5.905276
```

$n = 1000$

```
> x_prom
[1] 7.00282

> x_desv
[1] 5.859092
```

$n = 10,000$

Por tanto, cuando $n \to \infty$, los valores se aproximan a $E(\bar{X})$, y $Var(\bar{X})$. Podemos también que, a partir de 10,000, tenemos una buena aproximación de los valores.