MAF: Tarea 4

Tomás Ricardo Basile Álvarez
316617194

30 de enero de 2021

1. En una ciudad se publican tres periódicos A, B, C. Una encuesta reciente muestra que el 20 % de los habitantes adultos de la ciudad lee el periódico A, el 16 % lee el periódico B, el 14 % lee el periódico C, el 8 % lee los periódicos A y B, el 5 % lee los periódicos A y C, el 4 % lee los periódicos B y C, y el 2 % lee los tres periódicos. Si se sabe que el total de habitantes en la ciudad es de 20,000 y se elige un adulto al azar,

a) **¿Cuántos habitantes no lee ninguno de los periódicos?**

Denotamos por A, B , C los eventos de que una persona lea los periódicos A,B,C respectivamente.
Por el problema, sabemos que:

$$P(A) = 0{,}2 \qquad P(B) = 0{,}16 \qquad P(C) = 0{,}14$$
$$P(A \cap B) = 0{,}08 \qquad P(A \cap C) = 0{,}05 \qquad P(B \cap C) = 0{,}04 \qquad P(A \cap B \cap C) = 0{,}02$$

Luego, en clase vimos la fórmula para calcular la probabilidad de una unión es:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \\ &= 0{,}2 + 0{,}16 + 0{,}14 - 0{,}08 - 0{,}05 - 0{,}04 + 0{,}02 = 0{,}35 \end{aligned}$$

Este $P(A \cup B \cup C) = 0{,}35$ es la probabilidad de que una persona lea el periódico A o el B o el C.
Es decir, están consideradas todas las personas que leen por lo menos un periódico.
Las personas que no leen ningún periódico son entonces el 65 % restante.
Es decir, $20000 * 0{,}65 = \boxed{\text{13000 personas.}}$

b) **¿Cuántos habitantes lee exactamente uno de los periódicos?**

Para ello, necesitamos calcular la probabilidad de que una persona lea solamente el periódico A pero no lea el B y no lea el C. Y luego hacer lo mismo para personas que

lean solamente el B y solamente el C.
Entonces, queremos calcular P(sólo leer el A) $= P$(leer el A y no leer ni el B ni leer el C) $=$ $P(A \cap B^c \cap C^c)$
Entonces calculamos esta probabilidad usando algunos de los teoremas vistos en clase:

$$\begin{aligned}
P(A \cap B^c \cap C^c) &= P(A \cap (B \cup C)^c) \quad \text{Por ley de De Morgan} \\
&= P(A/(B \cup C)) \quad \text{por cómo se define la diferencia de conjuntos} \\
&= P(A) - P(A \cap (B \cup C)) \quad \text{por cómo se calcula la prob. de una diferencia} \\
&= P(A) - P((A \cap B) \cup (A \cap C)) \\
&= P(A) - [P(A \cup B) + P(A \cup C) - P((A \cap B) \cap (A \cap C))] \\
&\qquad \text{por cómo se calcula la probabilidad de una unión} \\
&= P(A) - P(A \cup B) - P(A \cup C) + P(A \cap B \cap C) \quad \text{porque } (A \cap B) \cap (A \cap C) = A \cap B \cap C
\end{aligned}$$

Y esto ya lo podemos calcular porque son puras probabilidades que ya conocemos.
Es decir, encontramos que:

$$P(\text{sólo leer A}) = P(A) - P(A \cup B) - P(A \cup C) + P(A \cap B \cap C)$$

En ningún momento especificamos quienes eran los eventos A, B, C. Por lo que la fórmula se vale si cambiamos de nombre los conjuntos $A \to B, B \to C, C \to A$ y podemos obtener así una fórmula para la probabilidad de que una persona lea solamente B:

$$P(\text{sólo leer B}) = P(B \cap C^c \cap A^c) = P(B) - P(B \cup C) - P(B \cup A) + P(A \cap B \cap C)$$

Y podemos hacer de nuevo el mismo cambio para obtener la probabilidad de leer solamente C:

$$P(\text{sólo leer C}) = P(C \cap A^c \cap B^c) = P(C) - P(C \cup A) - P(C \cup B) + P(A \cap B \cap C)$$

Ahora sí, la probabilidad de que una persona lea solamente un periódico es la suma de la probabilidad de que lea sólo A más la probabilidad de que lea sólo B más la probabilidad de que lea sólo C:

$$\begin{aligned}
P(\text{leer sólo un periódico}) &= P(\text{sólo leer A}) + P(\text{sólo leer B}) + P(\text{sólo leer C}) \\
&= P(A) - P(A \cup B) - P(A \cup C) + P(A \cap B \cap C) \\
&\quad + P(B) - P(B \cup C) - P(B \cup A) + P(A \cap B \cap C) \\
&\quad + P(C) - P(C \cup A) - P(C \cup B) + P(A \cap B \cap C) \\
&= P(A) + P(B) + P(C) - 2P(A \cup C) - 2P(B \cup C) - 2P(A \cup C) + 3P(A \cap B \cap C)
\end{aligned}$$

Y ahora ya podemos sustituir los valores del problema:

$$\begin{aligned}
&P(A) + P(B) + P(C) - 2P(A \cup C) - 2P(B \cup C) - 2P(A \cup C) + 3P(A \cap B \cap C) \\
&\quad = 0{,}2 + 0{,}16 + 0{,}14 - 2(0{,}08) - 2(0{,}05) - 2(0{,}04) + 3(0{,}02) \\
&= 0{,}22
\end{aligned}$$

Como hay 20000 personas y la probabilidad de que una persona lea sólo un periódico es 0.22, entonces la cantidad de personas que leen un sólo periódico es:
$20000 * 0{,}22 = \boxed{4400 \text{ personas}}$

c) **Si A, B son periódicos que se publican por la mañana y C se publica por la tarde, ¿cuántos habitantes leen al menos un periódico de mañana y uno de tarde?**

Para que lea un periódico de la tarde, tiene que necesariamente leer el periódico C (que es el único de la tarde).
Y para que lea uno de la mañana tiene que leer ya sea el periódico A o el B (o ambos).

Es decir, el enunciado es equivalente a preguntar la cantidad de personas que leen el C y además leen el A o el B.
O bien, escrito en conjuntos, es el evento $C \cap (A \cup B)$.
Por lo que necesitaremos calcular la probabilidad de $P(C \cap (A \cup B))$.

Lo calculamos usando los teoremas vistos en clase:
En específico usaremos que para dos eventos R, Q se tiene que $P(R \cup Q) = P(R) + P(Q) - P(R \cap Q)$ y entonces $P(R \cap Q) = P(R) + P(Q) - P(R \cup Q)$

$$\begin{aligned}
&P(C \cap (A \cup B)) = \\
&\quad = P(C) + P(A \cup B) - P(C \cup (A \cup B)) \qquad \text{por la fórmula de arriba para } P(R \cap Q) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{con } \; R = C \; , \; Q = A \cup B \\
&\quad = P(C) + P(A \cup B) - P(A \cup B \cup C) \\
&\quad = P(C) + P(A) + P(B) - P(A \cap B) - P(A \cup B \cup C)
\end{aligned}$$

y ahora ya podemos sustituir los valores del problema y usar que en el primer inciso calculamos que $P(A \cup B \cup C) = 0{,}35$

$$\begin{aligned}
P(C \cap (A \cup B)) &= P(C) + P(A) + P(B) - P(A \cap B) - P(A \cup B \cup C) \\
&= 0{,}14 + 0{,}2 + 0{,}16 - 0{,}08 - 0{,}35 = 0{,}07
\end{aligned}$$

Por lo que el 0.07 % recibe el periódico a la mañana y a la tarde, lo que equivale a:
$0{,}07 * 20000 = \boxed{1400 \text{ personas}}$

2. Se consideran dos dados A, B. El dado A tiene 4 caras rojas y 2 blancas, mientras que el dado B tiene 2 rojas y 4 blancas. Se hace un volado de una moneda justa. Si sale sol se usa el dado A, mientras que si sale águila, se usa el dado B. Se repite sucesivamente el experimento
Asumiré que los dados son justos.

a) **Demostrar que la probabilidad de que salga rojo en cualquier tirada es** $1/2$

Denotamos los siguientes eventos:

- A = Se escoge el dado A (sale sol en la moneda)
- B = Se escoge el dado B (sale águila en la moneda)
- R_j = En la tirada j-ésima del dado sale j

Entonces, como indica el problema, se lanza una vez la moneda y con ello se decide cuál de los dados se va a utilizar para el resto de las tiradas.
La probabilidad de que se escoja cualquiera de los dados es la misma (porque la moneda es justa) y es igual a 1/2:

$$P(A) = P(B) = \frac{1}{2}$$

Si sucedió el evento A (se escogió el dado A), entonces todas las tiradas se hacen con este dado.
En este caso, la probabilidad de que la j-ésima tirada del dado salga rojo la denotaríamos por $P(R_j|A)$.
Y como el dado A tiene 4 de las 6 caras rojas, la probabilidad de que salga rojo en la jésima tirada es simplemente $\frac{4}{6} = \frac{2}{3}$ para todo j (porque si ya sabemos que tenemos el dado A, todas las tiradas del dado tienen probabilidad 2/3 de salir rojo). Por lo tanto:

$$P(R_j|A) = \frac{2}{3}$$

Por otro lado, si sucedió el evento B (la moneda escogió al dado B), entonces todas las tiradas se hacen con este dado. En este caso, la probabilidad de que la j-ésima tirada del dado salga rojo la denotaríamos como $P(R_j|B)$.
Y como el dado B tiene 2 de sus 6 caras rojas, la probabilidad de que salga rojo en la j-ésima tirada es $\frac{2}{6} = \frac{1}{3}$ para todo j. Entonces:

$$P(R_j|B) = \frac{1}{3}$$

Con esto ya podemos calcular la probabilidad de que en cualquier tirada del dado salga rojo (sin saber qué dado estamos usando), es decir, la probabilidad $P(R_j)$.

Para ello, usamos la ley probabilidad total. Sabemos que los eventos A, B son una partición del espacio muestral (pues son dos eventos disjuntos que llenan todo el espacio muestral porque la moneda tiene que salir cara o cruz). Entonces, podemos usar esta partición para escribir la ley de probabilidad total para el evento R_j:

$$\begin{aligned} P(R_j) &= P(R_j|A)P(A) + P(R_j|B)P(B) \\ &= \frac{2}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2} \\ &= \frac{1}{2} \end{aligned}$$

Con lo que ya probamos que la probabilidad de que salga rojo en cualquier tirada (sin tener información sobre tiradas anteriores o sobre qué dado se escogió) es de $\boxed{P(R_j) = \frac{1}{2}}$

b) **Si en las dos primeras tiradas salen rojos, ¿cuál es la probabilidad de que en la tercera tirada salga un rojo también?**

Con la notación que hemos estado usando, tenemos como hipótesis que las dos primeras tiradas salieron rojas (aunque desconocemos qué dado se está utilizando), es decir, sabemos que sucede el evento $R_1 \cap R_2$ (la tirada uno es roja y la tirada dos es roja). Luego, queremos calcular la probabilidad de que dada esta hipótesis, la tercera tirada sea roja, es decir, queremos calcular $P(R_3|R_1 \cap R_2)$

Para calcularlo, podemos usar primero que nada la defnción de la probabilidad condicional. Que dice que si tenemos dos eventos H, K entonces $P(H|K) = \dfrac{P(H \cap K)}{P(K)}$. En este caso, para $H = R_1$, $K = R_1 \cap R_2$, tenemos:

$$P(R_3|R_1 \cap R_2) = \frac{P(R_3 \cap (R_1 \cap R_2))}{P(R_1 \cap R_2)} = \frac{P(R_1 \cap R_2 \cap R_3)}{P(R_1 \cap R_2)}$$

Luego, podemos usar la ley de la probabilidad total usando la partición A, B como habíamos hecho antes. Y lo hacemos tanto para el numerador como para el denominador:

$$\begin{aligned} P(R_3|R_1 \cap R_2) &= \frac{P(R_1 \cap R_2 \cap R_3)}{P(R_1 \cap R_2)} \\ &= \frac{P(R_1 \cap R_2 \cap R_3|A)P(A) + P(R_1 \cap R_2 \cap R_3|B)P(B)}{P(R_1 \cap R_2|A)P(A) + P(R_1 \cap R_2|B)P(B)} \qquad (1) \end{aligned}$$

Entonces, necesitamos encontrar cada una de estas probabilidades.

Para ello, notamos que para todo i, j se cumple que $P(R_i \cap R_j|A) = P(R_i|A)P(R_j|A)$. Y lo mismo si cambiamos A por B.
Esto porque en estas probabilidades ya tenemos escogido el dado A. Y dos tiradas distintas del dado son independientes, pues que caiga en rojo no nos da información extra sobre qué dado podremos estar utilizando (pues ya sabemos cuál estamos usando).
Por esta independencia, la probabilidad de la intersección se separa en el producto de las probabilidades y tenemos ya que $P(R_i \cap R_j|A) = P(R_i|A)P(R_j|A)$.
Entonces podemos calcular las probabilidades que necesitamos para (1) usando algunas de las probabilidades que ya calculamos en el inciso anterior:

- $P(R_1 \cap R_2 \cap R_3|A) = P(R_1|A)P(R_2|A)P(R_3|A) = \dfrac{2}{3} \cdot \dfrac{2}{3} \cdot \dfrac{2}{3} = \dfrac{8}{27}$
- $P(R_1 \cap R_2 \cap R_3|B) = P(R_1|B)P(R_2|B)P(R_3|B) = \dfrac{1}{3} \cdot \dfrac{1}{3} \cdot \dfrac{1}{3} = \dfrac{1}{27}$
- $P(R_1 \cap R_2|A) = P(R_1|A)P(R_2|A) = \dfrac{2}{3} \cdot \dfrac{2}{3} = \dfrac{4}{9}$
- $P(R_1 \cap R_2|B) = P(R_1|B)P(R_2|B) = \dfrac{1}{3} \cdot \dfrac{1}{3} = \dfrac{1}{9}$

Por lo que al reemplazar en (1) tenemos que:

$$\begin{aligned}
P(R_3|R_1 \cap R_2) &= \frac{P(R_1 \cap R_2 \cap R_3|A)P(A) + P(R_1 \cap R_2 \cap R_3|B)P(B)}{P(R_1 \cap R_2|A)P(A) + P(R_1 \cap R_2|B)P(B)} \\
&= \frac{\frac{8}{27} \cdot \frac{1}{2} + \frac{1}{27} \cdot \frac{1}{2}}{\frac{4}{9} \cdot \frac{1}{2} + \frac{1}{9} \cdot \frac{1}{2}} \\
&= \frac{3}{5}
\end{aligned}$$

Por tanto, la probabilidad de que la tercera tirada sea roja dado que las primeras dos lo son es de $\boxed{P(R_3|R_1 \cap R_2) = \frac{3}{5}}$

Vemos que esto es mayor a la probabilidad $P(R_3)$ (la probabilidad de que la tercera tirada sea roja).
Esto se debe a que en $P(R_3|R_1 \cap R_2)$ ya sabemos que las primeras dos tiradas son rojas, lo que nos da un poco de información extra y básicamente indica que es más probable que estemos usando el dado A que tiene más caras rojas.

c) **Si en las dos primeras tiradas salen rojos, ¿cuál es la probabilidad de que el dado que se esté usando en las dos tiradas sea el dado rojo?**

Lo que queremos calcular es $P(A|R_1 \cap R_2)$ (probabilidad de que el dado escogido al principio sea A dado que los primeros dos tiros fueron rojos).
Usaremos el teorema de Bayes, que dice que si K, H son eventos entonces $P(K|H) = \dfrac{P(H|K)P(K)}{P(H)}$. Usamos este teorema para $K = A$ y $H = R_1 \cap R_2$ y tenemos:

$$P(A|R_1 \cap R_2) = \frac{P(R_1 \cap R_2|A)P(A)}{P(R_1 \cap R_2)}$$

Para calcular la probabilidad del denominador, usamos la ley de probabilidad total con la partición A, B.

$$P(A|R_1 \cap R_2) = \frac{P(R_1 \cap R_2|A)P(A)}{P(R_1 \cap R_2|A)P(A) + P(R_1 \cap R_2|B)P(B)}$$

Y usamos las probabilidades que ya calculamos en el inciso anterior:

$$\begin{aligned} P(A|R_1 \cap R_2) &= \frac{\frac{4}{9} \cdot \frac{1}{2}}{\frac{4}{9} \cdot \frac{1}{2} + \frac{1}{9} \cdot \frac{1}{2}} \\ &= \frac{4}{5} \end{aligned}$$

Por lo que la probabilidad de que dado que las primeras dos tiradas sean rojas, el dado usado sea el A es de $\dfrac{4}{5} = 0{,}8$

3. Una urna contiene 5 bolas blancas y 5 bolas negras. Dos bolas se sacan aleatoriamente (sin reemplazo). Si son iguales, ganamos \$1.10, pero si no son iguales perdemos \$1.00

a) **Calcular la ganancia media esperada. ¿Es favorable el juego para el jugador?**

Primero planteamos un espacio muestral muy sencillo $\Omega = \{$ ganar, perder$\}$ donde 'ganar' indica el evento de ganar el juego (sacar dos bolas iguales) y 'perder' es el evento complementario, no ganar el juego.

Primero que nada, vemos que al tomar las 2 bolas, primero se tiene que escoger una bola de entre las 10 de la urna y luego se escoge una bola de entre las 9 restantes, para un total de $10 \cdot 9 = 90$ posibilidades para elegir las dos bolas.

De estas 90 formas de elegir dos bolas, contamos ahora el número en que se pueden sacar cada uno de los pares de bolas posibles:

- **Sacar dos bolas blancas:** Para sacar la primera bola blanca, la urna empieza con 10 bolas (5 blancas y 5 negras), por lo que tenemos 5 formas de sacar una bola blanca.
  Una vez hecho esto, quedan 9 bolas en la urna (4 blancas y 5 negras) por lo que hay 4 formas de escoger la segunda bola.
  Por lo que tenemos un total de $5 \cdot 4 = 20$ maneras de tomar dos bolas blancas de la urna.
- **Sacar dos bolas negras:** Como la cantidad de bolas negras originalmente en la urna es igual al de bolas blancas, el problema de tomar dos bolas negras es equivalente al de las dos bolas blancas. Por lo que del inciso anterior concluimos que hay 20 formas de elegir dos bolas negras de la urna.
- **Sacar primero una bola blanca y luego una negra:** Para sacar la primera bola blanca, la urna empieza con 10 bolas de las cuales 5 son blancas. Por tanto, tenemos 5 formas de escoger la primera bola.
  Luego, en la urna quedan 4 bolas blancas y 5 bolas negras. Entonces, hay 5 maneras de sacar la segunda bola para que sea negra.
  Entonces tenemos un total de $5 \cdot 5 = 25$ formas de sacar una bola blanca y luego un negra.
- **Sacar primero una bola negra y luego una blanca:** Esta situación es equivalente a la del inciso anterior, porque al inicio del experimento hay la misma cantidad de bolas blancas y negras. Y por tanto en el argumento del inciso anterior podemos cambiar todas las palabras 'blanca' por 'negra' y viceversa y calcular así que el número de maneras de sacar una bola negra y luego una blanca es también 25.

Ahora sí calculamos las probabilidades.
Como dijimos al principio, hay 90 formas de sacar dos bolas de la urna y de ésas, 20

corresponden a sacar dos blancas, 20 a sacar dos negras y un total de 50 a sacar una y una en cualquier orden.
Las 90 formas de sacar las dos bolas de la urna son equiprobables.

Entonces podemos calcular las probabilidades simplemente dividiendo los casos 'favorables' entre los casos posibles:

- **Probabilidad de Ganar (sacar dos bolas iguales):** Tenemos 40 casos favorables de los 90 (20 de sacar 2 bolas blancas y 20 de 2 negras). Por lo que la probabilidad de ganar es:
$$P(ganar) = \frac{40}{90} = \frac{4}{9}$$
- **Probabilidad de Perder (sacar dos bolas distintas):** Tenemos 50 casos en los que perdemos de los 90. Por lo que la probabilidad de perder es:
$$P(perder) = \frac{50}{90} = \frac{5}{9}$$

Entonces, tenemos una variable aleatoria con una probabilidad de $\frac{4}{9}$ de ganar, o bien, de obtener \$1,10. Y una probabilidad de $\frac{5}{9}$ de perder, es decir, perder $-\$1{,}00$.

Entonces, como vimos en clase, la ganancia media esperada es la probabilidad de ganar multiplicada por el dinero que se gana (\$1,10) más la probabilidad de perder por el dinero que se pierde ($\$ - 1{,}00$).
Entonces la ganancia media es:

$$\begin{aligned}\text{Ganancia Media} &= P(ganar)(\$1{,}10) + P(perder)(\$ - 1{,}00) = \frac{4}{9}(\$1{,}10) + \frac{5}{9}(\$ - 1{,}00)\\ &= -\$\frac{1}{15} = \underline{-\$0{,}0666}\end{aligned}$$

Lo que quiere decir que en promedio el jugador pierde casi 7 centavos cada vez que juega.

b) **Calcular la variancia de la cantidad que se gana.**

Como dijimos antes, tenemos una variable aleatoria con una probabilidad de $\frac{4}{9}$ de ganar, o bien, de obtener \$1,10. Y una probabilidad de $\frac{5}{9}$ de perder, es decir, obtener $-\$1{,}00$.

Entonces, calculamos su variancia según la fórmula:

$$\begin{aligned} V(X) &= \sum_i (x_i - \mu)^2 p_i \qquad \text{con } \ \mu \text{ el valor medio calculado antes} \\ &= (\$1{,}10 - \mu)^2 \frac{4}{9} + (-\$1{,}00 - \mu)^2 \frac{5}{9} \\ &= (\$1{,}10 + \$\frac{1}{15})^2 \frac{4}{9} + (-\$1{,}00 + \$\frac{1}{15})^2 \frac{5}{9} \\ &= \$\frac{49}{45} = \underline{\$1{,}08888} \end{aligned}$$

c) **Si llamamos c a la cantidad que ganamos y d a la cantidad que perdemos en el juego, ¿qué relación entre c y d debe ocurrir para que el juego sea favorable para el jugador?**

Como vimos en el inciso a), tenemos una probabilidad de ganar de $P(ganar) = \frac{4}{9}$. Es decir, esta es la probabilidad de ganar una cantidad c.
Y tenemos una probabilidad $P(perder) = \frac{5}{9}$ de perder una cantidad d de dinero. Entonces, la media de estos datos es:

$$\text{Ganancia media} = P(ganar)c + P(perder)(-d) = \frac{4c}{9} - \frac{5d}{9}$$

Para que el juego sea favorable para el jugador, nos gustaría que el jugador tenga una ganancia media positiva, indicando que en promedio el jugador gana dinero en cada partida.
Por ello imponemos que la ganancia media sea mayor a 0:

$$\begin{aligned} &\frac{4c}{9} - \frac{5d}{9} > 0 \\ &\Rightarrow \quad 4c > 5d \\ &\Rightarrow \quad c > \frac{5}{4}d \end{aligned}$$

Y eso nos da la relación entre c y d para que el juego sea favorable para el jugador.
Es decir, la cantidad de dinero que se obtiene al ganar debe de ser por lo menos 1.25 veces mayor a la cantidad de dinero que se pierde al perder.

4. El número de minutos X que juega un jugador de básquetbol en un partido aleatorio es una variable aleatoria continua con función de densidad dada por

$$f(x) = \begin{cases} 0{,}025 & , \quad 10 \le x \le 20 \text{ ó } \; 30 \le x < 40 \\ 0{,}050 & , \quad 2- \le x < 30 \end{cases}$$

a) **Comprobar que en efecto es una función de densidad y graficarla**

Para comprobar que efectivamente es una función de densidad, debemos de comprobar dos cosas:

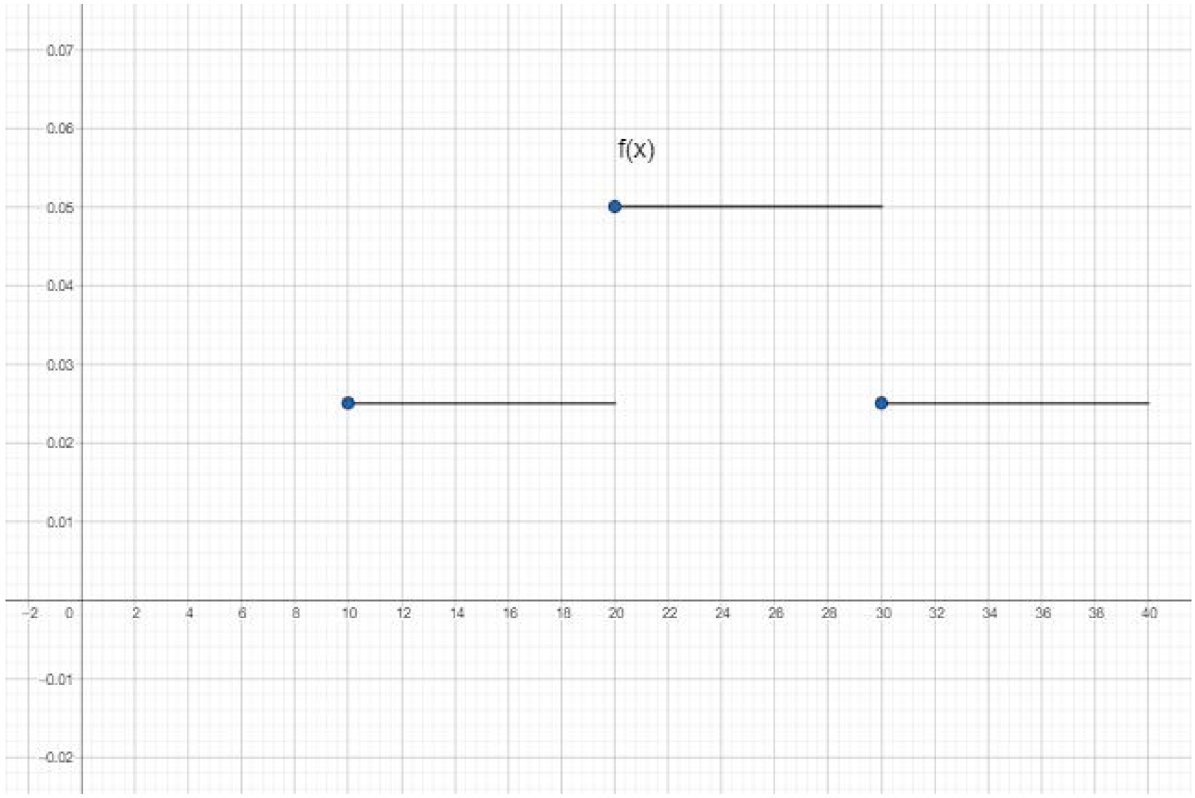
- **f es no negativa:** $(f(x) \ge 0 \quad \forall x)$
  Directamente de la definición vemos que f toma siempre valores positivos.
  Y podemos extender f a todos los reales definiendo que $f(x)$ valga 0 en todos los puntos no considerados (es decir $x \in (-\infty, 10) \cup (40, \infty)$).
  Que la función valga 0 en estos puntos no cambia la interpretación de la función pero nos permite tenerla definida para todos los reales.
- **La probabilidad total de la recta real es 1** $(\int_{\mathbb{R}} f(x)dx = 1)$:

$$\int_{\mathbb{R}} f(x)dx = \int_{10}^{20} 0{,}025dx + \int_{20}^{30} 0{,}05dx + \int_{30}^{40} 0{,}025dx$$

  Donde usamos la definición de f en cada uno de los intervalos en los que no es 0.

$$\begin{aligned} &= 0{,}025 \int_{10}^{20} dx + 0{,}05 \int_{20}^{30} dx + 0{,}025 \int_{30}^{40} dx \\ &= 0{,}025(20-10) + 0{,}025(30-20) + 0{,}025(40-30) \\ &= 10(0{,}025 + 0{,}025 + 0{,}05) = 10(0{,}1) = 1 \end{aligned}$$

  Lo que muestra que $\int_{\mathbb{R}} f(x)dx = 1$



Gráfica de la función.

b) **¿Cuál es la probabilidad de que el jugador juegue más de 15 minutos? ¿Y entre 20 y 35 minutos? ¿Y menos de 30 minutos? ¿Y más de 36 minutos?**

Simplemente usamos la definición de una función de densidad, que sabemos que la probabilidad de que la variable aleatoria X tome un valor entre a y b es de $\int_a^b f(x)dx$. Y para calcular esta integral usamos la definición de $f(x)$ en cada intervalo.
Entonces:

- **Probabilidad de que el jugador juegue más de 15 minutos:** Es la probabilidad de que X tome un valor mayor a 15:

$$\begin{aligned} P(X \geq 15) = \int_{15}^{\infty} f(x)dx &= \int_{15}^{20} 0{,}025dx + \int_{20}^{30} 0{,}05dx + \int_{30}^{40} 0{,}025dx \\ &= 0{,}025 \int_{15}^{20} dx + 0{,}05 \int_{20}^{30} dx + 0{,}025 \int_{30}^{40} dx \\ &= 0{,}025(5) + 0{,}05(10) + 0{,}025(10) = \underline{0{,}875} \end{aligned}$$

- **Probabilidad de que el jugador juegue entre 20 y 35 minutos:** Es la probabilidad de que X tome un valor entre 20 y 35:

$$\begin{aligned}
P(20 \leq X \leq 35) &= \int_{20}^{35} f(x)dx = \int_{20}^{30} 0{,}05dx + \int_{30}^{35} 0{,}025dx \\
&= 0{,}05 \int_{20}^{30} dx + 0{,}025 \int_{30}^{35} dx \\
&= 0{,}05(10) + 0{,}025(5) = \underline{0{,}625}
\end{aligned}$$

- **Probabilidad de que el jugador juegue menos de 30 minutos:** Es la probabilidad de que X tome un valor menor a 30:

$$\begin{aligned}
P(X \leq 30) &= \int_{-\infty}^{30} f(x)dx = \int_{10}^{20} 0{,}025dx + \int_{20}^{30} 0{,}05dx \\
&= 0{,}025 \int_{10}^{20} dx + 0{,}05 \int_{20}^{30} dx \\
&= 0{,}025(10) + 0{,}05(10) = \underline{0{,}75}
\end{aligned}$$

- **Probabilidad de que el jugador juegue más de 36 minutos:** Es la probabilidad de que X tome un valor mayor a 36:

$$\begin{aligned}
P(X \geq 36) &= \int_{36}^{\infty} f(x)dx = \int_{36}^{40} 0{,}025dx + \\
&= 0{,}025 \int_{36}^{40} dx \\
&= 0{,}025(4) = \underline{0{,}1}
\end{aligned}$$

c) **Calcular** $E(X)$ **y** $Var(X)$
Según la definición para calcular $E(X)$ tenemos:

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xf(x)dx = \int_{10}^{20} 0{,}025xdx + \int_{20}^{30} 0{,}05xdx + \int_{30}^{40} 0{,}025xdx \\
&= 0{,}025 \int_{10}^{20} xdx + 0{,}05 \int_{20}^{30} xdx + 0{,}025 \int_{30}^{40} xdx \\
&= 0{,}025 \frac{x^2}{2}\bigg|_{10}^{20} + 0{,}05 \frac{x^2}{2}\bigg|_{20}^{30} + 0{,}025 \frac{x^2}{2}\bigg|_{30}^{40} \\
&= 0{,}025 \frac{20^2 - 10^2}{2} + 0{,}05 \frac{30^2 - 20^2}{2} + 0{,}025 \frac{40^2 - 30^2}{2} \\
&= \boxed{25}
\end{aligned}$$

Para calcular $Var(X)$, calculamos primero el segundo momento central de X:

$$\begin{aligned}
E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx = \int_{10}^{20} 0{,}025x^2dx + \int_{20}^{30} 0{,}05x^2dx + \int_{30}^{40} 0{,}025x^2dx \\
&= 0{,}025 \int_{10}^{20} x^2dx + 0{,}05 \int_{20}^{30} x^2dx + 0{,}025 \int_{30}^{40} x^2dx \\
&= 0{,}025 \frac{x^3}{3}\Big|_{10}^{20} + 0{,}05 \frac{x^3}{3}\Big|_{20}^{30} + 0{,}025 \frac{x^3}{3}\Big|_{30}^{40} \\
&= 0{,}025 \frac{20^3 - 10^3}{3} + 0{,}05 \frac{30^3 - 20^3}{3} + 0{,}025 \frac{40^3 - 30^3}{3} \\
&= \frac{2050}{3}
\end{aligned}$$

Luego, usamos la fórmula de la varianza vista en clase:

$$\begin{aligned}
Var(X) &= E(X^2) - E(X)^2 \\
&= \frac{2050}{3} - 25^2 \\
&= \boxed{\frac{175}{3}} = 58{,}3333\cdots
\end{aligned}$$

5. a) Sea X una variable aleatoria normal de parámetros μ, σ^2. Demostrar que la variable aleatoria $Y = aX + b$ con $a > 0, b \in \mathbb{R}$, es una variable aleatoria normal de parámetros $a\mu + b$, $a^2\sigma^2$

Tenemos que $Y = g(X)$ donde $g(x) = ax + b$.
Luego, una fórmula clásica de probabilidad es que si $Y = g(X)$ con $g(x)$ una función estrictamente creciente y continuamente diferenciable (por lo que es invertible). Entonces $Y = g(X)$ tiene como función de densidad de probabilidad a:

$$f_Y(y) = f_X(g^{-1}(y)) \cdot (g^{-1})'(y)$$

con f_Y la función de densidad de Y y f_X la de X.
Para comprobar está fórmula, hay que demostrar que f_Y es la fórmula de densidad de Y. Es decir, demostrar que $P(c \leq Y \leq d) = \int_c^d f_Y(y)dy$.
Lo demostramos ahora:

$$\begin{aligned} P(c \leq Y \leq d) &= P(c \leq g(X) \leq d) \\ &= P(g^{-1}(c) \leq X \leq g^{-1}(d)) \qquad \text{porque g es creciente} \\ &= \int_{g^{-1}(c)}^{g^{-1}(d)} f_X(x)dx \\ &= \int_c^d f_X(g^{-1}(y)) \cdot (g^{-1})'(y)dy \qquad \text{hacemos el cambio de variable de la integral con y = g(x)} \\ &\qquad\qquad\qquad \text{entonces } x = g^{-1}(y) \ \Rightarrow \ dx = (g^{-1})'(y) \end{aligned}$$

Por lo que demostramos que $P(c \leq Y \leq d) = \int_c^d f_X(g^{-1}(y)) \cdot (g^{-1})'(y)dy$ y entonces por la definición de función de densidad,
$f_Y(y) = f_X(g^{-1}(y)) \cdot (g^{-1})'(y)$.

Armados con esta fórmula, teníamos que $Y = g(X)$ con $g(x) = ax + b$, $a > 0$.
Por lo que $g(x)$ claramente es creciente y diferenciable. Entonces podemos aplicar la fórmula de antes para encontrar f_Y en función de f_X.

Y como X se distribuye normal con parámetros μ, σ^2, su función de densidad es $f_X(x) = \dfrac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$. Entonces la función $f_Y(y)$ es:

$$
\begin{aligned}
f_Y(y) &= f_X(g^{-1}(y)) \cdot (g^{-1})'(y) \\
&= \frac{1}{\sigma\sqrt{2\pi}} e^{-(g^{-1}(y)-\mu)^2/(2\sigma^2)} \cdot (g^{-1})'(y) \\
&= \frac{1}{\sigma\sqrt{2\pi}} e^{-(\frac{y-b}{a}-\mu)^2/(2\sigma^2)} \cdot \frac{1}{a} \qquad \text{porque } y = g(x) = ax+b \\
& \qquad \Rightarrow\ g^{-1}(y) = \frac{y-b}{a} \ \Rightarrow\ (g^{-1})'(y) = \frac{1}{a} \\
&= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\left(\frac{y-b-\mu a}{a}\right)^2} \\
&= \frac{1}{a\sigma\sqrt{2\pi}} e^{-\frac{1}{2(a\sigma)^2}(y-(b+\mu a))^2}
\end{aligned}
$$

Vemos que esta función de densidad es la de una variable aleatoria normal con media $(b+\mu a)$ y con desviación estándar de $a\sigma$ (y por tanto varianza de $a^2\sigma^2$).
Lo que demuestra el ejercicio.

b) Sea X una variable aleatoria Poisson con parámetro λ. Calcular la probabilidad de que X tome sólo valores pares, i.e $P(X$ es par).

La función de probabilidad de $X \sim$ Poisson(λ) es $f(x) = e^{-\lambda}\dfrac{\lambda^n}{n!}$, $\quad n = 0, 1, 2, \cdots$.
Entonces, la probabilidad de que X tome el valor n es de $e^{-\lambda}\dfrac{\lambda^n}{n!}$.
Luego, para calcular la probabilidad de que X tome algún valor par, hay que sumar $P(X = n)$ para todo n par:

$$
\begin{aligned}
P(X\text{par}) &= \sum_{n\ par} P(X=n) = \sum_{n\ par} e^{-\lambda}\frac{\lambda^n}{n!} \\
&= \sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^{2n}}{(2n)!} \qquad \text{reemplazamos n por 2n para sumar sobre sólo los pares} \\
&= e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^{2n}}{(2n)!}
\end{aligned}
$$

Ahora reconocemos que $\sum_{n=0}^{\infty}\dfrac{\lambda^{2n}}{(2n)!} = 1 + \dfrac{\lambda^2}{2!} + \dfrac{x^4}{4!} + \cdots$ es la serie de Taylor de $\cosh(\lambda)$. Por tanto:

$$
\begin{aligned}
P(X\ par) &= e^{-\lambda}\cosh(\lambda) \\
&= e^{-\lambda}\frac{e^{\lambda}+e^{-\lambda}}{2} \\
&= \boxed{\frac{1+e^{-2\lambda}}{2}}
\end{aligned}
$$

6. La cantidad de lluvia que cae en Ciudad de México anualmente es una variable aleatoria normal de media 840 milímetros y desviación típica 150 milímetros.

a) **¿Cuál es la probabilidad de que en 2021 llueva más de 900 milímetros**

Sea X la variable aleatoria que da como resultado la cantidad de lluvia en la Ciudad de México en un año.
Entonces, la función de densidad es una función normal con parámetros $\mu = 840mm$, $\sigma = 150mm$, es decir:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{150\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-840}{150}\right)^2}$$

Queremos ver la probabilidad de que llueva más de 900 milímteros, es decir $P(X \geq 900)$, que se calcula como:

$$P(X \geq 900) = \int_{900}^{\infty} f(x)dx \ = \int_{900}^{\infty} \frac{1}{150\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-840}{150}\right)^2} dx$$

Podemos hacer un cambio de variable para convertir esta distribución normal a una distribución normal estándar ($\mu = 0, \sigma = 1$).
Para ello, hacemos el cambio de variable $y = \dfrac{x-840}{150}$, $dy = \dfrac{dx}{150}$
Y los límites ahora son $x = 900 \ \Rightarrow \ y = \dfrac{900-840}{150} = 0{,}4$, $x = \infty \ \Rightarrow \ y = \infty$.
Por tanto, la integral queda como:

$$\begin{aligned} P(X \geq 900) &= \int_{0,4}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}y^2} dy \\ &= \Phi(\infty) - \Phi(0{,}4) \end{aligned}$$

Donde Φ es la función de distribución acumulada correspondiente a la distribución normal estándar. Es decir, $\phi(x) = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy$.
Esta integral no se puede calcular analíticamente, pero se puede hacer en una computadora o usar una de las muchas tablas que existen con valores de esta función.
De esa forma, tenemos que:

$$P(X \geq 900) = \Phi(\infty) - \Phi(0{,}4) = 1 - 0{,}6552 = 0{,}3348$$

Por lo que la probabilidad de que se tenga más de 900 mm de lluvia es $\underline{0{,}3448}$

b) **¿Cuál es la probabilidad de que en los próximos 7 años haya exactamente 3 años donde se superen los 900 milímetros**

Como vimos en el inciso pasado, la probabilidad de que llueva más de 900 milímetros en un año es de 0.3448. Que a partir de ahora denotaremos como p.
Diremos que un año es 'Existoso' y lo marcamos con E si en dicho año llueve más de 900 milímetros. Y dirdemos que es un 'fracaso' y lo denotamos por F si llueve menos de 900 milímetros.

Entonces definimos un espacio muestral Ω que contenga los resultados de éxito o fracaso de los próximos años. Es decir, un elemento de Ω podría ser de la forma $EEFFFE$ que indica que llovió más de 900 milímetros en los primeros dos años y en el último pero llovió menos de 900 milímetros en los demás.

Así, podemos ver que lo que pide el problema es encontrar la probabilidad de que los próximos 7 años tengan exactamente 3 'éxitos' E y 4 'fracasos' F en cualquier orden.
Si suponemos que la cantidad de lluvia en un año es independiente de la cantidad de lluvia en los otros, entonces el espacio muestral denotado antes consiste en realizar 7 ensayos independientes de Bernoulli con probabilidad p de éxito y contar el número de éxitos totales, es decir, es una distribución binomial.

Entonces, para calcular la probabilidad de que haya 3 éxitos en los próximos 7 años simplemente necesitamos calcular:

$$\binom{7}{3} p^3(1-p)^4$$

Esto porque si los eventos son independientes, p^3 mide la probabilidad de tener 3 éxitos y $(1-p)^4$ la probabilidad de tener 4 fracasos (porque cada fracaso tiene una probabilidad $1-p$ y los eventos son independientes).
Sin embargo, hay muchas formas en las que se pueden reordenar estos 3 éxitos y 4 fracasos. De entre los 7 años, hay $\binom{7}{3}$ maneras de elegir en qué años van a quedar los éxitos (y en los demás años quedarán los fracasos).

Por tanto, la probabilidad que se nos pide de tener exactamente 3 años lluviosos en los próximos 7 es de:

$$\binom{7}{3} p^3(1-p)^4 = 35p^3(1-p)^4$$

Donde p es la probabilidad calculada en el inciso pasado, $p = 1 - \Phi(0{,}4) = 0{,}3448$. Si la sustituimos nos queda una probabilidad de

$$35(0{,}3348)^3(1-0{,}3348)^4 = 0{,}257$$

7. Calcular el valor esperado $E(X)$ y la varianza $Var(X)$ de las siguientes variables aleatorias. Además, para cada una de ellas, mostrar una aplicación real en la cual se usen dichas variables aleatorias.

a) **X es una variable aleatoria geométrica con distribución de probabilidad:**

$$P(X = n) = (1-p)^{n-1}p \quad , \quad n = 1, 2, \cdots , \quad 0 < p < 1$$

Por definición, tenemos que el valor esperado de X es: $E(X) = \mu = \sum_i x_i P(x_i)$
Donde x_i se calcula sobre todas las posibles imágenes de X. En este caso hay que sumar sobre $x_i = 1, 2, \cdots$
Entonces tenemos que:

$$\begin{aligned} E(X) &= \sum_{i=1}^{\infty} x_i P(x_i) = \sum_{n=1}^{\infty} n * P(X = n) \\ &= \sum_{n=1}^{\infty} n(1-p)^{n-1}p = p\sum_{n=1}^{\infty} n(1-p)^{n-1} \end{aligned}$$

La serie $\sum_{n=1}^{\infty} n(1-p)^{n-1}$ son un ejemplo de una **serie aritmético-geométrica**, que en general es una serie de la forma $\sum\limits_{n=1}^{\infty} nx^{n-1}$.
Resolveremos esta serie y luego lo aplicaremos para el caso particular que nos interesa aquí.

Para ver cuanto vale esta serie aritmético-geométrica, encontramos primero el valor de una suma conseguida al cortar esta serie en N términos (La N-ésima suma parcial). Es decir, buscamos calcular $\sum\limits_{n=1}^{N} nx^{n-1}$. Al valor que toma esta suma le llamaremos $S_N = \sum\limits_{n=1}^{N} nx^{n-1}$ y realizaremos un truco muy común para sumas de este estilo:
Primero consideramos la suma S_N:

$$S_N = \sum_{n=1}^{N} nx^{n-1} = 1 + 2x + 3x^2 + 4x^3 + \cdots + Nx^{N-1}$$

Ahora multiplicamos por x ambos lados:

$$xS_N = x + 2x^2 + 3x^3 + 4x^4 + \cdots + (N-1)x^{N-1} + Nx^N$$

Y restamos la primera suma menos la segunda agrupando los términos con la misma potencia de x:

$$\begin{aligned} S_N - xS_N &= 1 + x + x^2 + x^3 + \cdots + x^{N-1} - Nx^N \\ &= \frac{1 - x^{(N-1)+1}}{1-x} - Nx^N \\ &= \frac{1-x^N}{1-x} - Nx^N \qquad (1) \end{aligned}$$

Porque los primeros términos $1 + x + x^2 + \cdots + x^{N-1}$ una suma geométrica que llega hasta la potencia $N - 1$ y la suma tiene el resultado conocido de $\dfrac{1 - x^{(N-1)+1}}{1 - x}$

Despejando S_N de (1) tenemos que:

$$S_N = \frac{1 - x^N}{(1-x)^2} - \frac{Nx^N}{1-x}$$

Con lo que hemos encontrado el valor de la N-ésima suma parcial de $\sum_{n=1}^{\infty} nx^{n-1}$. Para encontrar esta serie infinita, simplemente tomamos el límite conforme $N \to \infty$.
Y vamos a centrarnos en el caso en que $\underline{|x| < 1}$, lo que nos va a asegurar la convergencia.
Si $|x| < 1$, $\Rightarrow$

$$\begin{aligned}
\sum_{n=1}^{\infty} nx^{n-1} &= \lim_{N\to\infty} \sum_{n=1}^{N} nx^{n-1} = \lim_{N\to\infty} S_n \\
&= \lim_{N\to\infty} \frac{1 - x^N}{(1-x)^2} - \frac{Nx^N}{1-x} \\
&= \frac{1}{(1-x)^2}
\end{aligned}$$

* Como $|x| < 1$, entonces conforme $N \to \infty$, $x^N \to 0$ e incluso $Nx^N \to 0$ por lo que podemos desechar esos términos.
Con lo que ya tenemos el resultado de la serie aritmético-geométrica:

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2} \quad , \quad |x| < 1$$

Si ahora retomamos el caso especial de la **esperanza** que estábamos estudiando, teníamos que:

$$\begin{aligned}
E(X) &= p \sum_{n=1}^{\infty} n(1-p)^{n-1} \\
&= p \frac{1}{(1-(1-p))^2} \\
&= \frac{p}{p^2} = \frac{1}{p}
\end{aligned}$$

Porque $\sum_{n=1}^{\infty} n(1-p)^{n-1}$ es el caso especial de la suma aritmético-geométrica estudiada antes cuando $x = 1 - p$ (y converge porque cumple que $|x| = |1-p| < 1$ porque $0 < p < 1$).

$$\boxed{E(X) = \frac{1}{p}}$$

Varianza: Como vimos en clase, la varianza se puede calcular como $E(X^2) - E(X)^2$.
Donde $E(X^2) = \sum_i x_i^2 p_i$ con x_i todas las posibles imágenes de X, que en este caso son $x_i = 1, 2, \cdots$ y p_i sus probabilidades
Por lo que tenemos que:

$$\begin{aligned} E(X^2) &= \sum_{i=1}^{\infty} x_i^2 p_i = \sum_{n=1}^{\infty} n^2 P(X = n) \\ &= \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} p = p \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} \end{aligned}$$

Para resolver esto, nos interesa ser capaces de resolver series de la forma $\sum_{n=1}^{\infty} n^2 x^{n-1}$.
Primero notamos que podemos hacer un cambio de variable $n \to n+1$ para escribir la serie equivalentemente como $\sum_{n=0}^{\infty} (n+1)^2 x^n$
Y ahora es una serie de potencias como las que se estudian en cálculo, con coeficientes $c_n = (n+1)^2$.
En cálculo se demuestra que estas series tienen un radio de convergencia dado por el ratio test como $R = \lim_{n\to\infty} \left|\frac{c_n}{c_{n+1}}\right| = \lim_{n\to\infty} \frac{(n+1)^2}{(n+2)^2} = 1$.
Por lo que nuestra serie $\sum_{n=0}^{\infty} (n+1)^2 x^n$ o equivalentemente $\sum_{n=1}^{\infty} n^2 x^{n-1}$ converge en $|x| < 1$.
Y no sólo eso, sabemos del cálculo que no sólo converge, sino que converge uniformemente en dicha bola $|x| < 1$.
Por lo que podemos integrar la serie y luego intercambiar la integral y la suma sin problemas (por convergencia uniforme), hacemos eso:

$$\begin{aligned} \int \sum_{n=0}^{\infty} (n+1)^2 x^n dx &= \sum_{n=0}^{\infty} \int (n+1)^2 x^n dx \\ &= \sum_{n=0}^{\infty} (n+1) x^{n+1} \end{aligned}$$

Queremos calcular esta serie $S = \sum_{n=0}^{\infty} (n+1) x^{n+1}$ que es la integral de la serie que nos interesa.
Para ello, expandimos esta serie:

$$S = \sum_{n=0}^{\infty} (n+1) x^{n+1} = x + 2x^2 + 3x^3 + 4x^4 + \cdots$$

Luego multiplicamos ambos lados por x:

$$xS = \sum_{n=0}^{\infty} (n+1) x^{n+2} = x^2 + 2x^3 + 3x^4 + 4x^5 + \cdots$$

Y entonce tenemos que:

$$\begin{aligned}
S - xS &= x + x^2 + x^3 + x^4 + \cdots \\
&= -1 + 1 + x + x^2 + x^3 + x^4 + \cdots \qquad \text{sumamos y restamos 1} \\
&= -1 + \frac{1}{1-x} \quad , \quad |x| < 1 \qquad \text{porque nos queda una serie geométrica} \\
&= \frac{x}{1-x}
\end{aligned}$$

Entonces, despeando S tenemos que:

$$S = \sum_{n=0}^{\infty} (n+1)x^{n+1} = \frac{x}{1-x}\frac{1}{1-x} = \frac{x}{(1-x)^2} \qquad , \ |x| < 1$$

Pero como habíamos notado antes, la suma que nos interesa es $\sum_{n=0}^{\infty}(n+1)^2 x^n$, que como habíamos notado, al integrarla nos daba la suma $S = \sum_{n=0}^{\infty}(n+1)x^{n+1}$.

Por lo que la suma que nos interesa $\sum_{n=0}^{\infty}(n+1)^2 x^n$ es la derivada de S:

$$\begin{aligned}
\sum_{n=0}^{\infty}(n+1)^2 x^n &= \frac{d}{dx} S \\
&= \frac{d}{dx} \sum_{n=0}^{\infty}(n+1)^2 x^n \\
&= \frac{d}{dx} \frac{x}{(1-x)^2} \\
&= \frac{1}{(1-x)^2} + \frac{2x}{(1-x)^3}
\end{aligned}$$

Por lo que en conclusión, tenemos que para $|x| < 1$:

$$\sum_{n=0}^{\infty}(n+1)^2 x^n = \frac{1}{(1-x)^2} + \frac{2x}{(1-x)^3}$$

Regresando al segundo momento central, teníamos que $E(X^2) = p\sum_{n=1}^{\infty} n^2 (1-p)^{n-1} = p\sum_{n=0}^{\infty}(n+1)^2(1-p)^n$
y esta suma es igual a la que acabamos de calcular paro para $x = 1-p$, que cumple la condición de $|x| < 1$. Entonces tenemos ya el resultado:

$$\begin{aligned}
E(X^2) = p\sum_{n=0}^{\infty}(n+1)^2(1-p)^n &= p\frac{1}{(1-(1-p))^2} + p\frac{2(1-p)}{(1-(1-p))^3} \\
&= \frac{1}{p} + \frac{2}{p^2} - \frac{2}{p} \\
&= \frac{2}{p^2} - \frac{1}{p}
\end{aligned}$$

Con esto y la media calculada antes, ya podemos calcular la varianza:

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} \\ &= \frac{1}{p^2} - \frac{1}{p} \\ &= \boxed{\frac{1-p}{p^2}} \end{aligned}$$

Aplicación:
Esta variable aleatoria se utiliza cuando tenemos una secuencia de ensayos independientes de Bernoulli (eventos con una probabilidad p de éxito y $(1-p)$ de fracaso)
La variable aleatoria X cuenta el número de intentos hasta llegar al primer éxito. Es decir, $X = n$ indica que se necesitaron realizar n ensayos de Bernoulli de tal forma que los primeros $(n-1)$ fueron fracasos y el último fue el primer éxito.
Por ello la probabilidad $P(X = n)$ se calcula como $(1-p)^{n-1}p$.
Ya que $(1-p)^{n-1}$ calcula la probabilidad de tener $n-1$ fallos en ensayos de Bernoulli independientes (cada uno tiene probabilidad $1-p$). Y luego el p indica la probabilidad de que el n-ésimo ensayo sea un éxito.

Ejemplo: Una persona compra un boleto de lotería cada semana. La probabilidad de que gane la lotería en cualquiera de estas semanas es de $p = 1/1000000$. ¿Cuál es la probabilidad de que gane la lotería por primera vez justo en la semana $n = 10$? ¿Cuál es el tiempo medio que debe de esperar hasta ganar?

Tenemos varios eventos de Bernoulli (ganar o perder la lotería en una semana dada, con probabilidad de éxito p). Por lo dicho antes, podemos definir la variable aleatoria X como el número de semanas que le toma conseguir su primer éxito. Y entonces usaríamos la distribución geométrica $P(X = n) = (1-p)^{n-1}p$ para calcular la probabilidad de que pierda los primeros $n-1$ intentos y gane la lotería en el n-ésimo.
Entonces, la probabilidad de que gane la lotería por primera vez en la semana 10 es de $P(X = 10) = (1-p)^{10-1}p = (1 - \frac{1}{1000000})^9 * \frac{1}{1000000} \simeq 9{,}99 \times 10^{-7}$

Luego, como la esperanza de esta variable aleatoria es $E(X) = \dfrac{1}{p}$. Esperaríamos que en promedio le tome $\dfrac{1}{p}$ intentos hasta ganar la lotería por primera vez. Es decir, tendría que jugar $\dfrac{1}{\frac{1}{1000000}} = 1000000$ de veces.
Si juega una vez por semana, esto corresponde a 19231 años.

b) X es una variable aleatoria binomial negativa con distribución de probabilidad:

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}\ , \qquad n = r, r+1, \cdots\ ,\ 0 < p < 1,\ \ r \in \mathbb{N}$$

Calculamos la esperanza según la definición como $E(X) = \sum_{x_i} x_i f(x_i)$ donde x_i toma los valores que puede tomar la variable aleatoria, $r, r+1, r+2, \cdots$
Entonces, tenemos:

$$\begin{aligned}
E(X) &= \sum_{x=r}^{\infty} x f(x) = \sum_{x=r}^{\infty} x \binom{x-1}{r-1} p^r (1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} x \frac{(x-1)!}{(r-1)!(x-r)!} p^r (1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{x!}{(r-1)!(x-r)!} p^r (1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{x!r}{r!(x-r)!} p^r (1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \binom{x}{r} r p^r (1-p)^{x-r} \\
&= \frac{r}{p} \sum_{x=r}^{\infty} \binom{x}{r} p^{r+1} (1-p)^{x-r}
\end{aligned}$$

Y ahora hacemos un cambio de índice en la suma, cambiamos $x \to x-1$. Y además, definimos $k := r+1 \ \Rightarrow\ r = k-1$, por lo que la suma nos queda como:

$$\begin{aligned}
E(X) &= \frac{r}{p} \sum_{x=r+1}^{\infty} \binom{x-1}{r} p^{r+1} (1-p)^{x-1-r} \\
&= \frac{r}{p} \sum_{x=k}^{\infty} \binom{x-1}{k-1} p^k (1-p)^{x-k}
\end{aligned}$$

Podríamos notar que esta última suma no es otra cosa que una suma sobre todas las probabilidades de una distribución normal negativa (pero ahora con parámetro k), por lo que debe de dar como resultado 1 ya que la distribución binomial negativa es una distribución de probabilidad. Sin embargo, resolvamos explícitamente la suma para demostrar esto:

$$\begin{aligned}
E(X) &= \frac{r}{p}\sum_{x=k}^{\infty}\binom{x-1}{k-1}p^k(1-p)^{x-k} \\
&= \frac{r}{p}\sum_{x=k}^{\infty}\binom{x-1}{x-k}p^k(1-p)^{x-k} \\
&= \frac{r}{p}\sum_{y=0}^{\infty}\binom{y+k-1}{y}p^k(1-p)^{y} && \text{definimos } y=x-k \text{ y sumamos sobre y} \\
&= \frac{r}{p}\sum_{y=0}^{\infty}p^k(-1)^y\binom{-k}{y}(1-p)^{y} && \text{por la propiedad } \binom{y+k-1}{y}=(-1)^y\binom{-k}{y} \\
&= \frac{r}{p}p^k\sum_{y=0}^{\infty}(-1)^y\binom{-k}{y}(1-p)^{y} \\
&= \frac{r}{p}p^k\sum_{y=0}^{\infty}\binom{-k}{y}(p-1)^{y}
\end{aligned}$$

Pero por el teorema binomial (generalizado para exponentes negativos),
sabemos que $(1+z)^{-k}=\sum_{y=0}^{\infty}\binom{-k}{y}z^y$.
En este caso tenemos esa suma para $z=p-1$, por lo que concluimos que:

$$\begin{aligned}
&= \frac{r}{p}p^k(1+(p-1))^{-k} \\
&= \frac{r}{p}p^k(p)^{-k} \\
&= \frac{r}{p}(1)=\frac{r}{p}
\end{aligned}$$

Por lo que la esperanza queda como:

$$\boxed{E(X)=\frac{r}{p}}$$

Varianza: Para calcular la varianza, primero calcularemos $E(X(X+1))$ (que es más fácil de calcular) y usaremos este resultado para obtener el segundo momento central $E(X^2)$ y luego con ello obtener la varianza.

$$\begin{aligned}
E(X(X+1)) &= \sum_{x=r}^{\infty} x(x+1)f(x) = \sum_{x=r}^{\infty} x(x+1)\binom{x-1}{r-1}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} x(x+1)\frac{(x-1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{(x+1)!}{(r-1)!(x-r)!}p^r(1-p)^{x-r} \\
&= \sum_{x=r}^{\infty} \frac{(x+1)!(r+1)r}{(r+1)!(x-r)!}p^r(1-p)^{x-r} \\
&= (r+1)r\sum_{x=r}^{\infty} \frac{(x+1)!}{(r+1)!(x-r)!}p^r(1-p)^{x-r} \\
&= (r+1)r\sum_{x=r}^{\infty} \binom{x+1}{r+1}p^r(1-p)^{x-r}
\end{aligned}$$

Cambiamos el índice x por $x-2$ en la suma, por lo que nos queda:

$$E(X(X+1)) = (r+1)r\sum_{x=r+2}^{\infty} \binom{x-1}{r+1}p^r(1-p)^{x-2-r}$$

ahora definimos k = r+2 $\Rightarrow$ r = k-2 y lo sustituimos en la suma

$$\begin{aligned}
&= (r+1)r\sum_{x=k}^{\infty} \binom{x-1}{k-1}p^{k-2}(1-p)^{x-k} \\
&= \frac{r(r+1)}{p^2}\sum_{x=k}^{\infty} \binom{x-1}{k-1}p^{k}(1-p)^{x-k}
\end{aligned}$$

Notamos que esta última suma es una suma sobre todas las probabilidades de una variable binomial negativa (con parámetro k). Y como se trata de una distribución de probabilidad (y como demostramos antes), esta suma debe de valer 1. Por lo tanto, tenemos que:

$$E(X(X+1)) = \frac{r(r+1)}{p^2}$$

Ahora usamos que la esperanza es lineal, por lo que $E(X(X+1)) = E(X^2+X) = E(X^2)+E(X)$.
Por lo que $E(X^2) = E(X(X+1)) - E(X)$. Y estos resultados ya los tenemos, entonces $E(X^2) = \dfrac{r(r+1)}{p^2} - \dfrac{r}{p} = \dfrac{r(r+1)-rp}{p^2}$.

Finalmente, calculamos la varianza como:

$$\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \frac{r(r+1) - rp}{p^2} - \left(\frac{r}{p}\right)^2 \\
&= \frac{r(r+1) - rp - r^2}{p^2} \\
&= \frac{r - rp}{p^2} \\
&\Rightarrow \boxed{V(X) = r\frac{1-p}{p^2}}
\end{aligned}$$

Aplicación

Recordamos la distribución binomial negativa $P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r} \quad , \quad n = r, r+1, r+2, \cdots$

Digamos que repetimos varias veces un ensayo de Bernoulli con probabilidad p (donde cada repetición es independiente de las demás). Definimos la variable aleatoria X como el número de veces que hay que repetir el ensayo de Bernoulli para conseguir r éxitos (donde r es un parámetro definido antes).
Entonces, por ejemplo, si $r = 4$, $P(X = 6)$ nos da la probabilidad de que en un total de 6 intentos consigamos los 4 éxitos que buscamos.
Con esta definición, es fácil argumentar por qué se define así la función de probabilidad. Si queremos la probabilidad de que nos tome n intentos conseguir r éxitos (donde obviamente debemos de tener $n \geq r$), entonces calculamos la probabilidad de tener r éxitos, que es p^r y la multiplicamos por la probabilidad de que el resto $n - r$ de los intentos sean fallos $(1-p)^{n-r}$.
Luego, por cómo definimos la variable aleatoria, el último intento tiene que ser un éxito (porque al llegar al r-ésimo éxito, dejamos de hacer los ensayos de Bernoulli y ya tenemos la cantidad n de intentos que nos tomó conseguir los r éxitos). Pero los otros $r - 1$ éxitos se pueden acomodar como sea entre los primeros $n - 1$ intentos, y es por ello que multiplicamos por el factor combinatorio $\binom{n-1}{r-1}$.

Ejemplo (Muestreo): Esta variable aleatoria se puede usar para hacer muestreos.
Por ejemplo, digamos que 10 % de la población de la facultad de ciencias lleva aprobadas todas sus materias del semestre actual. Vamos preguntando a las personas de la facultad si llevan sus materias aprobadas Y queremos encontrar 20 personas que cumplan con esta característica para hacerles un cuestionario. En promedio, ¿A cuántas debemos de preguntar si tienen sus materias aprobadas para conseguir a las 20 personas solicitadas?

Si consideramos que preguntar a una persona si tiene sus materias aprobadas es un ensayo de Bernoulli con una probabilidad de éxito de $p = 0{,}10$ y suponemos que estos ensayos son independientes entre sí (para lo que sería adecuado que la población de la facultad fuera lo

más grande posible). Entonces el ejercicio consiste en encontrar el promedio de cantidad de ensayos de Bernoulli que necesitaremos hacer antes de conseguir r= 20 éxitos.
Pero esto es justo lo que mide la distribución binomial negativa y nos da una experanza de:

$$E(X) = \frac{r}{p} = \frac{20}{0{,}1} = 200$$

Por lo que tendríamos que acudir con aproximadamente 200 personas. Lo cual coincide con lo que esperaríamos para conseguir 20 personas que tienen una característica de una prevalencia del 10 %.

c) X es una variable aleatoria con distribución Gamma con función de densidad:

$$f(x) = \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \qquad , \ \alpha, \lambda > 0 \quad , \quad x > 0$$

Esperanza: Se trata de una distribución continua y la función $f(x)$ es su función de densidad. Entonces, la esperanza se calcula como:

$$\begin{aligned}
E(X) &= \int_{\mathbb{R}} x f(x) dx \\
&= \int_0^\infty x \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{y}{\lambda}\right)^\alpha e^{-y} \frac{1}{\lambda} dy \qquad \text{hacemos el cambio } y = \lambda x \ \Rightarrow \ dy = \lambda dx \\
& \qquad\qquad\qquad\qquad\qquad \text{y como } \lambda > 0 \ \text{, entonces } x = 0 \ \Rightarrow \ y = 0 \ , \ x = \infty \ \Rightarrow \ y = \infty \\
&= \frac{1}{\lambda \Gamma(\alpha)} \int_0^\infty y^\alpha e^{-y} dy \\
&= \frac{1}{\lambda \Gamma(\alpha)} \Gamma(\alpha + 1) \qquad \text{reconocemos que la integral es la función Gamma evaluada en } \alpha + 1 \\
&= \boxed{\frac{\alpha}{\lambda}} \qquad\qquad \text{porque } \Gamma(\alpha+1) = \alpha\Gamma(\alpha)
\end{aligned}$$

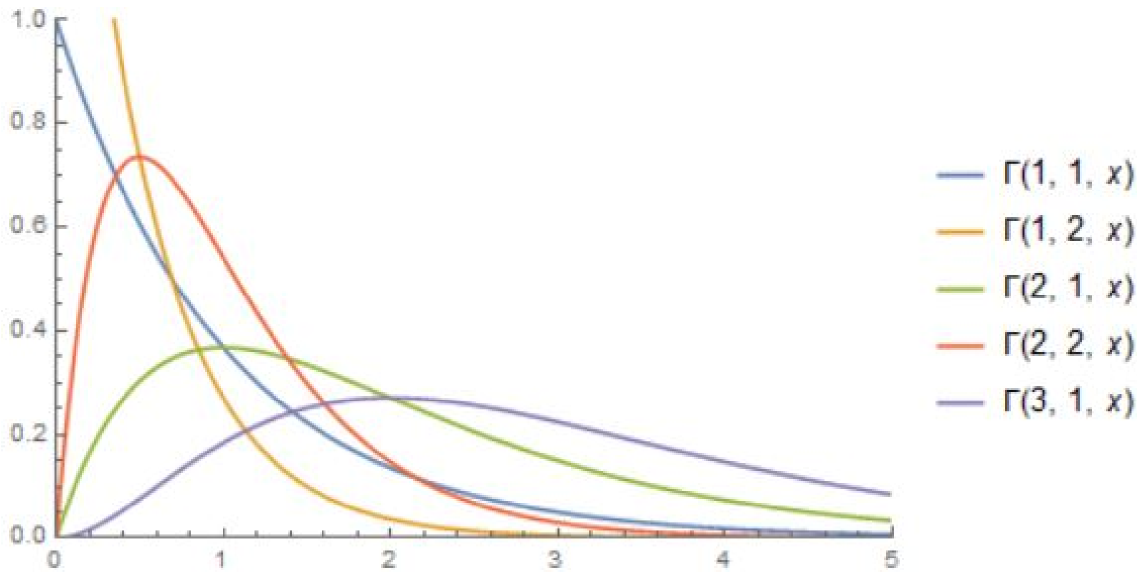
Varianza: Antes de calcular la varianza, calculamos el segundo momento central $E(X^2)$.

$$\begin{aligned}
E(X^2) &= \int_{\mathbb{R}} x^2 f(x) dx \\
&= \int_0^\infty x^2 \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1} e^{-\lambda x} dx \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{y}{\lambda}\right)^{\alpha+1} e^{-y} \frac{1}{\lambda} dy \qquad \text{hacemos el cambio } y = \lambda x \ \Rightarrow \ dy = \lambda dx \\
& \qquad\qquad\qquad\qquad\qquad \text{y como } \lambda > 0 \ \text{, entonces } x = 0 \ \Rightarrow \ y = 0 \ , \ x = \infty \ \Rightarrow \ y = \infty \\
&= \frac{1}{\Gamma(\alpha)\lambda^2} \int_0^\infty y^{\alpha+1} e^{-y} dy \\
&= \frac{1}{\Gamma(\alpha)\lambda^2} \Gamma(\alpha + 2) \\
&= \frac{1}{\Gamma(\alpha)\lambda^2} (\alpha+1)(\alpha)\Gamma(\alpha) \\
&= \frac{(\alpha+1)\alpha}{\lambda^2}
\end{aligned}$$

Entonces ya podemos calcular la varianza:

$$\begin{aligned}V(X) &= E(X^2) - E(X)^2 \\ &= \frac{\alpha^2+\alpha}{\lambda^2} - \frac{\alpha^2}{\lambda^2} \\ &= \boxed{\frac{\alpha}{\lambda^2}}\end{aligned}$$

Aplicación: La distribución γ cuenta con dos parámetros. El parámetro α llamado parámetro de forma y el parámetro λ que a veces se conoce como parámetro de escala. Graficamos algunas de estas funciones para distintos valores de los parámetros:



Como vemos, dependiendo del valor de los parámetros, la función de distribución puede tomar varias formas distintas por lo que puede modelar muchos eventos. Además, a diferencia de la distribución normal por ejemplo, la distribución Gamma sirve para modelar sucesos con algún sesgo en alguna dirección, ya que como se ve en las gráficas, no son funciones simétricas en general.

Ejemplo: Se puede usar para modelar tiempos de espera antes de que suceda algo. Por ejemplo, las horas de vida de una batería de un aparato electrónico quizá podría ser modelada por una distribución Gamma como $\Gamma(\alpha = 5, \lambda = 1/4)$.

Entonces, la batería tiene en promedio una vida de $E(X) = \dfrac{\alpha}{\lambda} = \dfrac{5}{1/4} = 20$ horas.

Con una varianza de $V(X) = \dfrac{\alpha}{\lambda^2} = \dfrac{5}{(1/4)^2} = 80$ (una desviación estándar de $\sqrt{80} = 8{,}94$ horas).

Y por ejemplo, podemos calcular la probabilidad de que la batería dure más de 50 horas al integral la distribución Gamma desde 50 hasta infinito $P(X \geq 50) = \int_{50}^{\infty} f(x\,; \alpha = 5, \lambda = 1/4)dx \simeq 0{,}0053$.

d) X es una variable aleatoria con distribución Cauchy con función de densidad:

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2} \quad , \quad x \in \mathbb{R}$$

Esperanza: Calculamos la esperanza como hemos visto $E(X) = \int_{\mathbb{R}} xf(x)dx$

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} xf(x) = \int_{-\infty}^{\infty} x\frac{1}{\pi}\frac{1}{1+x^2}dx \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x}{1+x^2} \\
&= \frac{1}{\pi}\int_{-\infty}^{0}\frac{x}{1+x^2}dx + \frac{1}{\pi}\int_{0}^{\infty}\frac{x}{1+x^2}dx \\
&= \lim_{a\to\infty}\frac{1}{\pi}\int_{-a}^{0}\frac{x}{1+x^2}dx + \lim_{b\to\infty}\int_{0}^{b}\frac{1}{\pi}\frac{x}{1+x^2}dx \quad \text{por cómo se define una integral impropia} \\
&= \lim_{a\to\infty}\frac{1}{\pi}\left[\frac{1}{2}\log(1+x^2)\right]_{-a}^{0} + \lim_{b\to\infty}\frac{1}{\pi}\left[\frac{1}{2}\log(1+x^2)\right]_{0}^{b} \\
&= \lim_{a\to\infty}\frac{1}{\pi}\left[-\frac{1}{2}\log(1+(-a)^2)\right] + \lim_{b\to\infty}\frac{1}{\pi}\left[\frac{1}{2}\log(1+b^2)\right] \\
&= \lim_{a\to\infty}\frac{1}{\pi}\left[-\frac{1}{2}\log(1+a^2)\right] + \lim_{b\to\infty}\frac{1}{\pi}\left[\frac{1}{2}\log(1+b^2)\right]
\end{aligned}$$

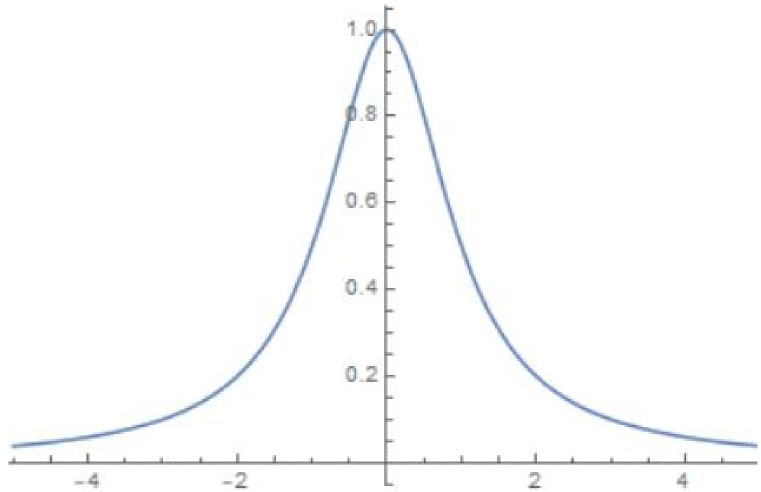
Conforme a tiende a infinito, claramente el límite $\log(1+a^2)$ se hace infinitamente grande y por tanto $\lim_{a\to\infty}\frac{1}{\pi}\left[-\frac{1}{2}\log(1+a^2)\right] = -\infty$.
Y conforme $b \to \infty$, el límite de $\log(1+b^2)$ tiende a infinito y por tanto el límite es $\lim_{b\to\infty}\frac{1}{\pi}\left[-\frac{1}{2}\log(1+b^2)\right] = \infty$
Por lo que un límite tiende a $-\infty$ y el otro a ∞. Y entonces la suma de ambos límites $-\infty+\infty$ no está definida.
Por lo tanto, la media no existe.

Varianza: La varianza también está indefinida pues para empezar, no podemos usar la definición que tenemos, ya que involucra conocer la media. Aún así, si la intentáramos calcular, tenemos el mismo problema:

$$\begin{aligned}
V(X) = E((x-\mu)^2) &= \int_{-\infty}^{\infty}\frac{1}{\pi}\frac{1}{1+x^2}(x-\mu)^2dx \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x^2}{1+x^2} - \frac{2\mu x}{1+x^2} + \frac{\mu}{1+x^2}dx
\end{aligned}$$

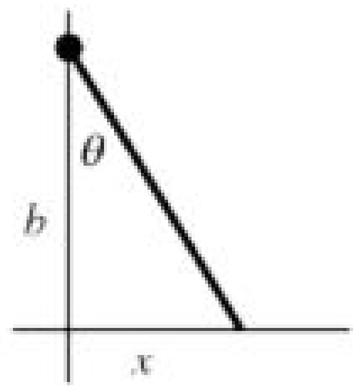
La segunda integral $-\frac{2\mu}{\pi}\int_{-\infty}^{\infty}\frac{1}{1-x^2}dx$ ya la habíamos calculado para la media y está indefinida.

Aplicación:
Vemos primero una gráfica de la función para tener una idea de cómo se ve:



La distribución de Cauchy puede ser útil para resolver el problema siguiente:

Tenemos un punto en $(0, b)$ como se ve en la imagen. Pensamos en que este punto es una fuente de luz, como puede ser un faro por ejemplo o algún tipo de partícula que emite radiación. La luz de esta fuente es lanzada con una simetría esférica, por lo que sale en todas las direcciones angulares θ, $\theta \in [-\pi/2, \pi/2]$ (con θ el ángulo que se muestra en la figura) con la misma intensidad.
Nos preguntamos cómo se distribuye la densidad de probabilidad de encontrar un rayo de luz en el eje x.



La probabilidad de que un rayo de luz específico salga con un ángulo de entre θ y $\theta + d\theta$ es de $\frac{1}{\pi}d\theta$.
Esto para que así la probabilidad en cualquier ángulo es la misma (porque supusimos que la fuente de luz tenía esta simetría) y el factor de $\frac{1}{\pi}$ aparece para que la probabilidad total de encontrar un rayo en todo el espacio $\theta \in [-\pi/2, \pi/2]$ sea de $\int_{-\pi/2}^{\pi/2} \frac{1}{\pi}d\theta = 1$.
Por lo tanto, la función de densidad de probabilidad para el ángulo es de $f_\theta(\theta) = \frac{1}{\pi}$.
Entonces, la probabilidad de que el rayo tenga un ángulo entre θ y $\theta + d\theta$ es de $f(\theta)d\theta = \frac{1}{\pi}d\theta$.

Ahora nos fijamos en el eje x y nos preguntamos, dado un rayo de luz que sale de la fuente, cuál es la función de densidad que describe la probabilidad de encontrar el rayo en cada intervalo del eje x.

Es decir, queremos convertir la función de densidad $f_\theta(\theta) = \dfrac{1}{\pi}$ para que ahora aplique a la variable x.

Por la figura, vemos que $x = b\tan\theta \quad \Rightarrow \quad \theta = \arctan(x/b)$.

Luego, por cómo se transforman las funciones de densidad de una variable a otra, tal como vimos en el ejercicio 5a), tenemos que la función de densidad para x es:

$$
\begin{aligned}
f_x(x) &= f_\theta(b\tan\theta)\cdot\frac{d\theta}{dx} \\
&= \frac{1}{\pi}\frac{d}{dx}(\arctan(x/b)) \\
&= \frac{1}{\pi}\frac{1}{b\left(1+\dfrac{x^2}{b^2}\right)} \\
&= \frac{b}{\pi b^2(1+\frac{x^2}{b^2})} \\
&= \frac{b}{\pi(b^2+x^2)}
\end{aligned}
$$

Esta es la función de densidad que expresa la probabilidad de encontrar un rayo de luz en el eje X. Vemos que esta función de distribución es la de Cauchy (para $b = 1$ coincide con el caso particular del ejercicio).

Con lo que probamos que la distribución de Cauchy puede modelar los impactos de energía sobre una línea recta que llegan emitidos de una fuente puntual.

e) X es una variable aleatoria con distribución Beta con función de densidad:

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} \quad , \quad a,b > 0 \quad , \qquad 0 < x < 1$$

Esperanza: Calculamos la esperanza directamente como lo hacemos para variables aleatorias continuas.

$$\begin{aligned}
E(X) &= \int_{\mathbb{R}} x f(x) \\
&= \int_0^1 x \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} dx \\
&= \frac{1}{B(a,b)} \int_0^1 x^{a}(1-x)^{b-1} dx \\
&= \frac{1}{B(a,b)} B(a+1,b) \qquad \text{por la def. de Beta } B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx \\
&\qquad\qquad\qquad\qquad\qquad\quad \text{con } p = a+1 > 0, q = b > 0 \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+1+b)} \qquad \text{por la relación entre } \Gamma \text{ , } B \\
&= \frac{\Gamma(a+b)\Gamma(a)a\Gamma(b)}{\Gamma(a)\Gamma(b)\Gamma(a+b)(a+b)} \\
&= \boxed{\frac{a}{a+b}}
\end{aligned}$$

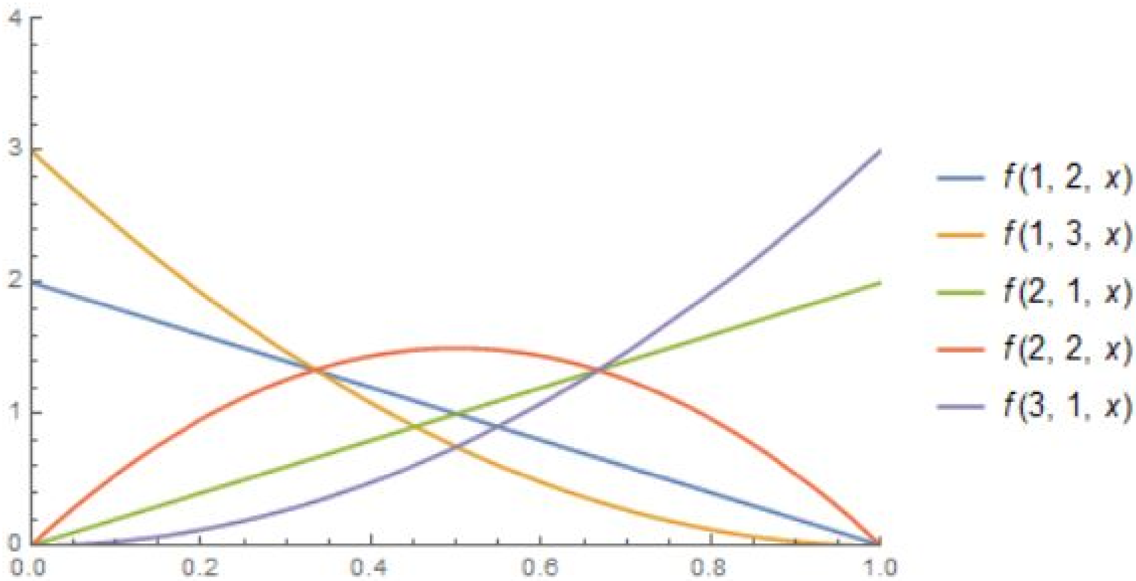
Varianza: Para calcular la varianza calculamos primero el segundo momento central:

$$\begin{aligned}
E(X^2) &= \int_{\mathbb{R}} x^2 f(x) \\
&= \int_0^1 x^2 \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1} dx \\
&= \frac{1}{B(a,b)} \int_0^1 x^{a+1}(1-x)^{b-1} dx \\
&= \frac{1}{B(a,b)} B(a+2,b) \qquad \text{por la def. de Beta } B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx \\
&\qquad\qquad\qquad\qquad\qquad\quad \text{con } p = a+2 > 0, q = b > 0 \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)} \qquad \text{por la relación entre } \Gamma, B \\
&= \frac{\Gamma(a+b)\Gamma(a)(a+1)(a)\Gamma(b)}{\Gamma(a)\Gamma(b)\Gamma(a+b)(a+b+1)(a+b)} \\
&= \boxed{\frac{(a+1)a}{(a+b+1)(a+b)}}
\end{aligned}$$

Con esto podemos calcular la varianza con la relación $V(X) = E(X^2) - E(X)^2$.

$$\begin{aligned} V(X) &= E(X^2) - E(X)^2 \\ &= \frac{(a+1)a}{(a+b+1)(a+b)} - \frac{a^2}{(a+b)^2} \\ &= \frac{(a+1)a(a+b) - a^2(a+b+1)}{(a+b+1)(a+b)^2} \\ &= \frac{a^3 + a^2 + ab + a^2b - (a^3 + a^2b + a^2)}{(a+b+1)(a+b)^2} \\ &= \boxed{\frac{ab}{(a+b+1)(a+b)^2}} \end{aligned}$$

Aplicaciones:
Como vemos, la función de distribución cuenta con dos parámetros a, b.
Como se puede ver en la siguiente gráfica, al variar estos parámetros se pueden crear muchas distribuciones distintas.



Una característica de la distribución Beta que la diferencia de la Gamma y de otras es que está restringida a $x \in [0, 1]$. Lo que significa que es útil para variables aleatorias restringidas a cierto intervalo.

La distribución Beta se utiliza generalmente para obtener una idea de la probabilidad de un evento si lo único que conocemos son algunos resultados.

Para ser más específicos, supongamos que tenemos una moneda pesada con una probabilidad p de salir sol y una probabilidad $1 - p$ de salir águila. Digamos que en principio nosotros no conocemos el valor de p, pero tiramos varias veces la moneda y contamos el número de soles.

Entonces la distribución Beta nos servirá para ver cuáles son los valores más esperados de p a partir de estos resultados.

Definimos la variable aleatoria S_n como el número de soles que se obtiene al tirar la moneda n veces. Y digamos que después de tirar la moneda n veces, obtuvimos k soles, es decir $S_n = k$.
Si supiéramos ya que la probabilidad de un sol es p, entonces la variable aleatoria S_n se distribuye con una distribución binomial, es decir:

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Sin embargo, a priori no conocemos el valor de p y justamente queremos usar el resultado de que de n lanzamientos se obtuvieron k soles para acercarnos al valor de p.

Definimos $f(p) : [0,1] \to \mathbb{R}$ como una función de densidad para la variable p. Es decir, usaremos $f(p)$ para determinar la certeza que tenemos de que la probabilidad de que salga sol sea p. El valor $f(p)dp$ nos indicará la 'probabilidad' (o más bien la certeza que tenemos) de que la moneda tenga una probabilidad p de salir sol.
A priori no sabemos nada sobre la moneda, sólo que puede estar pesada de cualquier forma. Por lo que a priori, la función $f(p)$ debe de ser $f(p) = 1 \ \ , \ \ 0 \leq p \leq 1$, indicando que todos los valores que puede tomar p son igual de inciertos, ya que no tenemos conocimiento sobre la moneda.

Sin embargo, al ir obteniendo resultados sobre la moneda vamos a ir actualizando el valor de $f(p)$, indicando que tenemos más certeza sobre el valor p de la probabilidad de que salga sol. Por ejemplo, si tras muchos experimentos vemos que salen la misma cantidad de soles que águilas, $f(p)$ va a ser una función muy alta en $1/2$ y va a valer casi 0 en los extremos de $[0,1]$, indicando que nuestros experimentos nos indican que el valor de p probablemente está cerca de $1/2$. Mientras que si tras muchos tiros salen casi puros soles, $f(p)$ va a ser más alta cerca del 1, indicando que estamos casi seguros que p debe de tener un valor cercano a 1 y la moneda favorece más al sol

Nos interesa saber exactamente cómo se actualiza la función $f(p)$ tras registrar varios resultados. Por ejemplo, si como dijimos antes, de n lanzamientos, obtenemos k soles, nos interesa conocer la función $f(p|S_n = k)$ que indica nuestra certeza sobre el valor de p dado que sabemos que $S_n = k$.

Para ello, necesitaremos primero calcular $P(S_n = k)$. Es decir, la probabilidad de que de n tiros salgan k soles a priori, sin saber el valor de p. Como dijimos antes, el valor de p a priori está totalemente indeterminado y su función de distribución sería $f(p) = 1 \ \ , \ \ 0 \leq p \leq 1$.
Para calcular $P(S_n = k)$, como no sabemos el valor de p, hay que considerar todos los valores posibles de p y usar la ley de probabilidad total (pero para el caso continuo en vez de discreto,

en el que la suma de la LPT se convierte en integral). Con esta ley, se puede ver que:

$$\begin{aligned}
P(S_n = k) &= \int_0^1 P(S_n = k \mid p) f(p) dp \\
&= \int_0^1 P(S_n = k \mid p) dp \qquad \text{a priori f(p) = 1} \\
&= \int_0^1 \binom{n}{k} p^k (1-p)^{n-k} dp \qquad \text{por lo visto antes para } P(S_n = k|p)
\end{aligned}$$

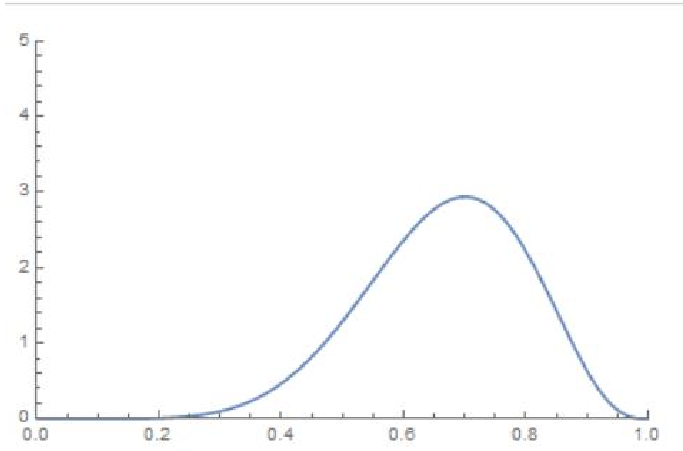
Finalmente, para calcular $f(p|S_n = k)$ (nuestra correción sobre la certeza que tenemos del valor de p, dado que vimos k soles en n tiradas) usamos el teorema de Bayes:

$$\begin{aligned}
f(p|S_n = k) = \frac{P(S_n = k|p)}{P(S_n = k)} &= \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\int_0^1 \binom{n}{k} p^k (1-p)^{n-k} dp} \\
&= \frac{p^k (1-p)^{n-k}}{\int_0^1 p^k (1-p)^{n-k} dp} \\
&= \frac{p^k (1-p)^{n-k}}{B(k+1, n-k+1)} \qquad \text{vemos que la integral de abajo es la def de función beta}
\end{aligned}$$

Y con esto, nos queda que la función de densidad del valor de p tras conocer que $S_n = k$ es $f(p|S_n = k) = \dfrac{p^k (1-p)^{n-k}}{B(k+1, n-k+1)}$, que es justamente una distribución Beta con $a = k+1$, $b = n - k$.

Es decir, tras n tiros de la moneda con k soles, nuestra certeza sobre el valor de p (el peso de la moneda o probabilidad de que salga sol) se actualiza y es descrito por una función B

Por ejemplo, si tenemos una moneda de la que no conocemos su peso p pero tras $n = 10$ tiradas resultan $k = 7$ caras, entonces la función $f(p)$ actualizada sería $f(p|S_{10} = 7) = \dfrac{p^7(1-p)^3}{B(8,4)}$ que tiene la siguiente gráfica:



Vemos que $f(p|S_{10} = 7)$ tiene un pico cerca de $p = 0{,}7$, indicando que tras los 10 tiros, tenemos más certeza de que el valor de p está cerca de 0.7 (Que es nuestra mejor conjetura al tener 7 soles en 10 tiros).

Además, esta explicación nos ayuda a darnos cuenta que si aumentamos n (por ejemplo $n = 100$) y aumentamos k (por ejemplo $k = 70$), entonces la curva $f(p|S_n = k)$ se va a hacer más delgada y alta alrededor de $p = 0{,}7$ porque tenemos más certeza de que el peso de la moneda p sea cercano a 0.7. Graficamos ahora $f(p|S_{100} = 70)$ y vemos que efectivamente se cumple esto:

.

Vemos que efectivamente la certeza de que p esté cerca de 0.7 se hace mucho más grande.

Esto incluso puede explicar que la distribución Beta se acerca a una delta de Dirac cuando $n \to \infty$. Esto porque después de tantos intentos ya tendremos mucha certeza sobre el valor de p, por lo que por ejemplo la curva $f(p|S_{1000} = 700)$ se acerca mucho a un pico muy alto y delgado en 0,7 y eventualmente se aproximará a una delta de Dirac en $p = 0{,}7$