

Intro a la Probabilidad y estadística



Martes y Jueves Aula B17
Dra Ana Georgina Flesia

Espacio vs Variable

- Si tiro un dado 20 veces consecutivas, el espacio muestral Ω asociado al experimento consta de vectores 20-dimensionales con seis posibilidades para cada entrada. Todos esos puntos son igualmente probables de ser observados, por lo cual $P(\omega) = 1/\#\Omega = 1/6^{20}$
- - **X= número de 6 en las 20 tiradas**
 - **Y= número de valores pares en las 20 tiradas**
 - **Z= número de rachas de tres 6 consecutivos**
- son variables aleatorias que concentran información de interés, y los valores que toman esas variables no tienen igual probabilidad, porque refieren a conjuntos de valores equiprobables.

Experimento hipergeométrico

- ▶ Consideremos una población de r objetos, de los cuales r_1 son de un tipo y $r_2 = r - r_1$ son de otro tipo.
- ▶ Supongamos que extraemos un subconjunto de objetos $\{x_1, \dots, x_n\}$ de tamaño $n \leq r_1$ de la población.
- ▶ La extracción de un subconjunto de una población se realiza seleccionando elementos al azar sin reponerlos a la población, sin considerar el orden de extracción.

Variable hipergeométrica $\mathcal{H}(r_1, r, n)$

- ▶ Sea X el número de los objetos del primer tipo en la muestra. Entonces X es una variable aleatoria cuyos valores posibles son $0, 1, \dots, n$.
- ▶ Calculemos su función de densidad discreta.

Función de masa hipergeométrica $\mathcal{H}(r_1, r, n)$

Observemos que

1. Todas las posibles muestras son equiprobables, pues son elegidas al azar, así que la probabilidad de elegir una muestra particular es

$$P(\{x_1, \dots, x_n\}) = \frac{1}{\binom{r}{n}}$$

2. El evento $(X = x)$ contiene todas las muestras (subconjuntos) con x elementos del tipo 1, y $r - x$ elementos del tipo 2.
3. Hay $\binom{r_1}{x}$ formas de elegir los x elementos del tipo 1 y $\binom{r - r_1}{n - x}$ formas de elegir los elementos del tipo 2 entre $r - x$ elementos.

Función de masa hipergeométrica $\mathcal{H}(r_1, r, n)$

1. Entonces

$$p_X(x) = P(X = x) = \frac{\binom{r_1}{x} \binom{r - r_1}{n - x}}{\binom{r}{n}}$$

Una variable X con esta distribución se dice que tiene densidad Hipergeométrica de parámetros (r_1, r, n) .

Esperanza de una variable $\mathcal{H}(r_1, r, n)$

Si X tiene distribución hipergeométrica $\mathcal{H}(r_1, r, n)$ entonces

$$E(X) = \frac{nr_1}{r} \quad Var(X) = \frac{r_1 n (r - r_1) (r - n)}{r^2 (r - 1)}$$

Ejemplo de una variable $\mathcal{H}(r_1, r, n)$

La cámara legislativa de cierta ciudad está compuesta por 30 concejales de los cuales 18 pertenecen al partido A y los otros 12 al partido B. Se elige al azar una comisión de 5 concejales.

1. ¿Cuál es la probabilidad de que todos los miembros de la comisión elegida sean del mismo partido?
2. ¿Cuál es la probabilidad de que la mayoría de los miembros de la comisión pertenezcan al partido A?

Ejemplo de una variable $\mathcal{H}(r_1, r, n)$

- Consideremos la variable aleatoria X número de concejales del partido A de la comisión, es una variable hipergeométrica de parámetros $r_1 = 18$, $r = 30$ y $n = 5$.
- La probabilidad de que todos los concejales elegidos sean de un mismo partido es $P(X = 5) + P(X = 0)$ pues ocurre que los 5 pertenecen al partido A o bien los 5 no pertenecen al partido A .

$$P(X = 5) + P(X = 0) = \frac{\binom{18}{5} \binom{12}{0}}{\binom{30}{5}} + \frac{\binom{18}{0} \binom{12}{5}}{\binom{30}{5}}$$

- La probabilidad de que la mayoría de los concejales sea del partido A es $P(X \geq 3)$, y se calcula

$$P(X \geq 3) = \frac{\binom{18}{3} \binom{12}{2}}{\binom{30}{5}} + \frac{\binom{18}{4} \binom{12}{1}}{\binom{30}{5}} + \frac{\binom{18}{5} \binom{12}{0}}{\binom{30}{5}}$$

Experimento geométrico

- ▶ Supongamos que podemos repetir experimentos Bernoulli, cada uno con valores posibles E o F , con probabilidades $P(E) = p$, $P(F) = 1 - p$, en forma independiente hasta obtener el primer éxito.
- ▶ Este experimento produce sucesiones finitas de fracasos terminados en éxito, de cualquier tamaño. Es un experimento con infinitos resultados.
- ▶ Sea X el número de experimentos Bernoulli necesarios hasta obtener un éxito.

Variable Geométrica $\mathcal{G}(p)$

- ▶ Sea X el número de experimentos Bernoulli necesarios hasta obtener un éxito.
- ▶ X es una variable con infinitos valores posibles $1, 2, 3, \dots$
- ▶ Si X toma el valor k , entonces hay $k - 1$ fracasos y 1 éxito en la sucesión de experimentos independientes. Ese punto tiene probabilidad $(1 - p)^{k-1}p$. Entonces

$$P(X = k) = (1 - p)^{k-1}p \quad k = 1, 2, 3, \dots$$

- ▶ Esta es una distribución bien definida pues $p_X(k) \geq 0$ y

$$\begin{aligned} \sum_{k=1}^{\infty} p_X(k) &= \sum_{k=1}^{\infty} (1 - p)^{k-1}p = p + p \sum_{n=1}^{\infty} (1 - p)^n \\ &= p + \frac{(1 - p)}{1 - (1 - p)}p = p + (1 - p) = 1 \end{aligned}$$

usando el límite de la serie geométrica de razón $(1 - p)$.

Esperanza Variable Geométrica $\mathcal{G}(p)$

- Sea X una variable con distribución Geométrica $\mathcal{G}(p)$, su esperanza es

$$E(X) = \frac{1}{p} \quad Var(X) = \frac{(1-p)}{p^2}$$

Ejemplo Variable Geométrica $\mathcal{G}(p)$

La sincronización de los semáforos del Boulevard Chacabuco esta hecha de tal forma que un semáforo da paso con probabilidad 0.2.

1. ¿Cuántos semáforos esperaría encontrar hasta el primer semáforo en verde (inclusive)?
2. ¿Cual es la probabilidad de que el primer semáforo en verde sea el tercero que encuentre?

Ejemplo Variable Geométrica $\mathcal{G}(p)$

- ▶ Se sabe entonces que $p = 0.2$ es la probabilidad de éxito, no tener que detenerse en un semáforo.
- ▶ Si consideramos un fracaso tener que detenernos, la probabilidad de fracaso es $q = 0.8$.
- ▶ Sea X el número de semáforos que pasamos hasta encontrar el primero en verde, es una variable geométrica de parámetro $p = 0.2$ e incluye al semáforo en verde.

Ejemplo Variable Geométrica $\mathcal{G}(p)$

- ¿Cuántos semáforos espera encontrar hasta el primer semáforo en verde? Es la esperanza de X ,

$$E(X) = \frac{1}{p} = \frac{1}{0.2} = 5$$

Son cinco semáforos, inclusive el semáforo en verde. Si contamos los semáforos en rojo, es la esperanza de $Y = X - 1$, el número de fracasos, por lo cual

$$E(Y) = E(X - 1) = \frac{1}{p} - 1 = \frac{1}{0.2} - 1 = 5 - 1 = 4$$

- ¿Cual es la probabilidad de que el primer semáforo en verde sea el tercero que encuentre? Esto es $P(X = 3)$

$$P(X = 3) = 0.2(0.8)^2 = 0.128$$

Observación Variable Geométrica $\mathcal{G}(p)$

1. Supongamos que arrojamamos un dado hasta observar el primer as.
2. La probabilidad de observar un as en una sola tirada es $1/6$, por lo cual si X es el número de tiradas necesarias hasta observar el primer as, X tiene distribución geométrica de parámetro $p = 1/6$.
3. La variable X no puede tomar el valor 0, pues eso implicaría que el dado nunca se arrojó. Sin embargo, en otros casos, el valor cero implicaría que el proceso no empezó.

Ejemplo Variable Geométrica $\mathcal{G}(p)$

- Sea X el número de horas completas hasta que se apaga una lámpara. Esta variable mide el número de intentos hasta encontrar un éxito, en este caso, que se apague la lámpara y es una variable Geométrica clásica.

$$p_X(k) = p(1 - p)^{k-1} \quad k = 1, 2, \dots$$

- Si consideramos Y el número de horas completas hasta que falla el tubo de un televisor, podemos considerar que $Y = 0$ representa el evento “el televisor falló antes de la primera hora”, lo cual incluye el hecho de que el televisor puede no encender.
- En ese caso, Y mide el número de fracasos hasta obtener un éxito, y tiene función de masa

$$p_Y(k) = P(Y = k) = p(1 - p)^k \quad k = 0, 1, 2, 3, \dots$$

- Es importante observar que $Y = X - 1$.

Propiedades Variable Geométrica $\mathcal{G}(p)$

Sea X una variable con distribución geométrica de parámetro p . Entonces se cumple que:

1. La función de distribución de X es

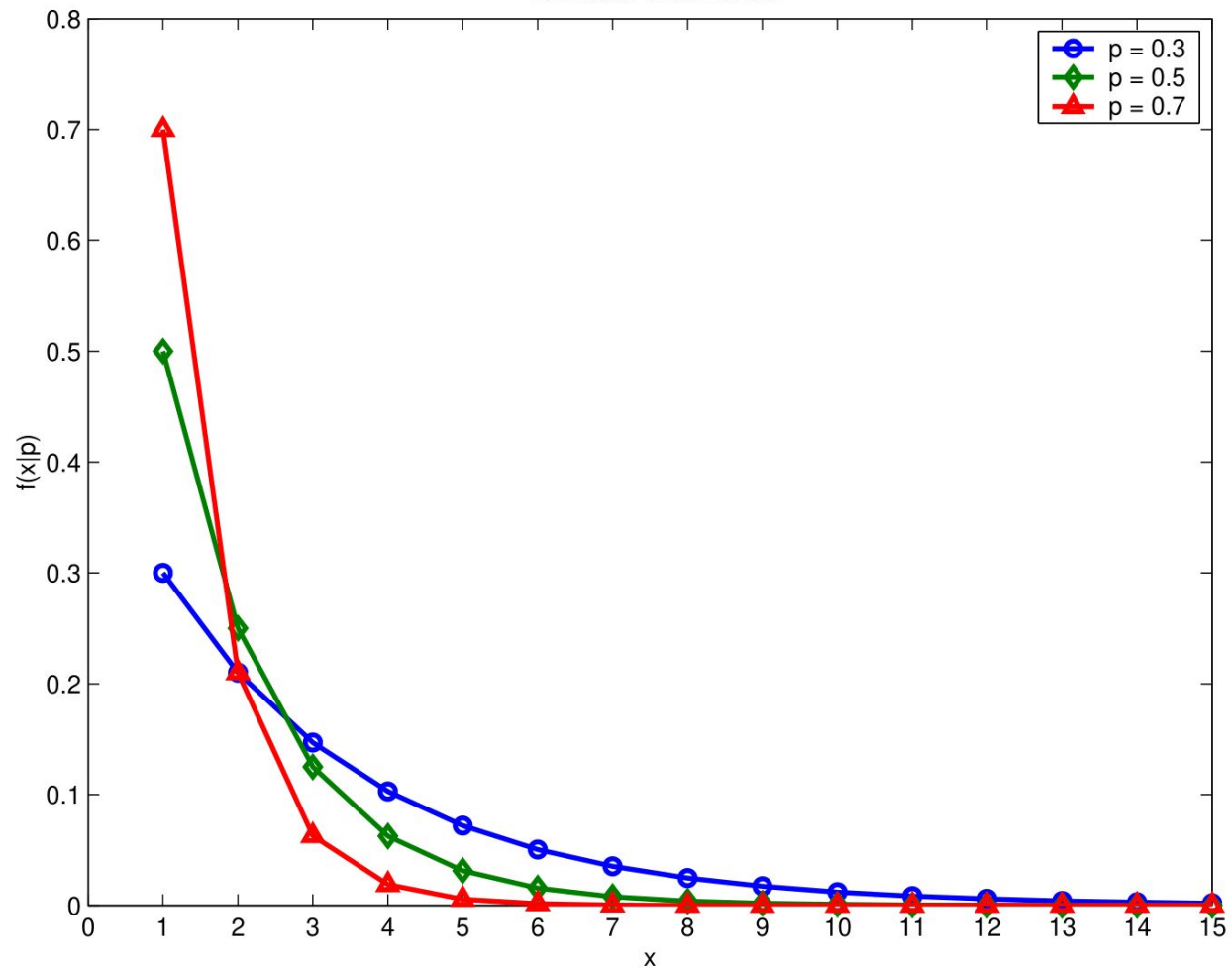
$$F_X(t) = 1 - (1 - p)^{[t]} \quad t \geq 1, t \in \mathbb{R}$$

2. $P(X \geq k) = (1 - p)^{k-1}$ para todo $k \in \mathbb{N}$.

3. La distribución geométrica sufre de amnesia, esto es

$$P(X \geq s + t | X > t) = P(X \geq s)$$

Densidad Geometrica



Experimento Binomial Negativa $\mathcal{BN}(r, p)$

- ▶ Supongamos que tenemos una sucesión de experimentos Bernoulli independientes con probabilidad de éxito p .
- ▶ Consideremos el evento donde k experimentos son necesarios para obtener r éxitos.
- ▶ Toda secuencia de experimentos de ese evento tiene probabilidad $(1 - p)^{k-r} p^r$.
- ▶ La cantidad de secuencias del evento es el número de subconjuntos donde puedo posicionar los $r - 1$ éxitos.

Distribución Binomial Negativa $\mathcal{BN}(r, p)$

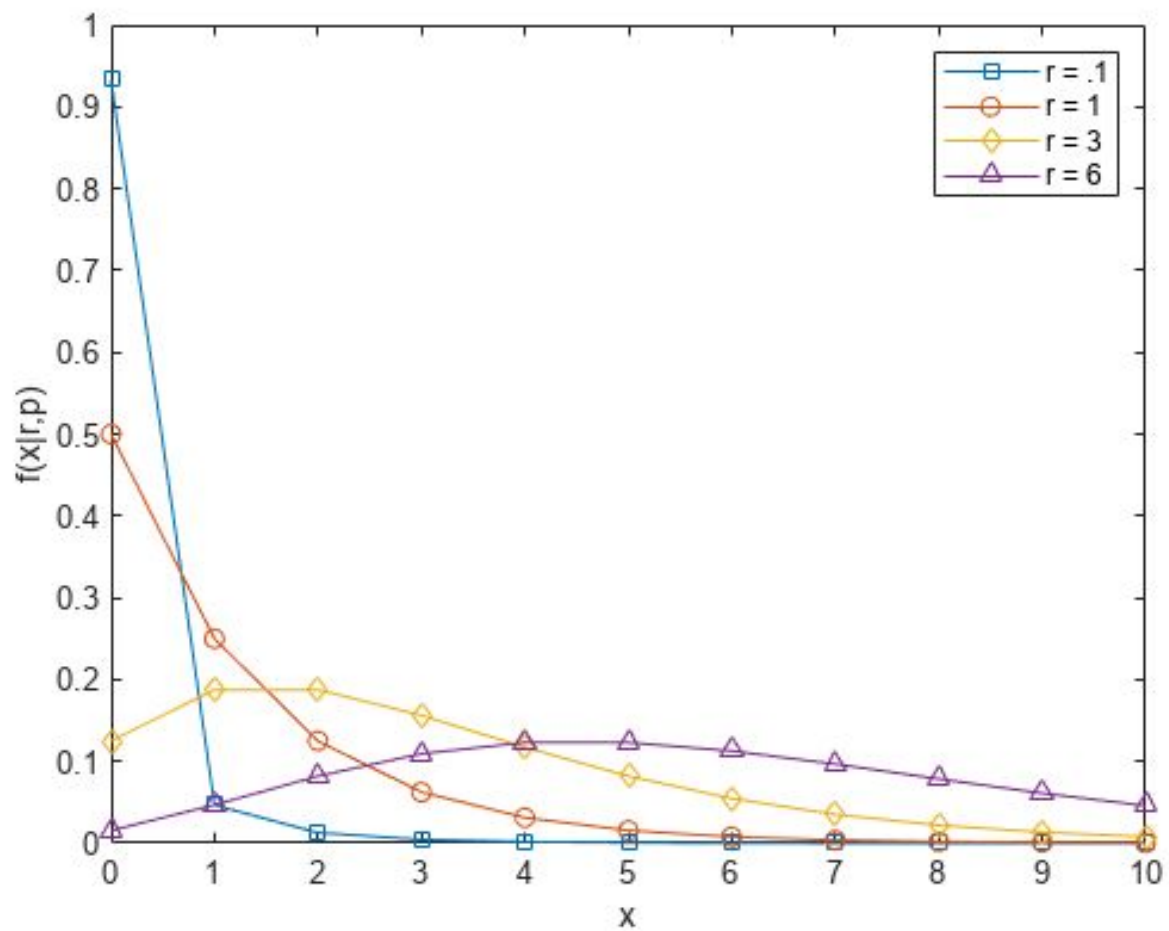
1. Sea X el número de experimentos independientes necesarios para obtener r éxitos, esta variable tiene Distribución Binomial Negativa con parámetros r y p .

$$p_Y(k) = \binom{k-1}{r-1} (1-p)^{k-r} p^r, \quad k = r, r+1, r+2, \dots,$$

$$E(Y) = \frac{r}{p} \quad Var(Y) = r \frac{(1-p)}{p^2}$$

2. Si X es el número de fracasos observados hasta obtener r éxitos, entonces

$$X = Y - r$$



$p=0.5$

Ejemplo Binomial Negativa $\mathcal{BN}(r, p)$

En una ruleta americana, hay 38 espacios: 18 negros, 18 rojos, y 2 verdes. Suponga que ya hace un tiempo que esta apostando en el casino, y que decide que se va a ir cuando gane 3 apuestas al rojo.

1. ¿Cual es la probabilidad de que deje el casino después de apostar 5 veces al rojo?

Ejemplo Binomial Negativa $\mathcal{BN}(r, p)$

- ▶ La probabilidad de ganar apostando al rojo es $p = 18/38$.
- ▶ Sea X la variable número de apuestas al rojo necesarias para ganar tres veces.
- ▶ ¿Cual es la probabilidad de que deje el casino después de apostar 5 veces al rojo? Esto es, quiero calcular la probabilidad de que $X = 5$.

$$\begin{aligned} p_X(5) &= \binom{5-1}{3-1} (1-p)^{5-3} p^3 \\ &= \binom{4}{2} \left(\frac{20}{38}\right)^2 \left(\frac{18}{38}\right)^3 \approx .1766. \end{aligned}$$