Intro a la Probabilidad y estadistica

Martes y Jueves Aula B17

Dra Ana Georgina Flesia

Experimentos geométricos

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 $\mathcal{G}(p)$

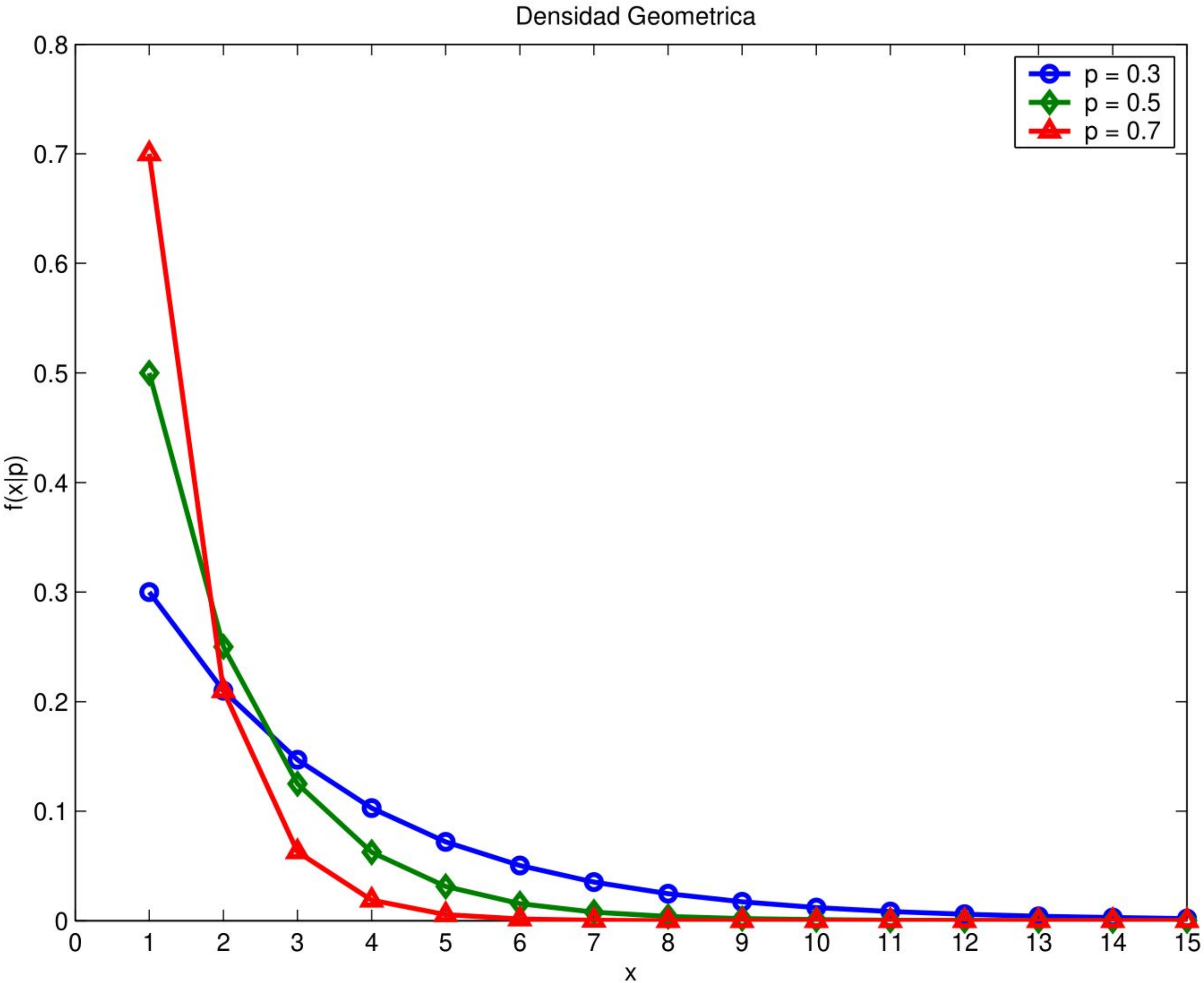
- Sea X el número de experimentos Bernoulli necesarios hasta obtener un éxito.
- X es una variable con infinitos valores posibles 1,2,3,..
- Si X toma el valor k, entonces hay $k-1$ fracasos y 1 éxito en la sucesión de experimentos independientes. Ese punto tiene probabilidad $(1-p)^{k-1}p$. Entonces

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

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

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

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


Densidad Geometrica
p = 0.3
p = 0.5
p = 0.7
f(x|p)
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

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. ¿Cuantos 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 $\mathcal{G}(p)$

- ¿Cuantos 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 arrojamos 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.

$$p_X(k) = p(1-p)^{k-1} \qquad k = 1, 2, \ldots$$

- 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

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

- 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
$$F_X(t) = 1 - (1-p)^{[t]} \qquad t \geq 1, t \in \mathbb{R}$$
2. $P(X \geq k) = (1-p)^{k-1}$ para todo $k \in \mathbb{N}$.
3. La distribución geométrica sufre de amnesia, esto es
$$P(X \geq s + t | X > t) = P(X \geq s)$$

Binomial Negativa

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 $\mathcal{BN}(r, p)$

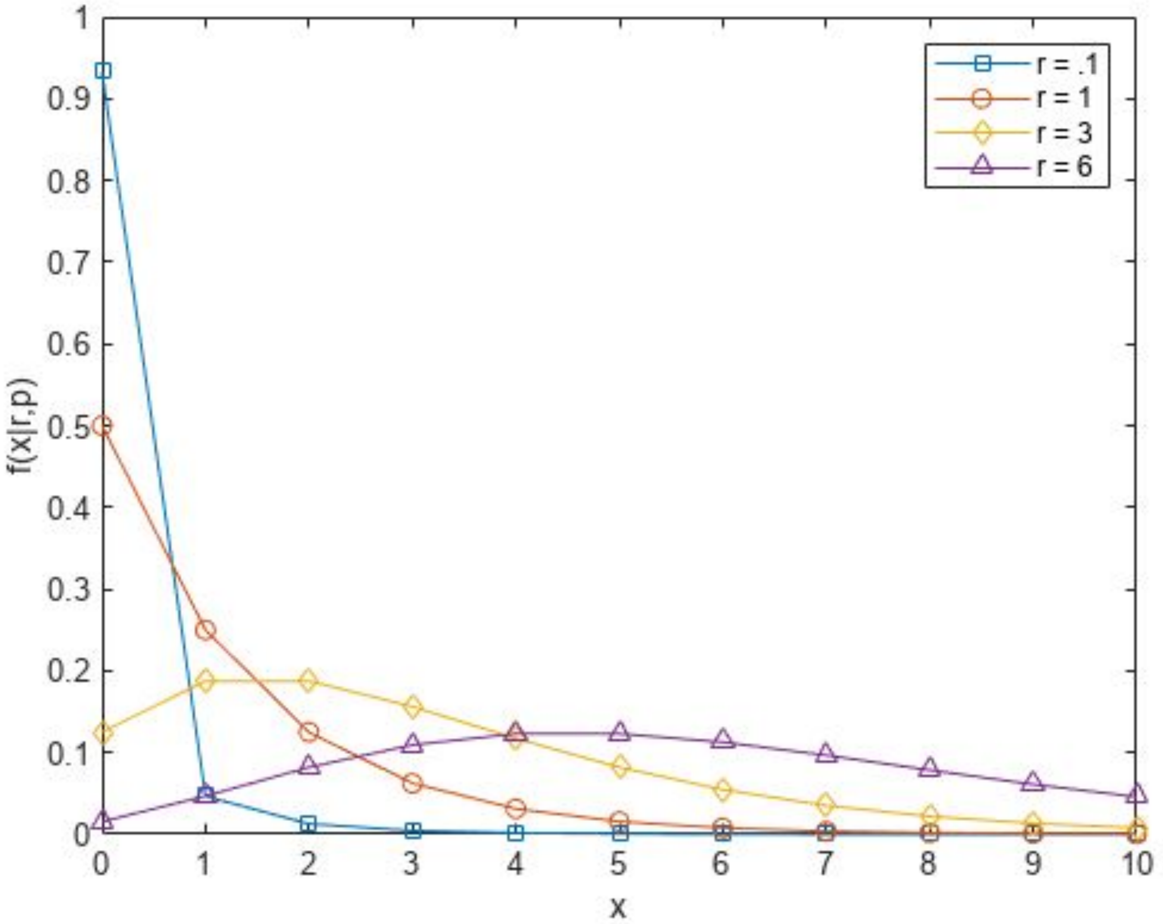
1. Sea X el número de experimentos independientes necesarios para obtener r éxitos, esta variable tiene Distribución Binomial Negativa con parámetros r y p.

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

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

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 $X = 5$.

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

Parametrizaciones y ambigüedades

Binomial Negativa y Pascal

Ciertos autores consideran que

- la distribución Binomial Negativa cuenta el número de fracasos hasta obtener r éxitos
- la distribución de Pascal cuenta el número de experimentos hasta obtener r éxitos

La distribución binomial negativa aparece en un estudio de Pierre Rémond de Montmort (1678-1719) sobre los juegos de azar en 1714, pero años antes ya había sido descrita por Blaise Pascal (1623-1662).

	X is counting...	Probability mass function	Formula	Alternate formula (using equivalent binomial)	Alternate formula (simplified using: $n = k + r$)	Support
1	k failures, given r successes	$f(k; r, p) \equiv \Pr(X = k) =$	$\binom{k+r-1}{k} p^r (1-p)^k$ [8][6][9]	$\binom{k+r-1}{r-1} p^r (1-p)^k$ [3][10][11][12]	$\binom{n-1}{k} p^r (1-p)^k$	for $k = 0, 1, 2, \ldots$
2	n trials, given r successes	$f(n; r, p) \equiv \Pr(X = n) =$	$\binom{n-1}{r-1} p^r (1-p)^{n-r}$ [6][12][13][14][15]	$\binom{n-1}{n-r} p^r (1-p)^{n-r}$		for $n = r, r+1, r+2, \ldots$
3	n trials, given r failures	$f(n; r, p) \equiv \Pr(X = n) =$	$\binom{n-1}{r-1} p^{n-r} (1-p)^r$	$\binom{n-1}{n-r} p^{n-r} (1-p)^r$	$\binom{n-1}{k} p^k (1-p)^r$	
4	k successes, given r failures	$f(k; r, p) \equiv \Pr(X = k) =$	$\binom{k+r-1}{k} p^k (1-p)^r$	$\binom{k+r-1}{r-1} p^k (1-p)^r$		for $k = 0, 1, 2, \ldots$

Ejemplos con resoluciones parciales para completar

Ejemplo 1

El temario de un examen para un proceso selectivo contiene 50 temas, de los cuales se elegirá uno por sorteo. Si una persona no ha estudiado los 15 últimos temas ¿cuál es la probabilidad de que salga un tema que haya estudiado?

Resolución

- La variable que representa el número del tema seleccionado para el examen sigue una distribución uniforme con parámetros a = 1 y b = 50.
- La persona ha estudiado los temas del 1 al 35; por tanto, la probabilidad que se pide es la cola a la izquierda de 35.
- La persona tiene una probabilidad del 70% de que el tema elegido sea uno de los que haya estudiado.

Ejemplo 2

En un examen formado por 20 preguntas, cada una de las cuales se responde declarando “verdadero” o “falso”, el alumno sabe que, históricamente, en el 75% de los casos la respuesta correcta es “verdadero” y decide responder al examen tirando dos monedas: pone “falso” si ambas monedas muestran una cara y “verdadero” si al menos hay una cruz. Se desea saber cual es la probabilidad de que tenga más de 14 aciertos.

Resolución

- X es el número de aciertos, y es una variable Binomial con n = 20, p = 0,75 y el punto a considerar es k = 14.
- La probabilidad de que el alumno tenga más de 14 aciertos es del 62%.

Ejemplo 3

Se sabe que el 7% de los útiles quirúrgicos en un lote de 100 no cumplen ciertas especificaciones de calidad.

Tomada una muestra al azar de 10 unidades sin reemplazo, interesa conocer la probabilidad de que no más de dos sean defectuosas

Resolución

- El número de útiles defectuosos en el lote es $R = 0{,}07 \times 100 = 7$.
- La variable número de defectuosos es hipergeométrica con parámetros (100, 7, 10)
- Para un tamaño muestral de n= 10, la probabilidad buscada es

$$P\{\text{número de defectuosos } \leq 2\}.$$

- La probabilidad de que, a lo sumo, haya dos útiles defectuosos en el lote es aproximadamente 0,98

Ejemplo 4

La probabilidad de que cierto examen médico dé lugar a una reacción “positiva” es igual a 0.8, ¿cuál es la probabilidad de que ocurran menos de 5 reacciones “negativas” antes de la primera positiva?

Resolución

La variable aleatoria “número de experimentos necesarios hasta obtener un positivo” sigue una distribución geométrica con parámetro p = 0,8.

Y si se quieren tener 5 experimentos negativos o menos, estoy buscando

$$P(\text{número de experimentos necesarios} \leq 6)$$

La probabilidad de que ocurran menos de 5 reacciones “negativas” antes de la primera positiva es casi 1 (0,9997).

Ejemplo 5

Se sabe que, en promedio, una de cada 100 placas de rayos X que se realizan es defectuosa. ¿Cuál es el número medio de placas útiles que se producen entre 10 defectuosas?

Resolución

- Si se considera el primer fallo como punto de inicio, hay que considerar la variable “número de placas hasta contar 9 defectuosas”, que sigue una distribución binomial negativa de parámetros $r = 9$ y $p = 0{,}01$.
- La esperanza de una binomial negativa es r/p por lo cual el número esperado de placas útiles es $(r/p)-9$