Distribución Geométrica

Bosquejo: Otra forma de visualizar esto es pensando que se "recorre el valor más bajo posible de X".

En el primer ejemplo numérico es como si desechará el 1 como origen y pusiera mi origen en 2 (que ahora sería 1). *Esto equivale a decir que si están haciendo inspecciones cada hora para verificar la vida útil de algún componente, el hecho de saber que no ha fallado en la observación 100, por ejemplo, (esto es, después de 100 horas de uso) no nos dice nada para responder a la pregunta: ¿cuál será la probabilidad de que falle en las siguientes 20 horas? La probabilidad de que falle es la misma que cuando el componente es nuevo y te haces la misma pregunta: ¿cuál será la probabilidad de que falle en las siguientes 20 horas?*

Esto sólo tiene sentido si se está trabajando con un proceso que se encuentra en estado de estabilidad.

Distribución Geométrica

Media y Varianza

Ya que

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

el valor esperado de $X \sim \text{Geom}(p)$ es

$$\begin{aligned} \mathrm{E}(X) &= \sum_{x=1}^{+\infty} x \cdot f(x) = \sum_{x=1}^{+\infty} x(\underbrace{1-p}_{q})^{x-1} p = p \sum_{x=1}^{+\infty} x q^{x-1} \\ &= p\left(\frac{1}{(1-q)^2}\right) = p\left(\frac{1}{p^2}\right) = \frac{1}{p}. \end{aligned}$$

Distribución Geométrica

Para el cálculo de la varianza, nos queda

$$
\begin{aligned}
\mathrm{V}(X) &= \mathrm{E}(X^2) - [\mathrm{E}(X)]^2, \\
\mathrm{E}(X^2) &= \mathrm{E}(X^2 - X + X) = \mathrm{E}(X^2 - X) + \mathrm{E}(X) \\
&= \mathrm{E}(X(X-1)) + \mathrm{E}(X),
\end{aligned}
$$

pero

$$
\begin{aligned}
\mathrm{E}(X(X-1)) &= \sum_{x=1}^{+\infty} x(x-1)\cdot f(x) = \sum_{x=1}^{+\infty} x(x-1)q^{x-1}p \\
&= p\sum_{x=1}^{+\infty} x(x-1)q^{x-1} = pq\sum_{x=1}^{+\infty} x(x-1)q^{x-2} \\
&= pq\left(\frac{2}{(1-q)^3}\right) = pq\left(\frac{2}{p^3}\right) = \frac{2(1-2p)}{p^2}.
\end{aligned}
$$

Distribución Geométrica

La cantidad $E(X(X-1))$ recibe el nombre de **Segundo Momento Factorial**, de aquí tenemos que

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

y entonces

$$V(X) = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{2-p-1}{p^2} = \frac{1-p}{p^2}.$$

Distribución Geométrica

Función Generatriz de Momentos:

$$M_X(t) = \mathrm{E}(e^{tX}) = \sum_{x=1}^{+\infty} e^{tx} \cdot f(x) = \sum_{x=1}^{+\infty} e^{tx} \cdot q^{x-1} p = p \sum_{x=1}^{+\infty} (e^t)^x \cdot q^{x-1}$$

$$= pe^t \sum_{x=1}^{+\infty} (e^t)^{x-1} \cdot q^{x-1} = pe^t \sum_{x=1}^{+\infty} (e^t q)^{x-1} \underset{(j=x-1)}{=} pe^t \sum_{j=0}^{+\infty} (e^t q)^j,$$

usando la identidad $\sum_{n=0}^{+\infty} z^n = \frac{1}{1-z}$, para $|z| < 1$, nos queda

$$M_X(t) = pe^t \left(\frac{1}{1 - qe^t} \right) = \frac{pe^t}{1 - qe^t} \qquad \text{si } \ |qe^t| < 1.$$

Distribución Geométrica

Ejemplo: Una tienda ofrece un premio a las personas que logren juntar las cinco letras de la palabra "valor"; para ello, se les regala una letra al azar de entre las cinco en cada compra que realicen. Lo que interesa aquí es la cantidad de viajes necesarios (compras) para juntar las cinco letras distintas.

Se podría pensar que la v.a. de interés, $X = \#$ de viajes necesarios hasta juntar las cinco letras de la palabra "valor", es geométrica ya que:

- cada resultado es independiente del anterior,
- el número posible de viajes, realizados entre cada consecución de letras distintas, se puede considerar como "infinito".

Distribución Geométrica

Sin embargo, la probabilidad de un "éxito" cada vez que compre irá cambiando: la primera vez la probabilidad de juntar una letra es 1; la segunda vez que compre ya tendrá en su poder una letra, con lo que la probabilidad de "éxito" cambiaría de 1 a la probabilidad de que reciba una letra distinta de la que se le dio inicialmente y esta sería $4/5$ (4 favorables de 5 posibles); y así seguiría hasta obtener otro "éxito" (letra distinta), con lo que p cambiaría a $3/5$, etc.

Entonces, se puede pensar que en cada viaje la probabilidad es mantenida fija hasta que haya "éxito", y juntando esto con los dos supuestos anteriores, se tendrían v.a.'s geométricas (exceptuando la primera), con p's distintas, cada vez que se lograra un "éxito". Se puede visualizar lo anterior usando una recta numérica:

éxito	viajes	éxito	viajes	éxito	viajes	éxito	viajes	éxito
$\bullet$	$\cdots$	$\bullet$	$\cdots$	$\bullet$	$\cdots$	$\bullet$	$\cdots$	$\bullet$
X_1	$X_2 \sim \text{Geom}(p_2)$		$X_3 \sim \text{Geom}(p_3)$		$X_4 \sim \text{Geom}(p_4)$		$X_5 \sim \text{Geom}(p_5)$	

Distribución Geométrica

Así, nuestra v.a. de interés podría considerarse como la suma de todos los viajes a la tienda, esto es, la suma de cada una de las v.a.'s geométricas definidas:

$$X = X_1 + X_2 + X_3 + X_4 + X_5,$$

donde $X_i \sim$Geom(p_i) $i = 2, 3, 4, 5$, con $p_2 = 4/5$, $p_3 = 3/5$, $p_4 = 2/5$, $p_5 = 1/5$. Entonces, podemos encontrar el **número esperado de viajes** calculando

$$\begin{aligned} \mathrm{E}(X) =& \ \mathrm{E}\left(\sum_{i=1}^{5} X_i\right) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \mathrm{E}(X_3) + \mathrm{E}(X_4) + \mathrm{E}(X_5) \\ =& \ 1 + \frac{1}{p_2} + \frac{1}{p_3} + \frac{1}{p_4} + \frac{1}{p_5} = 1 + \frac{1}{4/5} + \frac{1}{3/5} + \frac{1}{2/5} + \frac{1}{1/5} \\ =& \ 11.42 \text{ viajes}, \end{aligned}$$

esto es, entre 11 ó 12 viajes.

Distribución Geométrica

Ejemplo: Para el Sorteo Tec, si cada vez se expenden 150,000 boletos y 4 automóviles como premio, ¿en cuántos sorteos en promedio hay que participar hasta sacarse un automóvil? ¿Cuánto hay que invertir en promedio si cada boleto cuesta 1,050 pesos?

Como p es fija, hay independencia y "podemos" comprar boletos por siempre, entonces con $X = \#$ de participaciones hasta sacarse un automóvil, $\mathrm{E}(X) = \frac{150{,}000}{4} = 37,500$. Además, habría que invertir $1,050 \times 37,500 = 39,375,000$ pesos.

Con dos rifas por año, uno esperaría ganarse un auto en unos ¡18,750 años! Cuántos años tendrías que participar en promedio si quieres ganarte la casa del primer premio?

Distribución Binomial Negativa

Considera nuevamente el ejemplo del premio al juntar las letras de la palabra "valor", pero suponte que las probabilidades de éxito p_i se mantienen constantes (por ejemplo pensando que el premio se da a la persona que junte 5 letras "V", con lo que p_i siempre sería $1/5$). Como antes, la variable aleatoria $X =$ **# de viajes necesarios hasta obtener 5 "éxitos"**, entonces

$$X = \sum_{i=1}^{5} X_i, \quad \text{donde } X_i \sim \text{Geom}(p) \text{ son independientes y } p \text{ constante.}$$

Esto es equivalente al segundo criterio mencionado para la detención o ajuste de un proceso (ver diapositiva 56): hasta obtener r (por ejemplo 3) artículos defectuosos (debido tal vez a lo costoso de parar la línea al obtener el primer defectuoso), en donde

Distribución Binomial Negativa

- cada éxito tiene la misma probabilidad p, ya que se trata del mismo proceso y se asume que éste es estable (no hay fallas sistemáticas),
- cada X_i es independiente;

en este caso la variable aleatoria de interés es

$$X = \ \# \text{ de observaciones hasta encontrar } r \text{ ``éxitos''},$$

la cual se puede redefinir como

$$X = \sum_{i=1}^{r} X_i, \text{ donde cada } X_i \sim \mathrm{Geom}(p) \text{ es independiente y p constante.}$$

La v.a. X definida como la suma anterior se dice que sigue una distribución Binomial Negativa: $X \sim \mathrm{BN}(r, p)$. Aquí, inclusive $X_1 \sim \mathrm{Geom}(p)$.

Distribución Binomial Negativa

A continuación se muestra un dibujo similar al hecho para la distribución geométrica (ver diapositiva 76), aquí obs = observaciones y Geo=Geom:

obs	éxito 1	obs	éxito 2	$\cdots$	obs	éxito r-1	obs	éxito r
$\cdots$	$\bullet$	$\cdots$	$\bullet$	$\cdots$	$\cdots$	$\bullet$	$\cdots$	$\bullet$
$X_1 \sim$Geo(p)		$X_2 \sim$Geo(p)		$\cdots$	$X_{r-1} \sim$ Geo(p)		$X_r \sim$Geo(p)	

Media y Varianza

El valor esperado de la v.a. $X \sim \mathrm{BN}(r, p)$ es

$$\mathrm{E}(X) = \mathrm{E}\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} \mathrm{E}(X_i) = \sum_{i=1}^{r} \frac{1}{p} = \frac{r}{p};$$

para la varianza usamos el hecho de que las X_i's son independientes

$$\mathrm{V}(X) = \mathrm{V}\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} \mathrm{V}(X_i) = \sum_{i=1}^{r} \frac{1-p}{p^2} = r\left(\frac{1-p}{p^2}\right).$$

Distribución Binomial Negativa

Función Generatriz de Momentos

Usando la estructura de suma se tiene que la función generatriz de momentos de una distribución BN(r,p) está dada por

$$\begin{aligned} M_X(t) = \; & \mathrm{E}(e^{tX}) = \mathrm{E}\left(e^{t\sum_{i=1}^{r} X_i}\right) = \mathrm{E}\left(e^{tX_1+tX_2+\cdots+tX_r}\right) \\ =& \mathrm{E}(e^{tX_1})\,\mathrm{E}(e^{tX_2})\,\cdots\,\mathrm{E}(e^{tX_r}) \\ =& \left(\frac{pe^t}{1-qe^t}\right)\left(\frac{pe^t}{1-qe^t}\right)\cdots\left(\frac{pe^t}{1-qe^t}\right) = \left(\frac{pe^t}{1-qe^t}\right)^r, \end{aligned}$$

para $|qe^t| < 1$.

Distribución Binomial Negativa

Recalquemos los supuestos:

- Cada vez que se realice una observación la probabilidad de que ocurra un éxito es fija e igual al valor p:
  Nuevamente, el proceso "no sabe" o "no se acuerda" de lo que ha pasado antes de la observación que se esté realizando. Cuando uno detecte artículos defectuosos, sabiendo la probabilidad que el modelo asigna a ese caso, se toma una decisión respecto al proceso. Ver[15].
- Las observaciones son independientes.
  Como antes, el que una observación haya sido o no "éxito", no da información sobre el resultado de la siguiente observación.
- El primer posible valor de la variable aleatoria es r.
  Puesto que la condición es encontrar r defectuosos, se tienen que realizar mínimo r pruebas.
- El número de observaciones puede continuar indefinidamente antes de encontrar el r-ésimo defectuoso.

Distribución Binomial Negativa

Una forma de obtener explícitamente a la función de probabilidad de la v.a. Binomial Negativa se da a continuación.

El evento de que el r-ésimo "éxito" ocurra en $X = x$ equivale a pensar que en las $x-1$ observaciones anteriores hay $r-1$ "éxitos", los cuales pudieron ocurrir en cualquiera de las anteriores $x-1$ observaciones. Si agrupamos en casillas las observaciones, lo anterior puede visualizarse de la siguiente manera

$$\overbrace{\underbrace{|\_\_\_\_|\_\_\_\_|\_\_\_\_|\cdots\cdots|\_\_\_\_|}_{x-1\text{ casillas}}}^{r-1\text{ éxitos}} \qquad \overbrace{\underbrace{|\_\_\_\_|}_{\text{casilla }\#x}}^{r\text{-ésimo éxito}}$$

Distribución Binomial Negativa

Debido a la independencia sólo se tienen que multiplicar las probabilidades de cada casilla para obtener la probabilidad de r "éxitos" en x pruebas:

- En la última prueba hay "éxito" con probabilidad p.
- Los $r-1$ "éxitos" anteriores pueden ocurrir en cualquier orden. El número de maneras en que $r-1$ "éxitos" se pueden acomodar en $x-1$ casillas está dado por $\binom{x-1}{r-1}$, y cada uno de ellos tiene asignada una probabilidad p.
- Hay $(x-1)-(r-1)=x-r$ fracasos cada uno con probabilidad $(1-p)$.

Distribución Binomial Negativa

Entonces, la probabilidad de obtener r éxitos en x pruebas es

$$\begin{aligned} \mathrm{P}(X = x) &= p\left[\binom{x-1}{r-1} p^{r-1}\right](1-p)^{x-r} \\ &= \binom{x-1}{r-1}(1-p)^{x-r}p^{r}, \end{aligned}$$

con $x = r, r+1, r+2, \ldots$

Esta es la Distribución Binomial Negativa cuyos parámetros son r y p. Notemos que n sigue siendo aleatorio.

Distribución Binomial Negativa

Relación entre una v.a. $X \sim \mathbf{B}(n, p)$ con una v.a. $Y \sim \mathbf{BN}(r, p)$

Recuerda que

$X =$ # de éxitos en n pruebas Bernoulli independientes,

$Y =$ # de observaciones hasta tener r éxitos en pruebas Bernoulli independientes.

Consideremos la probabilidad de que en los n intentos se observen a lo más $r-1$ "éxitos", esto es,

$$\mathrm{P}(X \leq r-1).$$

Observemos que en los n intentos hay a lo más $r-1$ éxitos si y sólo si en los n intentos se observan menos de r éxitos, es decir que la probabilidad anterior es igual a la probabilidad

$$\mathrm{P}(Y > n).$$

Distribución Binomial Negativa

Entonces

$$P(X \leq r-1) = P(Y > n) = 1 - P(Y \leq n).$$

Cabe recalcar que esto **no** es una aproximación, es una identidad.

Ejemplo: Si un proceso nos da dos equipos con defectos "severos" se detiene la producción. La probabilidad de ese tipo de defectos es 0.05. ¿Cuál es la probabilidad de hacer más de 10 equipos sin detener la producción?

La variable aleatoria es $Y = \#$ de equipos revisados hasta obtener 2 con defectos "severos".

Distribución Binomial Negativa

Supuestos:

- probabilidad fija e igual a p (no hay fallas sistemáticas)
- resultados independientes

Entonces $Y \sim \mathrm{BN}(r = 2, p = 0.05)$, de donde

$$P(Y > 10) = 1 - P(Y \leq 10) = 1 - \sum_{y=2}^{10} \binom{y-1}{2-1} (0.05)^2 (0.95)^{y-2} = 0.914.$$

Utilizando la relación con la distribución binomial, para $X \sim \mathrm{B}(n = 10, p = 0.05)$ tenemos que

$$P(Y > 10) = P(X \leq 2 - 1) = \sum_{x=0}^{1} \binom{10}{x} (0.05)^x (0.95)^{10-x} = 0.914.$$

Distribución Binomial Negativa

Los cálculos de la generatriz de momentos, media y varianza se pueden obtener a partir de la función de probabilidad de la binomial negativa, aunque de hecho ya lo hicimos mediante la representación en sumas de la distribución binomial negativa. Por completez se incluye el siguiente este material.

Deduciremos la generatriz de momentos para obtener la media y la varianza de $X \sim \text{BN}(r, p)$. Para calcular la generatriz de momentos se utilizarán los siguientes resultados:

$$\binom{n}{z} = \binom{n}{n-z}, \tag{1}$$

$$(1-x)^{-r} = \sum_{j=0}^{+\infty} \binom{j+r-1}{j} x^j, \quad \text{para } |x| < 1. \tag{2}$$

Distribución Binomial Negativa

La generatriz de momentos es entonces

$$
\begin{aligned}
M_X(t) = \mathrm{E}(e^{tX}) &= \sum_{x=r}^{+\infty} e^{tx} \cdot f(x) = \sum_{x=r}^{+\infty} e^{tx} \binom{x-1}{r-1} q^{x-r} p^r \\
&= p^r \sum_{x=r}^{+\infty} e^{tx} \binom{x-1}{r-1} q^{x-r} = p^r (e^t)^r \sum_{x=r}^{+\infty} \binom{x-1}{r-1} (qe^t)^{x-r} \\
&= p^r (e^t)^r \sum_{j=0}^{+\infty} \binom{j+r-1}{r-1} (qe^t)^{x-r} \qquad j = x - r \\
&= (pe^t)^r \sum_{j=0}^{+\infty} \binom{j+r-1}{(j+r-1)-(r-1)} (qe^t)^{x-r} \qquad \text{de (1)} \\
&= (pe^t)^r \sum_{j=0}^{+\infty} \binom{j+r-1}{j} (qe^t)^{x-r} \\
&\overset{\text{de (2)}}{=} (pe^t)^r (1-qe^t)^{-r} = \frac{(pe^t)^r}{(1-qe^t)^r} = \left(\frac{pe^t}{1-qe^t} \right)^r, \qquad |qe^t| < 1.
\end{aligned}
$$

Distribución Binomial Negativa

Por lo tanto

$$M_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^r, \quad \text{para } |qe^t| < 1.$$

Cuando $r = 1$ la binomial negativa es equivalente a la Geométrica; es decir, la binomial negativa se puede considerar como una generalización de la geométrica. Así,

$$\mu = \mu_1' = \mathrm{E}(X) = M_X'(t)|_{t=0},$$

$$\begin{aligned} M_X'(t) &= r\left(\frac{pe^t}{1-qe^t}\right)^{r-1} \frac{(1-qe^t)pe^t - pe^t(-qe^t)}{(1-qe^t)^2} \\ &= r\left(\frac{pe^t}{1-qe^t}\right)^{r-1} \frac{pe^t - pqe^{2t} + pqe^{2t}}{(1-qe^t)^2} = r\left(\frac{pe^t}{1-qe^t}\right)^{r-1} \frac{pe^t}{(1-qe^t)^2}, \\ &= \frac{r(pe^t)^r}{(1-qe^t)^{r+1}}. \\ M_X'(0) &= \frac{r(pe^0)^r}{(1-qe^0)^{r+1}} = \frac{rp^r}{(1-q)^{r+1}} = \frac{rp^r}{p^{r+1}} = \frac{r}{p}. \end{aligned}$$

Distribución Binomial Negativa

Por lo tanto,

$$\mathrm{E}(X) = r/p.$$

Por otra parte,

$$\mu_2' = \mathrm{E}(X^2) = M_X''(t)|_{t=0},$$

$$M_X''(t) = \frac{(1-qe^t)^{r+1} r \cdot r(pe^t)^{r-1}pe^t - r(pe^t)^r(r+1)(1-qe^t)^r(-qe^t)}{[(1-qe^t)^{r+1}]^2},$$

$$M_X''(0) = \frac{(1-q)^{r+1}r^2p^{r-1}p - rp^r(r+1)(1-q)^r(-q)}{(1-q)^{2(r+1)}}$$

$$= \frac{p^{r+1}r^2p^r + r(r+1)p^rp^rq}{p^{2r+2}}$$

$$= \frac{p^{2r+1}r^2 + r(r+1)p^{2r}q}{p^{2r+2}} = \frac{rp^{2r}(rp+(r+1)q)}{p^{2r+2}}$$

$$= \frac{r\left(\overbrace{rp+rq}^{r}+q\right)}{p^2} = \frac{r(r+q)}{p^2}.$$

Distribución Binomial Negativa

Así que

$$\mathrm{E}(X^2) = \frac{r(r+q)}{p^2}.$$

Entonces,

$$\begin{aligned}
\mathrm{V}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 &= \frac{r(r+q)}{p^2} - \left(\frac{r}{p}\right)^2 \\
&= \frac{r(r+q) - r^2}{p^2} = \frac{r^2 + rq - r^2}{p^2} = \frac{rq}{p^2} = \frac{r(1-p)}{p^2},
\end{aligned}$$

lo que implica

$$\mathrm{V}(X) = \frac{r(1-p)}{p^2}.$$

Distribución Binomial Negativa

A modo de resumen, notemos lo siguiente:

- Notemos que en un contexto teórico el valor de r puede ser distinto de un número entero.
- Además, la variable aleatoria BINOMIAL, cuando $n = 1$, nos da una variable aleatoria que tiene distribución Bernoulli; es decir, la Bernoulli es un caso particular de la binomial.
- Similarmente, una variable aleatoria BINOMIAL NEGATIVA, con $r = 1$, nos da una variable aleatoria con distribución GEOMÉTRICA; es decir, la geométrica es un caso particular de la binomial negativa.

Distribución Poisson

Supóngase ahora que estamos interesado en el número de "éxitos", pero ahora preguntándonos por la posibilidad de que ocurran en un cierto intervalo de tiempo o espacio, por ejemplo

- Número de defectos en una plancha de acero de $1\,m^2$,
- Número de bacterias en $1\,cm^3$ de agua potable,
- Número de entregas de materia prima entre las 8:00 A.M. y la 1:00 P.M.,
- Número de llamadas que llegan a un conmutador en un minuto.

En estos casos interesa la variable aleatoria

$X = \#$ de "éxitos" obtenidos en un cierto intervalo (temporal o espacial).