Viernes 17 de septiembre de 2021.

2.5 Distribución Geométrica.

La distribución geométrica está relacionada con una secuencia de ensayos Bernoulli (independientes) con misma probabilidad de éxito pero el número de ensayos no es fijo. La variable aleatoria de interés X, es el número de ensayos necesarios hasta obtener el primer éxito. Los valores posibles que toma la variable aleatoria son $x = 1, 2, 3, \ldots$. Se dice que X tiene una distribución geométrica con parámetro p (probabilidad de éxito) lo cual se escribe como $X \sim G(x;p)$ si su f.d.p es de la forma:

$$\boxed{G(x;p) = P(X = x) = q^{x-1}p} \text{ con } x = 1, 2, \ldots$$

La f.g.m. para esta distribución es:

$$\psi_X(t) = E\left(e^{tX}\right) = \sum_{x=1}^{\infty} e^{tx} pq^{x-1} = p\left(e^t + e^{2t}q + e^{3t}q^2 + e^{4t}q^3 + \cdots\right) = pe^t\left(1 + \underbrace{e^t q + e^{2t}q^2 + e^{3t}q^3 + \cdots}\right)$$

$qe^t \quad (qe^t)^2 \quad (qe^t)^3 \quad (qe^t)^0$

$$= pe^t \sum_{i=0}^{\infty} e^{it} q^i = pe^t \sum_{i=0}^{\infty} \left(e^t q\right)^i = pe^t\left[\frac{1}{1 - qe^t}\right]$$

$p + q = 1$

$$\therefore \boxed{\psi_X(t) = \frac{pe^t}{1 - qe^t}}$$

Nota._ Utilizamos el resultado de la serie geométrica $\sum_{i=0}^{\infty} r^i = \frac{1}{1-r}$ converge si $0<|r|<1$

Media y varianza.

$p - pq + pq$

$$E(X) = \psi'_X(t=0) = \left.\frac{\left(1 - qe^t\right)pe^t - pe^t\left(-qe^t\right)}{\left(1 - qe^t\right)^2}\right|_{t=0} = \left.\frac{pe^t - pqe^{2t} + pqe^{2t}}{\left(1 - qe^t\right)^2}\right|_{t=0} = \left.\frac{pe^t}{\left(1 - qe^t\right)^2}\right|_{t=0} = \frac{p}{p^2}$$

$(1-q)^2$

La media de la v.a Geométrica es: $\boxed{E(X) = \frac{1}{p}}$

$$\psi''_X(t=0) = \left.\frac{\left(1 - qe^t\right)^2 pe^t + pe^t\, 2\left(1 - qe^t\right)qe^t}{\left(1 - qe^t\right)^4}\right|_{t=0} = \frac{p^3 + 2p^2q}{p^4} = \frac{p^2(p + 2q)}{p^4} = \frac{p + 2q}{p^2}$$

$-(1-p)+2q$

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

La varianza de la v.a Geométrica es: $\boxed{V(X) = \frac{q}{p^2}}$

Función de probabilidad acumulada.

$$F_X(x) = P(X \le x) = \sum_{i=1}^{x} pq^{i-1} = p\sum_{i=1}^{x} q^{i-1} = p\left[1 + q + q^2 + q^3 + \cdots + q^{x-1}\right] = p\sum_{i=0}^{x-1} q^i = p\left[\frac{1 - q^x}{1 - q}\right]$$

$q^0 \quad p$

La función de probabilidad acumulada de la v.a geométrica es $\boxed{F_X(x) = 1 - q^x}$

Nota._ Utilizamos el resultado de la serie geométrica $\sum_{i=0}^{x} r^i = \frac{1 - r^{x+1}}{1 - r}$

Ejemplos:

1._ Se va a realizar cierto experimento hasta que se obtenga un resultado exitoso. Los ensayos son independientes y el costo de efectuar el experimento es \$25,000 dólares, sin embargo, si se produce una falla, cuesta \$5,000 dólares iniciar el siguiente ensayo. Al experimentador le gustaría determinar el costo esperado del proyecto. Supón que $p = 0.25$.

Solución

La v.a es

X: Número de ensayos hasta obtener un experimento exitoso.

$X \sim G(x; p = 0.25)$

$C(X)$: Costo de experimento en función de la v.a.

Necesitamos construir el costo del proyecto como una función del número de ensayos requeridos hasta el primer éxito.

$C(X) = 25{,}000X + 5{,}000(X - 1) = 30{,}000X - 5{,}000$

$$E[C(X)] = 30{,}000E(X) - 5{,}000 = 30{,}000\left(\frac{1}{p}\right) - 5{,}000 = 30{,}000\left(\frac{100}{25}\right) - 5{,}000 = \$115{,}000.$$

2._ Una compañía aeroespacial ha construido cinco mísiles. La probabilidad de un disparo exitoso en cualquier prueba es 0.95. Supón lanzamientos independientes ¿Cuál es la probabilidad de que la primer falla ocurra en el quinto disparo?

Solución

El éxito es el evento de interés y en este problema nos interesa que el disparo sea no exitoso entonces:

$p = P(\text{éxito}) = P(\text{disparo no exitoso}) = 0.05$

$q = P(\text{fracaso}) = P(\text{disparo exitoso}) = 0.95$

X: Número de disparos hasta que ocurre la primera falla.

$X \sim G(x; p = 0.05)$

$P(X = 5) = (0.95)^4(0.05) = 0.0407$

3._ La compañía A planea visitar clientes potenciales hasta que se realice una venta considerable. Cada presentación de venta cuesta 1,000 dólares y cuesta 4,000 dólares viajar para visitar al siguiente cliente y realizar una nueva presentación.

a) ¿Cuál es el costo esperado de la realización de una venta si la probabilidad de hacer una venta después de cualquier presentación es 0.10?

b) Si la ganancia esperada en cada venta es 15,000 dólares, ¿Deben efectuarse los viajes?

c) Si el presupuesto para publicidad es sólo 100, 000 dólares ¿Cuál es la probabilidad de que esta suma sea gastada sin que se logre ningún pedido?

Solución

X: Número de visitas hasta realizar la primera venta importante.

$X \sim G(x; p = 0.1)$

Sabemos que $E(X) = \frac{1}{p} = \frac{1}{0.1} = 10$

$C(X)$: Costo de la realización de la primera venta en función de la v.a.

a) El costo en función del número de visitas es $C(X) = 1000X + 4000(X - 1) = 5000X - 4000$

Por lo tanto el costo esperado es $E[C(X)] = 5000E(X) - 4000 = 5000(10) - 4000 = \$46{,}000$.

b) El costo esperado para realizar la primera venta es de \$46,000 > \$15,000 que es la ganancia esperada en esa venta, por lo tanto no deben efectuarse los viajes.

c) El costo para realizar la primera venta está dado por

$C(X) = 5{,}000X - 4{,}000$

Entonces
$$\begin{aligned} P[C(X) > 100{,}000] &= P(5{,}000X - 4{,}000 > 100{,}000) \\ &= P(5{,}000X > 104{,}000) \\ &= P\left(X > \frac{104}{5}\right) = P(X > 20.8) = 1 - P(X \le 20) \\ &= 1 - F_X(20) = 1 - \left(1 - q^{20}\right) = (0.9)^{20} = 0.1216 \end{aligned}$$

20.8 20 21

4._ Dado que se ha lanzado una moneda normal 10 veces, y se han obtenido cero caras. ¿Cuál es la probabilidad de que se tenga que lanzar al menos dos veces más para obtener la primera cara?

Solución

X : Número de lanzamientos para obtener la primera cara.

$X \sim G(x; p = 0.5)$

$X \le 10 \quad X > 10 \quad X \le 11 \quad X \ge 12$

Entonces

$$P(X \ge 12 | X > 10) = \frac{P(X \ge 12,\ X > 10)}{P(X > 10)} = \frac{P(X \ge 12)}{P(X > 10)} = \frac{1 - F_X(11)}{1 - F_X(10)} = \frac{q^{11}}{q^{10}} = q = \frac{1}{2}$$

$P(X \ge 2) = 1 - F_X(1) = 1 - (1 - q^1) = q^1 = \frac{1}{2}$

Propiedad._ Si X tiene una distribución geométrica, para dos enteros positivos cualesquiera s y t, se tiene que

$$\boxed{P(X \ge s + t \mid X > s) = P(X \ge t)}$$

Esta propiedad se conoce como la propiedad de la falta de memorio o pérdida de la memoria.

Demostración:

$t = 2 \quad s \quad s+t \quad 10 \quad 12$

Observa que:
$$\begin{aligned} P(X \ge t) &= 1 - P(X \le t - 1) \\ &= 1 - F_X(t - 1) \\ &= 1 - \left(1 - q^{t-1}\right) = q^{t-1} \end{aligned}$$

Y por otro lado

$$P(X \ge s + t \mid X > s) = \frac{P(X > s, X \ge s + t)}{P(X > s)} = \frac{P(X \ge s + t)}{1 - P(X \le s)} = \frac{1 - P(X \le s + t - 1)}{1 - F_X(s)}$$

$$= \frac{1 - F_X(s + t - 1)}{1 - F_X(s)} = \frac{1 - \left(1 - q^{s+t-1}\right)}{1 - \left(1 - q^s\right)} = \frac{q^{s+t-1}}{q^s} = q^{t-1} = P(X \ge t)$$

Así por ejemplo el problema de las 10 cruces pudo resolverse usando esta propiedad de la pérdida de la memoria;

$$P(X \ge 12 \mid X > 10) = P(X \ge 2) = 1 - F_X(1) = 1 - (1 - q) = \frac{1}{2}$$

5._ Sea Y una v.a geométrica con probabilidad de éxito igual a p, demuestra que:

a) para un entero positivo a , $P(Y > a) = q^a$

b) para los enteros positivos a y b $P(Y > a + b \mid Y > a) = P(Y > b)$

Solución

a) $P(Y > a) = 1 - P(Y \le a) = 1 - F_Y(a) = 1 - \left(1 - q^a\right) = q^a$.

b) $P(Y > a + b \mid Y > a) = \frac{P(Y > a + b)}{P(Y > a)} = \frac{q^{a+b}}{q^a} = q^b = P(Y > b)$.

6._ Un contador público ha encontrado que 9 de 10 auditorias de compañías contienen errores importantes. Si el contador revisa la contabilidad de una serie de compañías.

a) ¿Cuál es la probabilidad de que la primera contabilidad con errores importantes se encuentre en la tercera contabilidad revisada?

b) ¿Cuál es la probabilidad de que la primera contabilidad con errores importantes se encuentre después de revisar la tercera?

c) Encuentra la media y la desviación estándar del número de contabilidades que hay que revisar para obtener la primera con errores importantes?

Solución

X: Número de contabilidades revisadas para obtener el primer error importante.

$X \sim G\left(x; p = \frac{9}{10}\right)$

a) $P(X = 3) = \left(\frac{1}{10}\right)^2\left(\frac{9}{10}\right) = \frac{9}{1000} = 0.009$

b) $P(X > 3) = 1 - F_X(3) = q^3 = \left(\frac{1}{10}\right)^3 = 0.001$

c) $E(X) = \frac{1}{p} = \frac{10}{9} = 1.1$

$V(X) = \frac{q}{p^2} = \frac{1}{10}\left(\frac{100}{81}\right) = \frac{10}{81} = 0.123$

Sacando la raíz positiva de la varianza obtenemos la desviación estándar, es decir $\sigma = \sqrt{V(X)} = 0.35$.

Binomial X : # de éxitos en n Ensayos Bernoulli independientes con prob. de éxito constante.

X : # de éxitos en un experimento binomial.

$x = 0, 1, 2, \ldots, n$

$X \sim b(x; n, p)$

$\psi_X(t) = (q + pe^t)^n$

$f_X(x) = \binom{n}{x} p^x q^{n-x}$

$\mu = np$

$\sigma^2 = npq$

$F_X(x) = \sum_{i=0}^{x} \binom{n}{i} p^i q^{n-i}$

Pensemos en el experimento : Lanzar una moneda con $P(\text{cara}) = p$ hasta la 1ª cara.

X : # de lanzamientos hasta la 1ª cara inclusive.

Eventos	$X = x$	
C	1	$P(X = 1) = P(C) = p$
X C	2	$P(X = 2) = P(XC) = P(X)P(C) = qp$
X X C	3	$P(X = 3) = P(XXC) = P(X)P(X)P(C) = q^2p$
X X X C	4	$P(X = 4)$ q^3p
X X X X C	5	
	⋮	

$P(X = x) = q^{x-1}p$

Geométrica

$X \sim G(x; p)$

$f_X(x) = q^{x-1}p$

$\mu = \frac{1}{p}$

$\sigma^2 = \frac{q}{p^2}$

$\psi_X(t) = \frac{pe^t}{1 - qe^t}$

$F_X(x) = 1 - q^x = P(X \le x)$

$P(X > x) = P(X \ge x + 1) = q^x$

# de experimentos	Costo	
1	25,000	$= (1)25{,}000 + 0(5{,}000)$
2	25,000 + 5,000 + 25,000	$= 2(25{,}000) + (1)5{,}000$
3	25,000 + 5,000 + 25,000 + 5,000 + 25,000	$= 3(25{,}000) + 2(5{,}000)$
⋮		
X		$= X(25{,}000) + (x-1)(5{,}000)$

$C(X) = 25{,}000X + 5{,}000(X - 1)$

$= 25{,}000X + 5{,}000X - 5{,}000$

$C(X) = 30{,}000X - 5{,}000$

$E(C(X)) = E(30{,}000X - 5{,}000)$

$= 30{,}000E(X) - 5{,}000$

$= 30{,}000\left(\frac{1}{0.25}\right) - 5{,}000$

$= 120{,}000 - 5{,}000 = 115{,}000.00$