CAPITULO IV

MODELOS DE DISTRIBUCIÓN DE PROBABILIDAD DE VARIABLES ALEATORIAS DISCRETAS

4-1 INTRODUCCIÓN

En las secciones anteriores se ha definido experimento aleatorio, variable aleatoria, distribución de probabilidad y se han desarrollado sus propiedades generales.

Al estudiar los distintos fenómenos aleatorios se empieza a observar que el comportamiento de muchos de ellos son bastantes análogos entre sí. Si se analizan las características de las variables aleatorias, estas se repiten, lo que conllevó a continuar el análisis de las características comunes de estas variables, este estudio llevó a definir modelos probabilísticos particulares que permiten explicar fenómenos aleatorios que tienen una realización similar entre sí.

En este capítulo se estudiará algunos modelos de distribución de probabilidad de variables aleatorias discretas específicas, que resultan ser útiles para la solución de situaciones problemáticas que se enfrenta diariamente el investigador. Los mismos son el resultado matemático de la aplicación de la teoría de la probabilidad a distintos fenómenos aleatorios, esta deducción parte de considerar como ciertas, determinadas premisas.

Es por eso que al seleccionar un determinado modelo para su aplicación a un caso práctico definido, deberá conocerse la naturaleza del fenómeno aleatorio y analizar si las premisas del modelo elegido se cumplen.

Algunos modelos de distribución de probabilidad de variable aleatoria discreta que son muy utilizados son el modelo Bernoulli, el Binomial, el modelo de Poisson, el hipergeométrico, el geométrico, el modelo binomial negativo entre otros. Con relación a los modelos de variable aleatoria continua, algunos que se estudiarán será el modelo normal, uniforme, exponencial, gamma, beta, t de Student, ji-cuadrado, F- Fisher-Snedecor y el modelo Weibull.

Veremos a continuación primeramente los modelos de distribución de probabilidad de una variable discreta y luego en la unidad siguiente los modelos utilizando variables aleatorias continuas.

4-2 MODELOS DE DISTRIBUCIÓN DE PROBABILIDAD DE UNA VARIABLE ALEATORIA DISCRETA

4-2-1 MODELO BERNOULLI

Este modelo se denomina Bernoulli en honor al matemático y científico James Bernoulli, (1654-1705), su obra maestra fue "El arte de la conjetura", un trabajo pionero en la teoría de la probabilidad.

Considere un experimento dicotómico, cuando se habla de experimento dicotómico nos referimos aquellos que sólo tienen dos posibles resultados, clasificados como éxito o fracaso.

Ejemplos 4-1:

El experimento dicotómico: *"El lanzamiento de una moneda insesgada y observar si se obtiene cara o cruz"*, es un proceso de Bernoulli.

El espacio muestral es: Ω= {cara, cruz}

Ejemplos 4-2:

El experimento dicotómico: *"Una persona elegida al azar, que solicita un empleo es entrevistada para una prueba de aptitud, y observar el éxito o fracaso de la entrevista*", también puede ser descripto como un proceso de Bernoulli.

El espacio muestral es: Ω= {éxito, fracaso}

Una variable aleatoria X tiene una distribución de Bernoulli si cumple las siguientes características:

1. La variable aleatoria es discreta.
2. El experimento aleatorio tiene solamente dos posibles resultados, mutuamente excluyentes.
3. La variable aleatoria X que representa el número de éxitos en una prueba. Su función de probabilidad está dada por la ecuación:

$$f_X(x) = \begin{cases} p^x(1-p)^{1-x}, & si \ \ x = 0, 1 \\ 0 & para \ \ otro \ valor \end{cases}$$

   Siendo p: el parámetro de la distribución de Bernoulli. Representa la probabilidad de que el suceso sea un éxito. Los valores que puede tomar son: $0 \leq p \leq 1$
   Su anotación es: X ~ Bernoulli (x, p)

Ejemplo 4-3:

En el ejemplo 4-1 de lanzar una moneda los posibles resultados son: cara o cruz; en el ejemplo 4-2, de la entrevista los posibles resultados son: éxito o fracaso. En general llamaremos éxito a la ocurrencia del evento que estamos estudiando, fracaso a la no ocurrencia.

La variable aleatoria siempre estará enunciada en función del evento éxito.

La variable aleatoria asigna al evento "éxito" el número 1 y al "fracaso" el 0, entonces:

$$X: \Omega \to R;$$

$$X(\text{é}xito) = 1,$$

$$X(fracaso) = 0$$

El recorrido de la variable X es:

$$X(\Omega) = \{0, 1\}$$

La probabilidad de éxito es p, y la probabilidad de fracaso es q

$$P(\{\text{é}xito\}) = P(X = 1) = f_X(1) = p$$

$$P(\{fracaso\}) = P(X = 0) = f_X(0) = q$$

Como el experimento tiene sólo dos posibles resultados entonces:

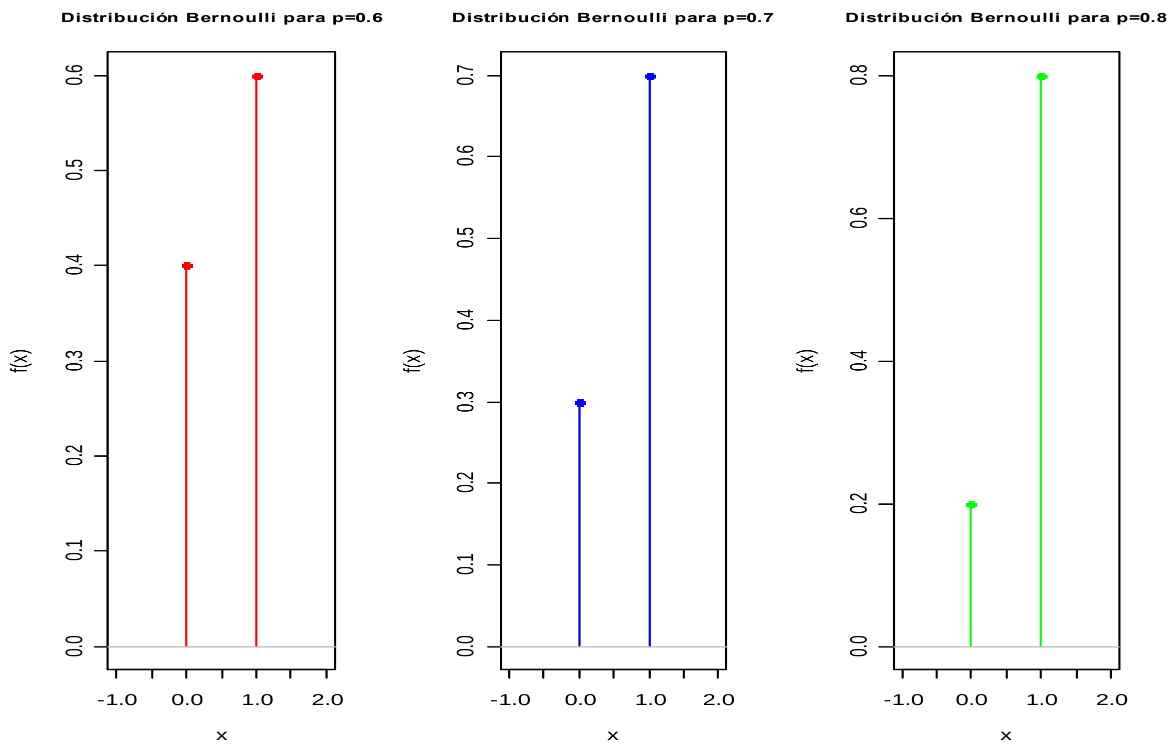
$$p + q = 1 \Rightarrow \mathrm{q} = 1 - \mathrm{p}$$

Nos queda:

$$f_X(1) = P(X = 1) = p$$

$$f_X(0) = P(X = 0) = 1 - p$$

Representación gráfica de la función de probabilidad de una variable aleatoria con distribución Bernoulli.



4-2-1-1 ESPERANZA DE LA VARIABLE BERNOULLI

A continuación se calculará algunas medidas importantes del Modelo Bernoulli, primeramente veremos esperanza de la variable Bernoulli:

$$E(X) = \sum_{x_j \in X(S)} x_j \; f_X(x_j) = 0 \cdot f_X(0) + 1 \cdot f_X(1) = 0 + p = p$$

$$E(X) = p$$

Luego la esperanza de una variable Bernoulli, es su parámetro p, de ahora en más cuando se sepa que la variable tiene distribución Bernoulli, directamente cuando se quiera calcular la esperanza se aplicará está relación.

4-2-1-2 VARIANZA DE LA VARIABLE BERNOULLI

Se calculará la varianza de una variable Bernoulli, para lo cual se partirá de la fórmula de cálculo de la varianza:

$$\sigma^2 = Var(X) = E(X^2) - E^2(X)$$

Calculemos primero $E(X^2)$

$$E(X^2) = \sum_{x_j \in X(S)} x_j^2 \; f_X(x_j) = 0^2 \cdot f_X(0) + 1^2 \cdot f_X(1) = 0 + p = p$$

$$Var(X) = E(X^2) - E^2(X) = p - p^2 = p(1-p)$$

$$\sigma^2 = Var(X) = p(1-p)$$

Entonces cuando se sepa que se está trabajando con una variable Bernoulli, igual como se hizo con la Esperanza, cuando debamos calcular varianza aplicaremos esta fórmula.

4-2-1-3 DESVIACIÓN ESTÁNDAR LA VARIABLE BERNOULLI

La desviación estándar es la raíz cuadrada de la varianza de la variable, por lo tanto, nos queda:

$$D.E.(X) = \sqrt{\sigma^2} = \sqrt{Var(X)}$$

$$\sigma = D.E.(X) = \sqrt{p(1-p)}$$

4-2-1-4 FUNCIÓN GENERADORA DE MOMENTOS

Si X es una variable aleatoria con distribución Bernoulli con parámetro p, luego su función generadora de momento es:

$$m_X: \varepsilon_0 \to R, \qquad siendo\ \varepsilon_0 = R, \qquad ya\ que\ e^t\ está\ definido\ para\ todo\ t\ real$$

$$m_X(t) = E(e^{tX})$$

$$m_X(t) = \sum_{x_j \epsilon X(\Omega)} e^{tx_j} f_X(x_j)$$

Reemplazamos x_j por los valores que toma, que son 0 y 1

$$m_X(t) = e^{t0} f_X(0) + e^{t1} f_X(1)$$

$m_X(t) = (1-p) + e^t p$ función generadora de momentos.

Ejemplo 4-4:
Considere el experimento aleatorio de una persona elegida al azar, que solicita un empleo es entrevistada para una prueba de aptitud, y observar el éxito o fracaso de la entrevista, como se dijo anteriormente puede ser descripto como un proceso de Bernoulli. La probabilidad de que la entrevista sea un éxito es p=0,42. ¿Cuántas entrevistas se esperan que sean exitosas?, ¿Con qué desviación estándar?

Resolución:
Considere la variable aleatoria X: "Número de entrevistas exitosas obtenidas", se observa que la distribución de probabilidades de X cumple las condiciones del modelo Bernoulli.

$$X: \Omega \to R;$$

$$X(éxito) = 1$$

$$X(fracaso) = 0$$

El recorrido de la variable X es:

$$X(\Omega) = \{0, 1\}$$

La probabilidad de éxito es p=0,42, y la probabilidad de fracaso es q=0,58

$$P(\{éxito\}) = P(X = 1) = f_X(1) = 0{,}42$$

$$P(\{fracaso\}) = P(X = 0) = f_X(0) = 0{,}58$$

Nos queda:

$$f_X(1) = P(X = 1) = p = 0{,}42$$

$$f_X(0) = P(X = 0) = 1 - p = 0{,}58$$

Función de probabilidad del Modelo Bernoulli:

$$f_X(x) = \begin{cases} p^x(1-p)^{1-x}, & si \;\; x = 0, 1 \\ \\ 0 & para \;\; otro\, valor \end{cases}$$

Para el ejemplo dado la función de probabilidad queda:

$$f_X(x) = \begin{cases} 0{,}42^x(1-0{,}42)^{1-x}, & si \;\; x = 0, 1 \\ \\ 0 & para \;\; otro\, valor \end{cases}$$

Esperanza:

$$E(X) = p$$

$$E(X) = 0{,}42$$

Se espera que en promedio el número de entrevistas exitosas sea 0,42.

Varianza:

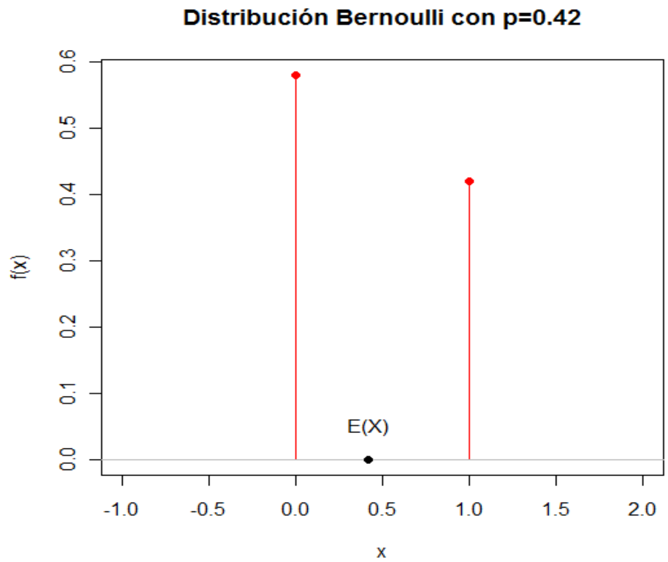
$$\sigma^2 = Var(X) = p(1-p)$$

$$Var(X) = 0{,}42.\,(1-0{,}42) = 0{,}2436$$

La desviación estándar es:

$$\sigma = \sqrt{Var(X)} = \sqrt{0{,}2436} = 0{,}4936$$

En promedio el número de entrevistas exitosas se desvía de su media en 0,4936.



4-2-2 MODELO BINOMIAL

Suponga ahora que se realiza n ensayos independientes, cada uno de los cuales tiene dos posibles resultados, “éxito” o “fracaso”. La probabilidad de obtener como resultado un éxito es p, y la probabilidad de obtener un fracaso es (1-p). Si X representa el número de éxitos que ocurren en n ensayos independientes, entonces X se dice que es una variable aleatoria binomial con parámetros: n y p.

Luego el espacio muestral es: $\Omega = \{0,1\}$

El recorrido de la variable aleatoria es $X(\Omega) = \{0,1,2,\dots,n\}$

Considere el ejemplo dado (4-2): una persona elegida al azar que solicita un empleo es entrevistada para una prueba de aptitud, observar el éxito o fracaso de la entrevista, es una variable Bernoulli.

Si ahora se realiza a n personas seleccionadas aleatoriamente, que solicitan un empleo una entrevista para una prueba de aptitud, si la proporción de personas que pasan con éxito la entrevista es constante durante un período razonable y se observa el número de personas que pasan con éxito la entrevista, entonces se puede aplicar el modelo Binomial, dado que las distintas personas tienen un criterio independiente al ser entrevistadas.

Una variable X tiene una distribución Binomial si cumple con las siguientes características:

1. Es una variable aleatoria discreta.
2. Cada ensayo es una variable Bernoulli, es decir tiene dos resultados posibles, mutuamente excluyentes, “éxito” o “fracaso”, con probabilidades p y (1-p) respectivamente.
3. Se realizan n ensayos estadísticamente independientes. (n es un número fijo para cada experimento, que denominaremos parámetro).

4. Como los ensayos son estadísticamente independientes entonces la probabilidad de éxito p, permanece constante durante los distintos ensayos.
5. La variable aleatoria X representa la cantidad de éxitos, en n ensayos independientes.
6. Su función de probabilidad o de cuantía es:

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & si\ x = 0, 1, \dots n \\ \\ 0 & paraotrovalor \end{cases} \qquad (4.1)$$

$\binom{n}{x}$ es el número combinatorio:

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

Siendo n y p: los parámetros de la distribución de Binomial. Siendo p la probabilidad de que el suceso sea un éxito, con $0 \leq p \leq 1$, y n la cantidad de ensayos independientes realizados, n pertenece al conjunto de los enteros positivos.

En forma simbólica se anota: $X \sim B(x, n, p)$

Para cada par de valores (n, p), se tendrá una función de densidad, todas estas definen una familia de distribuciones binomiales.

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & si\ x = 0, 1, \dots n \\ \\ 0 & para\ otro\ valor \end{cases} \qquad \text{(4-1)}$$

La validez de la ecuación (4-1) puede ser verificada, señalando que la probabilidad de cualquier secuencia particular de n resultados que contiene x éxitos y (n-x) fracasos es, por la independencia supuesta de los ensayos:

$$\underbrace{p.p \dots p}_{x\ éxitos} . \underbrace{(1-p).(1-p) \dots (1-p)}_{(n-x) fracasos} = p^x (1-p)^{n-x}$$

Luego debemos considerar las $\binom{n}{x}$ formas que se pueden presentar las diferentes secuencias de los n resultados que conducen a x éxitos y (n-x) fracasos.
Una de ellas por ejemplo es:

$$(1-p).\ \underbrace{p.p \dots p}_{(x-1) éxitos}\ .\underbrace{(1-p).(1-p) \dots (1-p)}_{(n-x-1) fracasos}.p = p^x (1-p)^{n-x}$$

Por lo tanto considerando todas las distintas formas, $\binom{n}{x}$, que se pueden presentar los n ensayos con x éxitos, llegamos a la fórmula (4.1).

Por ejemplo, en el caso que nos interese la probabilidad de que la variable aleatoria tome el valor de 2 éxitos en 4 ensayos independientes, es decir para x = 2 y n=4, entonces tenemos $\binom{4}{2} = 6$ formas en que los 4 ensayos pueden dar lugar a dos éxitos, a saber, cualquiera de los siguientes resultados:

$$(e,e,f,f); (e,f,e,f);\ (e,f,f,e); (f,e,e,f); (f,e,f,e); (f,f,e,e)$$

Como cada uno de estos resultados tiene una probabilidad de ocurrir de

$$P\big((e,e,f,f)\big) = P(E \cap E \cap F \cap F)$$

Debido a la independencia de sucesos tenemos:

$$P(E \cap E \cap F \cap F) = P(E).P(E).P(F).P(F)$$

$$Como \qquad P(E) = p, \qquad P(F) = 1 - p$$

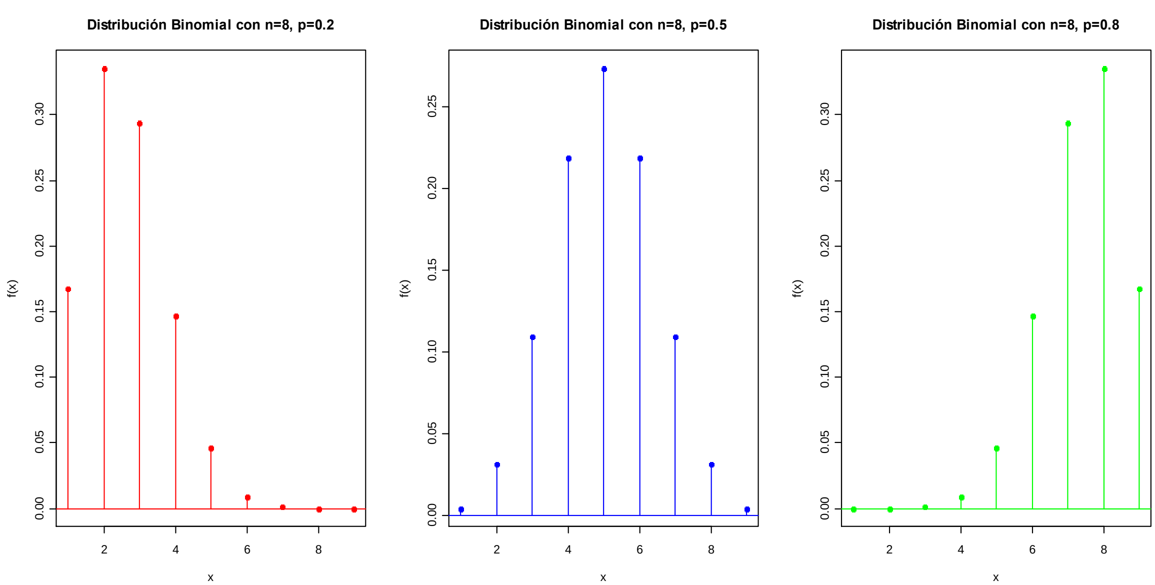
$$P(E \cap E \cap F \cap F) = P(E).P(E).P(F).P(F) = p.p.(1-p).(1-p)$$

$$P(E \cap E \cap F \cap F) = p^2(1-p)^2$$

Luego la probabilidad buscada de dos éxitos en cuatro pruebas es de:

$$f(2) = P(X = 2) = \binom{4}{2} p^2(1-p)^2$$

Representación gráfica de la función de probabilidad de una variable aleatoria con distribución Binomial:



Observe las que gráficas para $p = 0{,}2\ y\, para\, p = 0{,}8$, estas son asimétricas positiva (o a derecha) y negativa (o a izquierda) respectivamente, en cambio para el parámetro $p = 0{,}5$ es simétrica.

Son múltiples las aplicaciones prácticas de este modelo, en los distintos campos de las ciencias, por ejemplo en un proceso productivo algunas unidades resultan defectuosas, provenientes de distintas máquinas, y distintos operarios, si la proporción de estas unidades es constante durante un período razonable y se seleccionan aleatoriamente un determinado número de unidades, entonces se podrá emplear el modelo binomial para hacer proposiciones de probabilidad con respecto al número de artículos defectuosos.

Ejemplo 4-5:

Por experiencia previa se sabe que el 2 por ciento de los microprocesadores producidos por un fabricante de hardware para computadoras son defectuosos. Suponga que hay independencia entre los eventos. Un comprador ordena 100 de estos microprocesadores.

a) Enuncie la variable aleatoria en estudio.
b) ¿Cuál es la distribución de X? Determine sus parámetros.
c) ¿Cuál es la probabilidad de recibir 2 microprocesadores defectuosos?
d) ¿Cuál es la probabilidad de recibir menos de dos microprocesadores defectuosos?
e) ¿Cuál es la probabilidad de recibir más de 1 microprocesadores defectuoso?

Solución:

a) La variable aleatoria es:
X: "Número de microprocesadores defectuosos recibidos por el comprador de 100"
b) ¿Cuál es la distribución de X? Determine sus parámetros.
Se tiene n pruebas de Bernoulli, estadísticamente independientes con probabilidad de éxito constante, luego X se distribuye como una Binomial.

$$X \sim B(x, n = 100, p = 0{,}02)$$

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & si\ x = 0, 1, \ldots n \\ 0 & para\ otro\ valor \end{cases}$$

Sus parámetros son $n = 100, y\ \ p = 0{,}02$.
Por lo tanto la función de probabilidad es:

$$f_X(x) = \begin{cases} \binom{100}{x} 0{,}02^x (1-0{,}02)^{100-x}, & si\ x = 0, 1, \ldots 100 \\ 0 & para\ otro\ valor \end{cases}$$

c) ¿Cuál es la probabilidad de recibir 2 microprocesadores defectuosos?

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$$
$$f(2) = P(X = 2) = \binom{100}{2} 0{,}02^2 (1-0{,}02)^{100-2}$$
$$f(2) = P(X = 2) = 0{,}2734139$$

Interpretación: La probabilidad de recibir dos microprocesadores defectuosos es de 0,2734139.

Resolución con el software estadístico R:

Hacer todos los cálculos a veces es muy complejo, por lo que se puede utilizar un software estadístico para facilitarlos. En este caso se utilizará un software denominado "R". La sentencia para encontrar el valor de la función de probabilidad en el punto, es colocar d seguido de la distribución que se está buscando, en este caso "binom", entre paréntesis se colocará el valor del punto donde se quiere que se evalúe la función, y sus parámetros.
Nota: el software "R" trabaja utilizando puntos en los números decimales en vez de coma, y separa los números o sentencias con coma.
Para el presente ejercicio ítem c) la sentencia es:

- $dbinom(x, n, p)$
- $dbinom(2, 100, 0.02)$
- 0.2734139

d) ¿Cuál es la probabilidad de recibir menos de dos microprocesadores defectuosos?

$$P(X < 2) = P(\leq 1) = \sum_{x_j=0}^{x_j=1} P(X = x_j) = P(X = 0) + P(X = 1)$$

$$P(X < 2) = P(\leq 1) = \binom{100}{0} 0{,}02^0 (1 - 0{,}02)^{100-0} + \binom{100}{1} 0{,}02^1 (1 - 0{,}02)^{100-1}$$

$$P(X < 2) = P(\leq 1) = 0{,}1326196 + 0{,}2706522 = 0{,}4032718$$

Interpretación: La probabilidad de recibir menos de dos microprocesadores defectuosos es de 0,4032718.

Resolución con el software estadístico R:

En este caso se está aplicando la función de distribución acumulada en un determinado punto, para el cálculo con "R" la sentencia a utilizar es "p" seguida por el nombre de la distribución a utilizar, en este caso se aplica distribución binomial, abreviadamente se indica "binom", entonces:

- $pbinom(x, n, p)$
- $pbinom(1{,}100{,}0.02)$
- 0.4032717

e) ¿Cuál es la probabilidad de recibir más de 1 microprocesadores defectuoso?

$P(X > 1) = 1 - P(X \leq 1) = 1 - [P(X = 0) + P(X = 1)] = 1 - 0{,}4032718$
$P(X > 1) = 0{,}5967283$

La probabilidad de recibir más de un microprocesador defectuoso es de 0,5967283.

Resolución con el software estadístico R:

Como en "R" los operadores elementales son los mismos "+", "-", "*" y "/", en este caso tenemos que realizar una diferencia por lo que podemos colocar directamente la sentencia:

- $1 - pbinom(x, n, p)$
- $1 - pbinom(1{,}100{,}0.02)$
- $0{,}5967283$

Ejemplo 4-6:

Una empresa que produce tornillos sabe que ocasionalmente tendrán algunos defectuosos. La probabilidad de producir un tornillo defectuoso es de 0,01, independientemente uno de otro. La empresa vende los tornillos en paquetes de 10 y ofrece una garantía de reembolso que, como máximo 1 de los 10 tornillos es defectuoso. ¿Qué proporción de paquetes vendidos deberá la compañía reemplazar?

Sea X: "Número de tornillos defectuosos en un paquete"
X es una variable aleatoria binomial con parámetros n=10 y p=0,01. Anotaremos:

$$X \sim B(x, n = 10, p = 0{,}01)$$

Un paquete será reemplazado según la garantía de la empresa cuando tenga más de un tornillo defectuoso.

Por lo tanto la probabilidad que un paquete tenga que ser reemplazado es:

$$P(X > 1) = 1 - P(X = 0) - P(X = 1)$$

$$P(X > 1) = 1 - \binom{10}{0} 0{,}01^0 (1 - 0{,}01)^{10-0} - \binom{10}{1} 0{,}01^1 (1 - 0{,}01)^{10-1}$$

$$P(X > 1) = 1 - 0{,}904382 - 0{,}091352 = 0{,}004266$$

Por lo tanto la probabilidad de que un paquete tenga que ser reemplazado es de 0,0042.

Resolución con el software R:

Con el software estadístico "R", se puede calcular con la sentencia:

- $1 - pbinom(x, n, p)$

  Otra forma es utilizar la sentencia "$dbinom(x, n, p)$", que para nuestro ejemplo nos queda:
- $dbinom(0, n, p)$
- $dbinom(1, n, p)$

O directamente colocando:

- $1 - dbinom(0, n, p) - dbinom(1, n, p)$
- $1 - dbinom(0{,}10{,}0.01) - dbinom(1{,}10{,}0.01)$
- $0.0042662 \cong 0.004266$

Nota:

- La sentencia $dbinom(x, n, p)$, da la probabilidad de que una variable binomial tome exactamente un valor x, $f(x)$, cuando sus parámetros son n y p.

- También se puede utilizar la sentencia $pbinom(x, n, p)$, la cual da la probabilidad acumulada en el punto x, $F(x)$, de una variable binomial con parámetros n y p.

4-2-2-1 FUNCIÓN GENERADORA DE MOMENTOS

Se obtendrá la función generadora de momentos para luego a partir de ella obtener la esperanza y la varianza de la variable aleatoria.

Sea la función generadora de momentos:

$m_X: \varepsilon_0 \to R$ dada por: $m_X(t) = E(e^{tX})$, siendo $\varepsilon_0 = R$

$$m_X(t) = \sum_{x_j \epsilon X(\Omega)} e^{tx_j} f_x(x_j)$$

Se remplaza $f_x(x_j)$, por la función de probabilidad de la distribución binomial:

$$m_X(t) = \sum_{x_j=0}^{n} e^{tx_j} \binom{n}{x_j} p^{x_j} (1-p)^{n-x_j}$$

Reagrupando

$$m_X(t) = \sum_{x_j=0}^{n} \binom{n}{x_j} (e^t p)^{x_j} (1-p)^{n-x_j} \qquad (4.2)$$

Si se recuerda el binomio de Newton:

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j \, b^{n-j}$$

Ahora se compara la ecuación (4.2) con el binomio de Newton tenemos:

$$a = (e^t p); \;\; b = (1-p); \;\; y \; j = x_j$$

Luego la función generadora de momento en el modelo binomial es:

$$m_X(t) = \left(e^t p + (1-p)\right)^n$$

4-2-2-2 ESPERANZA DE LA VARIABLE BINOMIAL

Para obtener la esperanza de la variable binomial a partir de la función generadora de momentos, se calculará la derivada primera de la función generadora de momentos y luego se evalúa en t=0.

Función generadora de momentos de una variable binomial:

$$m_X(t) = \left(e^t p + (1-p)\right)^n$$

Se calcula su derivada primera:

$$m'_X(t) = \frac{d(m_X(t))}{dt} = n.\left(e^t p + (1-p)\right)^{n-1}.e^t p$$

Se evalúa la función $m'_X(t), \quad en\ t = 0$

$$t = 0 \Rightarrow m'_X(t=0) = n.p$$

Luego la esperanza de una variable aleatoria binomial es el producto de sus dos parámetros.

$$E(X) = n.p$$

4-2-2-3 VARIANZA DE LA VARIABLE BINOMIAL

Considerando la fórmula de varianza de una variable aleatoria tenemos:

$$Var(X) = E(X^2) - [E(X)]^2 \quad \textit{(4-2-1)}$$

Se calcula primeramente la esperanza de la variable al cuadrado $E(X^2)$, la cual se obtiene a partir de la función generadora de momentos. Como acabamos de hacer con la esperanza, se calcula la derivada segunda de la función generadora de momentos y luego se evalúa en $t = 0$.

$$m_X(t) = \left(e^t p + (1-p)\right)^n$$

Su derivada primera es:

$$m'_X(t) = \frac{d(m_X(t))}{dt} = n.\left(e^t p + (1-p)\right)^{n-1}.e^t p$$

Se vuelve a derivar para obtener la derivada segunda:

$$m''_X(t) = \frac{d(m'_X(t))}{dt^2} = n.(n-1).\left(e^t p + (1-p)\right)^{n-2}.e^{2t}p^2 + n.\left(e^t p + (1-p)\right)^{n-1}.e^t p$$

$$t = 0 \Rightarrow m''_X(t=0) = n.(n-1).p^2 + n.p = E(X^2)$$

$$E(X^2) = n.(n-1).p^2 + n.p \quad \text{(4-2-2)}$$

Recordando que la esperanza de un variable binomial es el producto de sus parámetros se tiene:

$$E(X) = n.p \quad \text{(4-2-3)}$$

Luego reemplazando en la fórmula de varianza (4-2-1) las ecuaciones (4-2-2) y (4-2-3), se obtiene la varianza de una variable aleatoria binomial:

$$Var(X) = E(X^2) - [E(X)]^2$$
$$Var(X) = n.(n-1).p^2 + n.p - (n.p)^2$$
$$Var(X) = n^2.p^2 - n.p^2 + n.p - n^2.p^2 = n.p - n.p^2$$
$$\sigma^2 = Var(X) = n.p(1-p)$$

4-2-2-4 DESVIACIÓN ESTÁNDAR DE LA VARIABLE BINOMIAL

Muy importante es calcular la desviación estándar, dado que cuando queremos estudiar la forma de la distribución se compara la media con la desviación estándar. No comparamos la media con la varianza, esto se debe que ambas medidas tienen distintas unidades.

$$\sigma = D.E.(X) = \sqrt{\sigma^2} = \sqrt{n.p(1-p)}$$

Ejemplo 4-7:

Considerando el ejemplo 4-6. Una empresa que produce tornillos sabe que ocasionalmente tendrán algunos defectuosos. La probabilidad de producir un tornillo defectuoso es de 0,01, independientemente uno de otro. La empresa vende los tornillos en paquetes de 10 y ofrece una garantía de reembolso que, como máximo 1 de los 10 tornillos es defectuoso.

a) ¿Cuál es el número esperado de tornillos defectuosos por paquete?
b) ¿Cuál es su varianza y desviación estándar?

Sea X: "Número de tornillos defectuosos en un paquete"

$$X \sim B(x, n = 10, p = 0{,}01)$$

a)

$$E(X) = n.p$$
$$E(X) = 10.0{,}01 = 0{,}1$$

En promedio el número de tornillos defectuosos que se espera encontrar en un paquete de 10 tornillos es de 0,1.

b)

$$Var(X) = n.p(1-p)$$
$$Var(X) = 10.0{,}01.(1-0{,}01) = 0{,}099$$

$$DE(X) = \sqrt{Var(X)} = \sqrt{n.p(1-p)} = \sqrt{0{,}099} = 0{,}3146$$

Se espera que en promedio el número de tornillos defectuosos que se encuentre en un paquete de 10, se desvíen respecto de su media en 0,3146 unidades.

Ejemplo 4-8:
Un proceso de fabricación de tornillos se comprueba, inspeccionando cada hora n tornillos seleccionados aleatoriamente entre los producidos durante dicho lapso. Si uno o más de los tornillos son defectuosos, el proceso se detiene y se examina. ¿Cuál debe ser el valor de n, si el fabricante desea que la probabilidad de que el proceso se detenga, cuando el 10% de los tornillos producidos sea defectuoso, sea alrededor del 95%? Suponga independencia de la calidad de cualquier tornillo respecto de los demás.

Sea X: "Número de tornillos defectuosos"

$$X \sim Binomial\ (x, n, p = 0{,}10)$$

$$p = 0{,}10;\ q = 0{,}90;\ n = ?$$

$$P(Se\ detenga\ el\ proceso) = P(X \geq 1) = 0{,}95$$

$$P(X \geq 1) = 1 - P(X = 0) = 0{,}95$$

$$1 - P(X = 0) = 1 - \binom{n}{0} . p^0 . (1-p)^{n-0} = 0{,}95$$

$$1 - (0{,}90)^n = 0{,}95$$

$$1 - 0{,}95 = (0{,}90)^n$$

$$0{,}05 = (0{,}90)^n$$

Aplicando logaritmo natural a ambos miembros de la igualdad:

$$ln0{,}05 = n . ln0{,}90$$

$$n = \frac{ln0{,}05}{ln0{,}90} = 28{,}43$$

Para asegurar una probabilidad de 0,95 debemos tomar una muestra de tamaño 29. Siempre se toma el número entero superior.

$$n = 29$$

4-2-3 MODELO POISSON O DE LOS SUCESOS RAROS

Se estudiará otra distribución de probabilidad discreta llamada distribución de Poisson, fue introducida por el francés Simeón Dennis Poisson (1781-1840), en un libro en que aplicó la teoría de probabilidades a los pleitos y juicios penales. Este libro, publicado en 1837, fue titulado "Reacherches sur la probabilité des jugements en matières criminelles et matière civile", (Investigación sobre la probabilidad de los juicios en materia criminal y civil).

El modelo de Poisson describe la probabilidad como un evento muy raro, acaecido en un tiempo o en un intervalo de espacio. La distribución de Poisson aparece a menudo cuando se trata del número de ocurrencias de un cierto suceso en un número muy grande de observaciones, cuando la probabilidad de que ocurra el suceso en una sola observación es muy pequeña.

En un número importante de aplicaciones, el experimento aleatorio fundamental consiste en observar el número de ocurrencias de un cierto suceso durante un intervalo de tiempo de duración t, donde la elección de t se deja a arbitrio del investigador. Esta situación se presenta por ejemplo en problemas de tráfico telefónico, donde se considera el número de llamadas telefónicas erróneas durante un intervalo de tiempo. Suponga que las veces que ocurre el suceso durante intervalos de tiempo no solapados son siempre independientes. Como se dijo la probabilidad de que ocurra un suceso en un intervalo de tiempo, Δt es un valor pequeño. Por lo tanto la probabilidad de que ocurran dos o más sucesos en un mismo intervalo tiende a cero (o es despreciable).

La probabilidad de una ocurrencia durante cualquier intervalo de tiempo muy pequeño debe ser aproximadamente proporcional a la longitud de ese intervalo. Una de las consecuencias de esta condición es que el proceso observado debe ser estacionario sobre el periodo de observación completo, esto es, la probabilidad de una ocurrencia debe ser la misma sobre el periodo completo. No puede haber intervalos con mucha ocurrencia durante los cuales se sabe de antemano que es probable que las ocurrencias sean más frecuentes, ni intervalos tranquilos, durante los cuales se sabe de antemano que es probable que las ocurrencias sean menos frecuentes. Esta condición se refleja en el hecho de que la constante λ, que indica el número promedio de ocurrencias por unidad de tiempo o espacio, en cualquier intervalo considerado durante el periodo completo de observación, se mantiene constante.

Este modelo tiene características análogas al Modelo Binomial la diferencia más importante es que se emplea cuando los sucesos no ocurren como consecuencia de la realización de n pruebas, sino que con relación a unidades de tiempo, longitud, superficie o espacio. Además la distribución de Poisson se aplica a sucesos que pueden tener una probabilidad muy baja.

Muchos fenómenos se modelan como un proceso de Poisson, por ejemplo, el número de grietas en un dique por unidad de superficie. El número de accidentes de tránsito en un cruce por semana. El número de automovilistas que entran en una estación de servicio por minuto. El número de clientes que entran en una oficina de correos por hora. El número de partículas α liberadas por un material radioactivo por minuto. El número de llamadas equivocadas recibidas en una central telefónica durante un periodo de tiempo fijo.

Una variable X tiene distribución de Poisson si cumple con las siguientes características:

1. Es una variable aleatoria discreta.
2. Cada ensayo es una variable Bernoulli, es decir tiene dos resultados posibles, mutuamente excluyentes, “éxito” o “fracaso”.
3. El número de ocurrencias en dos intervalos de tiempo (espacio o volumen), disjuntos deben ser estadísticamente independientes entre sí.
4. El número promedio λ, de ocurrencias en cada unidad de tiempo (espacio o volumen), es constante.

5. La probabilidad de que haya dos o más ocurrencias en cualquier intervalo de tiempo muy pequeño debe tener una magnitud de menor orden que la probabilidad de que haya sólo una ocurrencia.
6. La variable aleatoria X representa el número de ocurrencias de un evento aleatorio por unidad de tiempo (espacio, o volumen).
   Su función de probabilidad es:

$$f_X(x) = \begin{cases} \dfrac{e^{-\lambda}.\lambda^x}{x!} & para\ x = 0,1,2,\dots y\ \ \lambda > 0 \\ 0 & para\ otro\ caso \end{cases}$$

λ es la tasa de ocurrencia, es el parámetro de la distribución. Representa el número promedio de ocurrencia por unidad de tiempo, espacio o volumen.

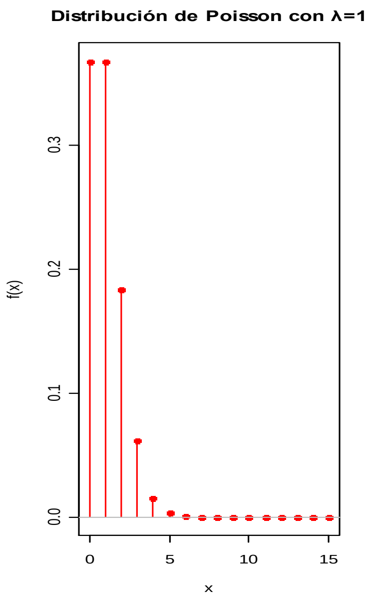
Lo anotamos en forma simbólica: $X \sim P(x,\ \lambda)$

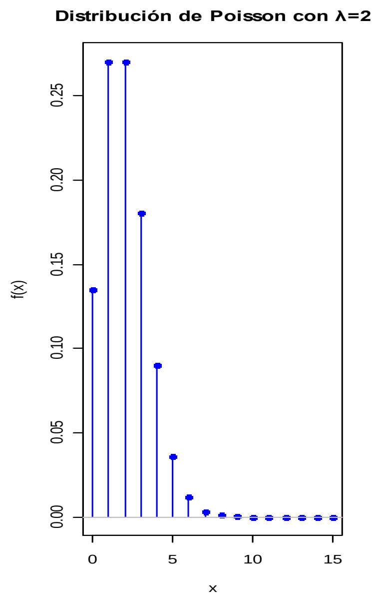
El símbolo e, conocido como número de Euler, es un número irracional cuyo valor es $e = 2{,}71828$ ....Muy usado en matemática, por ejemplo se lo encuentra en la base de los logaritmos naturales.

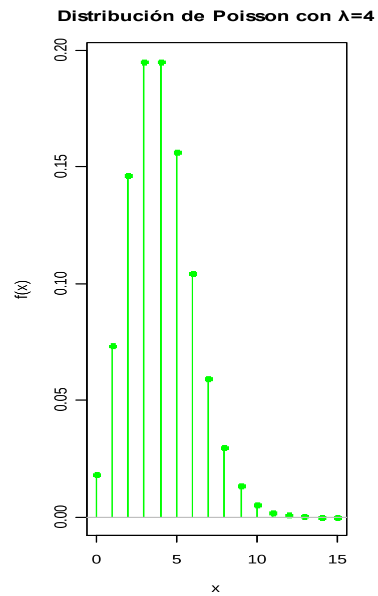
En principio, teóricamente es posible que suceda un número infinito de eventos en un intervalo dado. No hay límites al número de ensayos. Por lo que la variable aleatoria puede tomar cualquier valor natural incluido el cero.

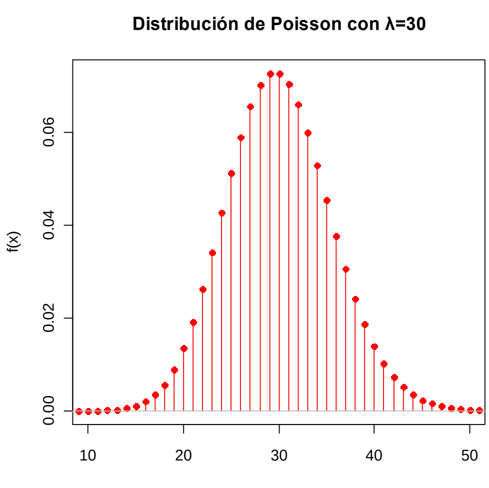
Para cada valor del parámetro λ, se tendrá definida una función de densidad, definiéndose una familia de distribución de Poisson.

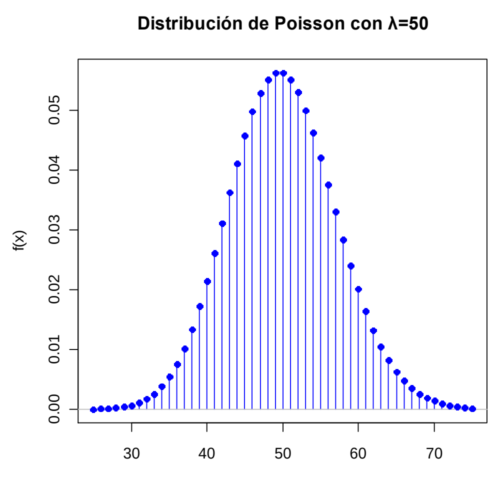
Representación gráfica de la función de probabilidad una variable aleatoria con distribución de Poisson:

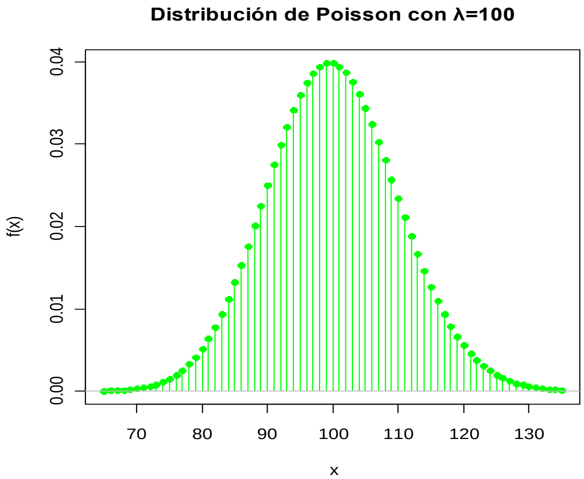












Obsérvese que cuando $\lambda = 1, \lambda = 2, \lambda = 4$, las distribuciones son asimétricas positivas o a derecha; y para $\lambda = 30$, $\lambda = 50$, $\lambda = 100$ las distribuciones se hacen simétricas. Es decir para valores pequeños del parámetro λ, la distribución es asimétrica a derecha o sesgada. A medida que el valor del parámetro aumenta, para $\lambda \geq 30$, la distribuciones se hace simétrica respecto a un eje $x = \lambda$.

Ejemplo 4-9:

El gerente de una fábrica que se dedica a la elaboración de productos químicos determina que en promedio el número de accidentes por semana es de 2. Considere que el número de accidentes por semana es una variable aleatoria con distribución de Poisson.

a) Enuncie la variable aleatoria en estudio. Indique los supuestos que se deben realizar para justificar el modelo elegido.
b) Determine la función de probabilidad de la variable aleatoria.
c) Encuentre la probabilidad de que ocurra un accidente en una semana.
d) Encuentre la probabilidad de que en una semana se produzca algún accidente.
e) Determine la probabilidad de que se produzcan cuatro accidentes en dos semanas.
f) Determine la probabilidad de que se produzcan dos accidentes en una semana y uno en la semana siguiente. ¿Es lo mismo que calcular la probabilidad de que se produzcan tres accidentes en dos semanas? ¿Por qué?
g) Si se sabe que ocurrió por lo menos un accidente en una semana particular. ¿Cuál es la probabilidad de que ocurran más de 2?

Resolución:

a) Enuncie la variable aleatoria en estudio. Indique los supuestos que se deben realizar para justificar el modelo elegido.

X: "Número de accidentes en una fábrica de productos químicos por semana"

$$X \sim P(x, \lambda = 2\ accidentes/semana)$$

Supuestos: Se supone que el número de accidentes es independiente de una semana a otra, y que ocurren a una razón constante λ.

b) Determine la función de probabilidad de la variable aleatoria.

$$f_X(x) = \begin{cases} \dfrac{e^{-2}.\,2^x}{x!} & para\ x = 0,1,2,\dots \\ 0 & para\ otro\ caso \end{cases}$$

c) Encuentre la probabilidad de que ocurra un accidente en una semana.

$$f_X(x) = P(X = 1) = \frac{e^{-2}.\,2^x}{x!} = \frac{2{,}718^{-2}.\,2^1}{1!} = 0{,}27067$$

Interpretación: La probabilidad de que ocurra un accidente en una semana es de 0,27067

Resolución con el software R:

- Para obtener la probabilidad en un punto en la distribución de Poisson colocamos "d" seguido del nombre de la distribución a utilizar en este caso, "pois" y entre paréntesis, el valor en el punto x, que buscamos y el valor del parámetro λ de la distribución. Es decir: $dpois(x,\ \lambda)$
  Entonces en "R" la sentencia es:
- $dpois(x,\ \lambda)$
- $dpois(1, 2)$
- $0{,}27067$

d) Encuentre la probabilidad de que en una semana se produzca algún accidente. Como en la distribución de Poisson la variable aleatoria puede tomar valores $x = 0,1,2,\dots$, es decir desde 0 hasta infinito, entonces cuando queremos encontrar probabilidades "mayor que", siempre lo vamos a calcular utilizando el complemento.

$$P(X > 0) = 1 - P(X \leq 0) = 1 - P(X = 0) = 1 - \frac{e^{-2}.\,2^0}{0!} = 1 - e^{-2} = 0{,}86766$$

Interpretación: La probabilidad de que en una semana se produzca algún accidente es de 0,86766

Resolución con el software R:

- Para obtener la probabilidad acumulada hasta un determinado punto x en la distribución de Poisson colocamos "p" seguido de "ppois" y entre paréntesis, el valor del punto x y el valor del parámetro λ de la distribución.
- R calcula siempre probabilidades "menores que", por lo tanto en este caso que queremos calcular una probabilidad de "mayor que", utilizaremos el complemento.

Por lo tanto para nuestro ejemplo para resolverlo utilizando el software "R" la sentencia a aplicar es:

- $1 - ppois(x, \lambda)$

- $1 - ppois(0, 2)$
- $0{,}86766$

e) Determine la probabilidad de que se produzcan cuatro accidentes en dos semanas.

Enunciaremos primeramente la variable aleatoria, la cual es:
Y= “Número de accidentes en una fábrica de productos químicos en dos semanas”
La unidad de tiempo ha variado, nuestra tasa de ocurrencia o promedio también cambia, se debe calcular primero este parámetro.
λ: es el número promedio de ocurrencia por unidad de tiempo.
$\lambda_1 = \lambda t$: es el número de ocurrencias en un intervalo t.
Luego la tasa de ocurrencia es:

$$\lambda_1 = 4\ (accidentes/2semanas)$$

$$Y \sim P(y, \lambda_1)$$

$$Y \sim P(y, 4)$$

$$f(4) = P(Y = 4) = \frac{e^{-4}.\,4^4}{4!} = 0{,}1953668$$

Interpretación: La probabilidad de que se produzcan cuatro accidentes en la fábrica de productos químicos en dos semanas es de 0,1953668

Resolución con el software R:

Con el software “R” la sentencia a utilizar es:

- $dpois(x, \lambda_1)$
- $dpois(4, 4) = 0.1953668$

f) Determine la probabilidad de que se produzcan dos accidentes en una semana y uno en la semana siguiente. ¿Es lo mismo que calcular la probabilidad de que se produzcan tres accidentes en dos semanas? ¿Por qué?
Sea X: “Número de accidentes en una fábrica de productos químicos en una semana”

$$\lambda = 2\ (accidentes/semana)$$

Sea Y: “Número de accidentes en una fábrica de productos químicos en la semana siguiente”

$$\lambda = 2\ (accidentes/semana)$$

Z: “Número de accidentes en una fábrica de productos químicos en dos semanas”

$$\lambda_1 = 4\ (accidentes/2semanas)$$

Se calculará a continuación la probabilidad de que se produzcan dos accidentes en una semana y uno en la siguiente:

$$P(X = 2 \cap Y = 1) \underbrace{=}_{ind.} P(X = 2).P(Y = 1) = \frac{e^{-2}.2^2}{2!}.\frac{e^{-2}.2^1}{1!} = 0{,}073263$$

Interpretación: La probabilidad de que se produzcan dos accidentes en una fábrica de productos químicos en una semana y un accidente en la semana siguiente es de 0,073263.

Ahora se calculará la probabilidad de que se produzcan tres accidentes en dos semanas:

$$f(3) = P(Z = 3) = \frac{e^{-\lambda_1}.\lambda_1{}^z}{z!} = \frac{e^{-4}.4^3}{3!} = \ 0{,}1953668$$

Interpretación: La probabilidad de que se produzcan tres accidentes en dos semanas es de 0,1953668.

Se concluye que no es lo mismo que se produzcan dos accidentes en una semana y uno en la semana siguiente, que se produzcan tres accidentes en dos semanas, hay más probabilidad en esta segunda opción. Debido que hay más casos a considerar en esta elección. Se pueden presentar los tres accidentes en la primera semana y ninguno en la segunda semana, o recíprocamente, o se pueden presentar dos accidentes en la primera semana y un accidente en la segunda semana y recíprocamente.

Resolución con el software R:

Se calculará a continuación la probabilidad de que se produzcan dos accidentes en una semana y uno en la siguiente, con el software “R”, la sentencia a utilizar es:

- $dpois(2,2) * dpois(1,2)$
- 0.073263

Ahora se calculará la probabilidad de que se produzcan tres accidentes en dos semanas, con el software “R” la sentencia a aplicar es:

- $dpois(x, \lambda)$
- $dpois(3,4)$
- 0.1953668

La interpretación y la conclusión es la misma aunque apliquemos distintos métodos de cálculo.

g) Si se sabe que ocurrió por lo menos un accidente en una semana particular. ¿Cuál es la probabilidad de que ocurran más de 2?

Sea X: “Número de accidentes en una fábrica de productos químicos en una semana”

$$\lambda = 2\ (accidentes/semana)$$

En este ítem tenemos un condicionante, se sabe que ocurrió por lo menos un accidente en una semana particular, por lo que la probabilidad pedida se expresa:

$$P(X > 2/X \geq 1) = \frac{P(X > 2 \cap X \geq 1)}{P(X \geq 1)} = \frac{(P(X > 2)}{P(X \geq 1)} = \frac{1 - P(X \leq 2)}{1 - P(X = 0)}$$

$$P(X > 2/X \geq 1) = \frac{0{,}323324}{0.864665} = 0{,}373929$$

Interpretación: La probabilidad de que ocurran más de dos accidentes dado que ocurrió por lo menos un accidente en una semana es de 0,373929

Resolución con el software R:

Si se utiliza el software "R", debemos tener en cuenta que no tenemos una sentencia particular para calcular una probabilidad condicionada en particular, por lo tanto aplicamos la definición y calculamos el cociente en R, entonces la sentencia a utilizar es:

- $\frac{1-\text{ppois}(2{,}2)}{1-\text{ppois}(0{,}2)}$
- 0.373929

4-2-3-1 APROXIMACIÓN DE LA DISTRIBUCIÓN BINOMIAL

Sea X una variable binomial con parámetros n y p, cuando el valor de n es grande $(n \to \infty)$ y el valor de p es chico, cercano a 0, $(p \to 0)$ la distribución binomial se puede aproximar a la distribución de Poisson con media $n.p$ y $(n.p \to \lambda)$.

Suponga que $X \sim B(n, p)$, recordando que la ecuación de la función binomial es:

$$P(X = x) = \frac{n!}{(n-x)!\,x!} p^x (1-p)^{n-x}$$

Si definimos $\lambda = n.p$ entonces podemos escribir la ecuación de la función binomial:

$$P(X = x) = \frac{n!}{(n-x)!\,x!} . \left(\frac{\lambda}{n}\right)^x . \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

$$P(X = x) = \frac{n.(n-1)\dots(n-x+1)}{x!} . \frac{\lambda^x}{n^x} . \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

Suponga ahora que $(n \to \infty)$ $y\, p \to 0$ de forma que $n.p$ permanece igual al valor fijo λ. Se sabe de cálculo elemental que:

$$\lim_{n\to\infty} P(X = x) = \lim_{n\to\infty} \frac{n.(n-1)\dots(n-x+1)}{x!} . \frac{\lambda^x}{n^x} . \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x}$$

$$\lim_{n\to\infty} P(X = x) = \lim_{n\to\infty} \frac{n.(n-1)\dots(n-x+1)}{x!} . \frac{\lambda^x}{n^x} . \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-x} \quad (3.1)$$

Resolviendo cada uno de estos límites por separado se obtiene

$$\lim_{n\to\infty} \frac{n.(n-1)\dots(n-x+1)}{n^x}.\frac{\lambda^x}{x!} = \frac{\lambda^x}{x!}; \quad \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{-x} = 1; \quad \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n} = e^{-\lambda}$$

Reemplazando en la ecuación (3.1) resulta:

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} \; ; (3.2)$$

Siendo la relación (3.2), la función de la distribución de Poisson con media λ. Esta relación expresa que la distribución binomial $B(n,p)$, converge a la distribución de Poisson $P(\lambda)$, cuando $n\to\infty,\ p\to 0\ y\ n.p\to\lambda$.

En la práctica para n ≥30 y p<0,05, se puede utilizar la aproximación a la distribución de Poisson.

Ejemplo 4-10:

Considere que la probabilidad de que un artículo producido por una determinada máquina sea defectuoso es 0,01. Encontrar la probabilidad de que una muestra de 35 artículos contenga como máximo 1 artículo defectuoso.

Sea X: “Número de artículos defectuosos producido por una máquina”

$$X\sim B(x, n=35, p=0{,}01)$$

$$P(X\le 1) = P(X=0) + P(X=1)$$

$$P(X\le 1) = \frac{35!}{(35-0)!\,0!}0{,}01^0(1-0{,}01)^{35-0} + \frac{35!}{(35-1)!\,1!}0{,}01^1(1-0{,}01)^{35-1}$$

$$P(X\le 1) = 0.9521413$$

Resolución utilizando la aproximación de Poisson:

Como n >30 y p<0,05, se puede utilizar la aproximación de la distribución de Poisson con:

$$\lambda = n.p = 35 * 0{,}01 = 0{,}35$$

Entonces: $X\to P(x,\lambda=0{,}35)$

$$P(X\le 1) = P(X=0) + P(X=1)$$

$$P(X\le 1) = \frac{e^{-0{,}35}0{,}35^0}{0!} + \frac{e^{-0{,}35}0{,}35^1}{1!} = 0{,}9513289$$

Obsérvese que es una muy buena aproximación.

Resolución con el software R:

- *pbinom(1,35,0.01)*
- 0.9521413

Resolución utilizando la aproximación de Poisson, con el software R:

>ppois(1,0.35)
> 0.9513289

Se puede observar que los valores aproximados obtenidos utilizando la distribución de Poisson, son bastantes cercanos al valor exacto que se obtuvo al aplicar la distribución binomial.

4-2-3-2 FUNCIÓN GENERADORA DE MOMENTOS

Se obtendrá la función generadora de momentos para la distribución de Poisson, para luego a partir de ella obtener la esperanza y la varianza de la variable aleatoria.

$$m_X(t) = E(e^{tX})$$

$$m_X(t) = \sum_{x=0}^{\infty} e^{tx} f_X(x)$$

Reemplazando $f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

$$m_X(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$m_X(t) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t\lambda)^x}{x!}$$

Por el desarrollo de Mc Laurin tenemos:

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x, \quad si\ consideramos\ x = e^t\lambda \ \ y\ n = x;\ para\ todo\ x$$

Se obtiene la función generadora de momentos de una variable Poisson:

$$m_X(t) = e^{-\lambda}.e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

Definida para todo t real, $\varepsilon_0 = R$

$$m_X(t) = e^{\lambda(e^t-1)}$$

4-2-3-2 PROPIEDAD REPRODUCTIVA DE VARIABLES ALEATORIAS DE POISSON

Si $X_1, X_2, \dots, X_n$ son n variables aleatorias independientes, cada variable X_i tiene distribución de Poisson con parámetro λ_i, $con\ i = 1,2, \dots . n$ entonces la variable aleatoria

$$Y = \sum_{i=1}^{n} X_i$$

También tiene distribución de Poisson con parámetro $\lambda = \sum_{i=1}^{n} \lambda_i$.

$$Y = P(\lambda)$$

Se comprobará por medio de la función generadora de momentos.

$$Sea\ X_1, X_2, \dots, X_n\ n\ variables\ independientes$$

$$X_1 \sim P(\lambda_1), X_2 \sim P(\lambda_2), \dots, X_n \sim P(\lambda_n)$$

Recordemos que la función generadora de momentos de una variable X_i con distribución de Poisson es:

$$m_{X_i}(t) = e^{\lambda_i(e^t-1)}$$

$$Sea\ Y = X_1 + X_2 + \cdots + X_n$$

Entonces la función generadora de momentos de la variable Y es:

$$m_Y(t) = m_{\sum_{i=1}^{n} X_i}(t) = E(e^{tY}) = E\left(e^{t\sum_{i=1}^{n} X_i}\right)$$

$$m_Y(t) = \prod_{i=0}^{n} E(e^{tX_i}) = \prod_{i=0}^{n} m_{X_i}(t)$$

$$m_Y(t) = m_{X_i}(t).m_{X_i}(t) \dots m_{X_n}(t)$$

$$m_Y(t) = e^{\lambda_1(e^t-1)}.e^{\lambda_2(e^t-1)} \dots . e^{\lambda_n(e^t-1)}$$

$$m_Y(t) = e^{(\lambda_1+\lambda_2+\cdots+\lambda_n)(e^t-1)}$$

$$m_Y(t) = e^{\sum_{i=1}^{n} \lambda_i(e^t-1)}$$

Luego

$$Y \sim P(\sum_{i=1}^{n} \lambda_i)$$

Esta es una propiedad muy utilizada en control estadístico de procesos. Por ejemplo considere un producto fabricado por un proceso. Por ejemplo una computadora, un LCD, u otro producto, donde se pueden producir múltiples defectos. El número de cada tipo de defecto es una variable aleatoria de Poisson. Entonces la propiedad reproductiva de la distribución de Poisson nos permite analizar la suma de todos los tipos defectos como una variable aleatoria de Poisson.

4-2-3-3 ESPERANZA DE LA VARIABLE POISSON

A partir de la función generadora de momentos se puede obtener la esperanza haciendo la derivada primera de dicha función y luego evaluándola en $t = 0$

$$m_X(t) = e^{-\lambda}.e^{e^t\lambda}$$

$$m'_X(t) = \frac{d(m_X(t))}{dt} = e^{-\lambda}.e^{e^t\lambda} \ .e^t\lambda$$

$$t = 0 \; entonces \;\; m'_X(t = 0) = \lambda$$

$$E(X) = \lambda$$

Por lo tanto tenemos que la esperanza de una variable de Poisson es su parámetro λ.

4-2-3-4 VARIANZA Y DESVIACIÓN ESTÁNDAR DE UNA VARIABLE POISSON

Si recordamos la fórmula para calcular la varianza:

$$Var(X) = E(X^2) - E^2(X)$$

Se observa que se debe calcular $\mathrm{E}(\mathrm{X}^2)$, que es el momento de segundo orden, para lo cual se debe obtener la derivada segunda de la función generadora de momentos y luego evaluarla en $t = 0$.

En símbolos: $m''_X(t = 0) = \mathrm{E}(\mathrm{X}^2)$

Su derivada primera es:

$$m'_X(t) = \frac{d(m_X(t))}{dt} = e^{-\lambda}.e^{e^t\lambda} \ .e^t\lambda$$

Entonces la derivada segunda es:

$$m''_X(t) = \frac{d(m'_X(t))}{dt^2} = e^{-\lambda}.e^{e^t\lambda} \ .(e^t\lambda \ )^2 + e^{-\lambda}.e^{e^t\lambda} \ e^t\lambda$$

$$t = 0 \quad entonces \quad m''_X(t = 0) = \lambda^2 + \lambda$$

$$E(X^2) = \lambda^2 + \lambda$$

$$Var(X) = E(X^2) - [E(X)]^2$$

$$Var(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$Var(X) = \lambda$$

$$\sigma^2 = Var(X) = \lambda$$

Obsérvese que la E(X) y la Var(X) son iguales al parámetro, esta es una propiedad de la distribución de Poisson.

Desviación estándar de una variable Poisson

$$DE(X) = \sqrt{Var(X)} = \sqrt{\lambda}$$
$$\sigma = \mathrm{DE(X)} = \sqrt{\lambda}$$

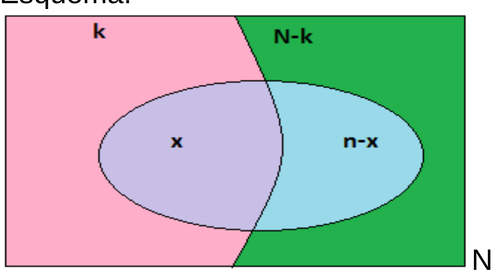
4-2-4 MODELO HIPERGEOMÉTRICO

Este modelo es muy importante cuando se selecciona una muestra sin reposición, en una población finita cuyos elementos pueden ser clasificados en dos categorías.

Considere una población finita de N unidades, de forma que k unidades poseen una determinada característica y $(N-k)$ de ellas no poseen dicha característica. Considere la experiencia de realizar una extracción al azar de una unidad aleatoriamente en dicha población, el resultado será una de las k unidades que poseen la característica en estudio o una de las $(N-k)$ unidades que no posee dicha característica. Si se toma una muestra aleatoria de n elementos sin reposición en esta población, cada extracción subsecuente es dependiente, por lo que la probabilidad de éxito cambia en cada extracción.

Se desea encontrar la probabilidad de que al extraer una muestra al azar de tamaño n, se obtenga x unidades del tipo k ó éxitos.

Esquema:



N: cantidad de unidades de la población finita.
k: unidades que cumplen una determinada característica en estudio, (éxito).
(N-k): unidades que no cumplen con la característica bajo estudio, (fracaso).
n: tamaño de la muestra.
x: cantidad de unidades que se extraen en la muestra que cumple con la característica en estudio, es decir de las k unidades.
(n-x): cantidad de unidades que se extraen de la muestra que no cumplen con la característica en estudio, es decir de las (N-k) unidades.

La variable siempre se enuncia en función del éxito, en este caso el éxito será obtener una unidad con la característica en estudio.

Características de una variable con distribución hipergeométrica:

1. Se realiza n ensayos, es decir se extrae n elementos sin reposición, por los que los sucesos son dependientes.
2. Cada ensayo es una variable aleatoria discreta Bernoulli, es decir tiene sólo dos posibles resultados mutuamente excluyentes: "éxito" o "fracaso".
3. La variable representa el número de éxitos en n realizaciones sin reposición.
4. Su función de probabilidad es:

$$f(x) = \begin{cases} \dfrac{\binom{k}{x} \cdot \binom{N-k}{n-x}}{\binom{N}{n}}, & si \;\; \max\{0, n-(N-k)\} \leq x \leq \; \min\{n, k\} \\ \\ 0 & en\ cualquier\ otro\ caso \end{cases}$$

Cuyos parámetros son N, k, n, que deben ser enteros positivos.
Lo anotamos en forma simbólica: $X \sim H(x, n, k, N)$
El valor de que puede tomar la variable aleatoria no puede exceder de n, ni de k. por lo tanto se debe verificar que el valor que puede tomar la variable X, es $X \leq \min\{n, k\}$, de la misma forma el número de fracasos que se obtienen, $(n-x)$, puede exceder a $(N-k)$; por otra parte el valor de X debe ser al menos $(n-(N-k))$, dado que el valor de x no puede ser menor que 0, se debe verificar que $X \geq \max\{0, (n-(N-k))\}$.
Por lo tanto el valor de X debe ser un número entero en el intervalo $\max\{0, (n-(N-k))\} \leq x \leq \min\{n, k\}$

Procedamos a interpretar la fórmula de su función de probabilidad la cual se obtiene a partir de la definición clásica de probabilidad "casos favorables" dividido por los "casos posibles". Siendo:

$\binom{k}{x}$: El número total de combinaciones posibles de extraer exactamente x unidades de k, siendo k las unidades que cumplen con cierta característica.

$\binom{N-k}{n-x}$: El número total de combinaciones posibles de extraer exactamente $(n-x)$, unidades de $(N-k)$, siendo $(N-k)$ las unidades que no cumplen con la característica en estudio.

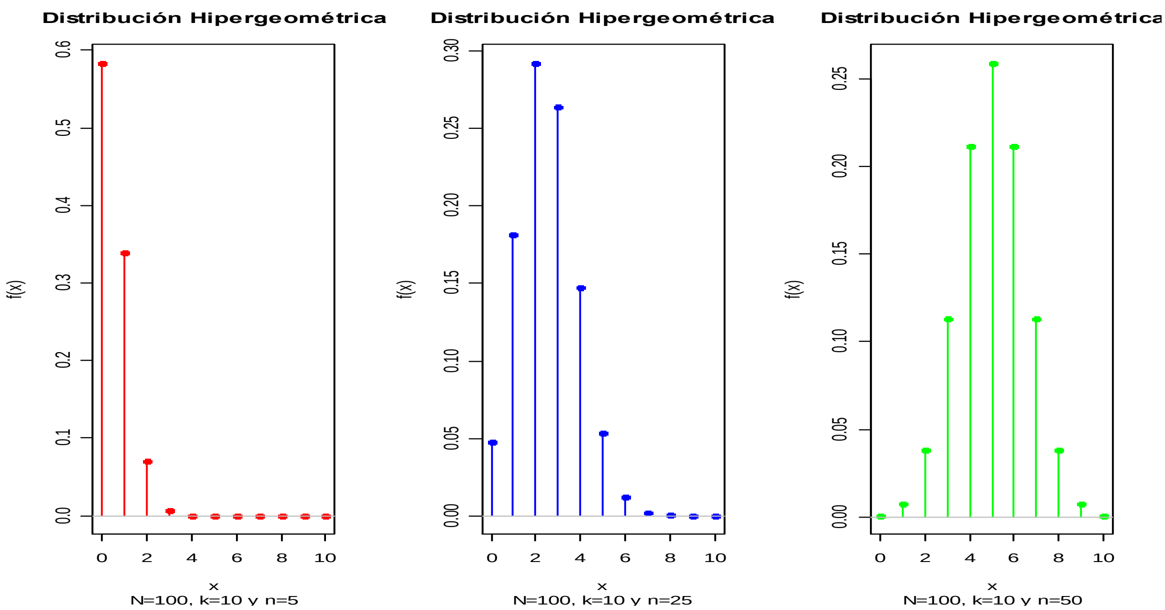
$\binom{N}{n}$: El número total de combinaciones para extraer de N.

Cada forma de $\binom{k}{x}$ combinada con cada forma de $\binom{N-k}{n-x}$ da el número total de formas mutuamente exclusivas de obtener x artículos con cierta característica en estudio en n extracciones sin reemplazo. Dado que $\binom{N}{n}$ es el número total de resultados y las extracciones son al azar, cada una de esas formas tiene la probabilidad $1/\binom{N}{n}$. De esto resulta la función de densidad de la distribución hipergeométrica.

Recuerde que $\binom{k}{x}$ es un número combinatorio que se calcula de la siguiente manera:

$$\binom{k}{x} = \frac{k!}{x!.(k-x)!}$$

Representación gráfica de la función de probabilidad de una variable aleatoria con distribución hipergeométrica:



Nota: Se puede observar que en el primero y segundo gráfico las distribuciones son asimétricas. Y en el tercer gráfico la distribución es simétrica. En general podemos concluir que a medida que el tamaño de muestra crece la distribución va tendiendo a ser más simétrica.

Ejemplo 4-11:
Un comprador de componentes electrónicos los compra en lotes de 10. Es su política de inspeccionar 3 componentes al azar de cada lote, acepta el lote sólo si todos los componentes inspeccionados son no defectuosos. Si el 30 por ciento de los lotes tienen 4 componentes electrónicos defectuosos y el 70 por ciento tiene sólo un componente electrónico defectuoso, ¿qué proporción de los lotes no rechazará el comprador?

Sea A: “El comprador acepta el lote”
B: “El lote tiene 4 componentes defectuosos”
C: “El lote tiene 1 componente defectuoso”

Sea X: “Número de componentes electrónicos defectuosos observados en una muestra de tamaño 3 sin reposición”.

$$X \sim Hipergeométrica\ (x, n = 3, k, N = 10)$$

El comprador acepta el lote si en su muestra tomada no hay componentes electrónicos defectuosos. En estos casos al inspeccionar una muestra esta se obtiene sin reposición, y un componente puede tener sólo dos posibles resultados: defectuoso o no defectuoso, por lo que se aplicará la distribución hipergeométrica.

Datos:

$P(B) = 0{,}30$; $P(C) = 0{,}70$

$$N = 10; n = 3; x = 0$$

$P(A) = P(X = 0)$ se calculará a partir de la distribución hipergeométrica.

Recordemos que:

$$P(A) = P(X = 0) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{M}{n}}$$

$P(A/B)$: significa que es la probabilidad de que el comprador acepte el lote dado que tiene 4 componentes defectuosos.

$P(A/C)$: significa que es la probabilidad de que el comprador acepte el lote dado que tiene un componente defectuoso.

$$P(A) = P(A/B).P(B) + P(A/C).P(C) \text{ (Teorema de Probabilidad Total)}$$

$$P(A) = \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}}.\left(\frac{3}{10}\right) + \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}}.\left(\frac{7}{10}\right) = 0{,}54$$

Por lo tanto la probabilidad de que el comprador acepte un lote es de 0,54 ó el 54% de los lotes serán aceptados.

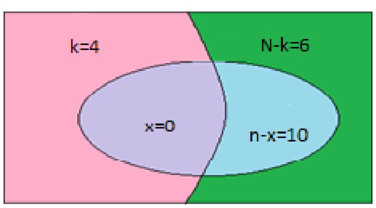


Fig.4-1

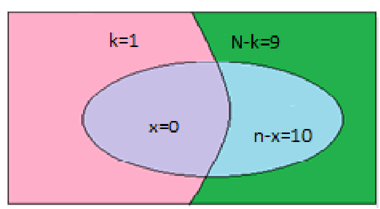


Fig.4-2

Esquemas de distribución hipergeométrica, en la fig. 4-1, se tiene k=4, que significa que existen 4 componentes electrónicos defectuosos en el lote y en la fig. 4-2, k=1, significa que en este esquema hay sólo un componente eléctrico defectuoso.

4-2-4-1 FUNCIÓN DE DISTRIBUCIÓN ACUMULATIVA DE LA DISTRIBUCIÓN HIPERGEOMÉTRICA

$$F(x) = P(X \leq x) = \sum_{i=0}^{x} \frac{\binom{k}{i} . \binom{N-k}{n-i}}{\binom{N}{n}}, \qquad si\ x = 0, 1, 2, \dots n$$

4-2-4-2 ESPERANZA DE UNA VARIABLE HIPERGEOMÉTRICA

Para determinar la esperanza de una variable hipergeométrica, aplicaremos la definición de esperanza:

$$E(X) = \sum_{x=0}^{n} x. f(x)$$

$$E(X) = \sum_{x=0}^{n} x. \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = k. \sum_{x=1}^{n} \frac{\binom{k-1}{x-1}\binom{N-k}{n-x}}{\binom{N}{n}}\ (4.4.1)$$

Puede demostrarse que:

$$\binom{N}{n} = \frac{N}{n}\binom{N-1}{n-1} \quad (4.4.2)$$

Luego reemplazando (4.2.4) en (4.4.1) tenemos

$$E(X) = k. \sum_{x=1}^{n} \frac{\binom{k-1}{x-1}\binom{N-k}{n-x}}{\frac{N}{n}\binom{N-1}{n-1}} = \frac{nk}{N}. \sum_{x=1}^{n} \frac{\binom{k-1}{x-1}\binom{N-k}{n-x}}{\binom{N-1}{n-1}}$$

Si reemplazamos: $M = N - 1;\ \ r = k - 1;\ \ s = n - 1; y = x - 1$

$$E(X) = n. \frac{k}{N}. \sum_{y=1}^{s} \frac{\binom{r}{y}\binom{M-r}{s-y}}{\binom{M}{s}}$$

Entonces:

$$\sum_{y=1}^{s} \frac{\binom{r}{y}\binom{M-r}{s-y}}{\binom{M}{s}} = 1$$

Esta sumatoria es igual a uno dado que es la suma de una función de probabilidad hipergeométrica con parámetros M, r, y s.
Por lo tanto:

$$E(X) = n.\frac{k}{N}$$

Nota: no se utiliza la función generadora de momentos para hallar la esperanza ni la varianza, en este caso porque es más complicado y tedioso, que partir de la definición de esperanza.

4-2-4-3 VARIANZA DE UNA VARIABLE HIPEGEOMÉTRICA

Con el mismo procedimiento utilizado para calcular la esperanza puede demostrarse que la varianza de una variable hipergeométrica es:

$$Var(X) = \left(\frac{N-n}{N-1}\right) n \left(\frac{k}{N}\right).\left(\frac{N-k}{N}\right)$$

Ejemplo 4-12:
Un comprador mayorista de componentes electrónicos tiene un plan de muestreo doble, que lleva a cabo cuando realiza compras mayores a sus gastos fijos mensuales. Si los lotes son de tamaño 20 su plan de muestreo es: a) selecciona una muestra de tamaño 2, si ambos son no defectuosos acepta el lote, b) si ambos son defectuosos se rechaza el lote, c) si un artículo es defectuoso y el otro es no defectuoso entonces toma una segunda muestra de tamaño uno. En esta segunda etapa su plan es aceptar si el artículo es no defectuoso y rechazar si el artículo es defectuoso. La probabilidad del lote de tener un componente electrónico defectuoso es 0,25. ¿Cuál es la probabilidad de aceptación de un lote?

X_1: "Número de componentes electrónicos defectuosos observados en una muestra de tamaño 2, en la primera etapa del muestreo"
X_2: "Número de componentes electrónicos defectuosos observados en una muestra de tamaño 1, en la segunda etapa del muestreo"

$$X_1 \sim Hipergeométrica(x_1, N = 20, k = 5, n = 2)$$
$$X_2 \sim Hipergeométrica(x_2, N = 18, k = 4, n = 1)$$

Datos:

$$N = 20;\ \ n = 2; p = 0{,}25$$

$$p = 0{,}25, como\ p = \frac{k}{N} \rightarrow k = p.N = 0{,}25\ .20 = 5;\ \ luego\ k = 5$$

Entonces:

a)

$$P_1(X_1 = 0) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\binom{5}{0}\binom{15}{2}}{\binom{20}{2}} = 0{,}55263$$

La probabilidad de aceptar el lote, si no se obtienen componentes electrónicos defectuosos es de 0,55263.

b)

$$P_1(X_1 = 2) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\binom{5}{2}\binom{15}{0}}{\binom{20}{2}} = 0{,}05263$$

Si se obtienen dos componentes defectuosos se rechaza el lote con una probabilidad de 0,05263.

c)

$$P_1(X_1 = 1) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\binom{5}{1}\binom{15}{1}}{\binom{20}{2}} = 0{,}39474$$

Si en la primera etapa se obtiene un componente no defectuoso y uno defectuoso, se toma una segunda muestra. En esta segunda etapa se toma una muestra de tamaño 1, recordando que hemos extraído dos componentes en la primera etapa del muestreo y no los hemos repuesto, por lo tanto el tamaño del lote es 18 y como se extrajo un componente defectuoso y un no defectuoso, por lo que nos queda 4 componentes defectuosos, k=4, y 14 no defectuosos.

$$P_2(X_2 = 0) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \frac{\binom{4}{0}\binom{14}{1}}{\binom{18}{1}} = 0{,}77777$$

La probabilidad de aceptar un lote es la probabilidad de obtener en la primera etapa una muestra con dos componentes no defectuosos o de obtener en la primera etapa un componente defectuoso y en la segunda etapa ningún componente defectuoso.

$$P(Aceptar) = P_1(X_1 = 0) + P[(X_1 = 1) \cap (X_2 = 0)]$$

$$P(Aceptar) = P_1(X_1 = 0) + P_1(X_1 = 1).P_2(X_2 = 0) = \ 0.85965$$

La probabilidad de aceptar un lote con el plan de muestreo doble analizado es de 0,85965.

Resolución con el software R:

Cuando se trabaja con la distribución hipergeométrica en "R", las sentencias a utilizar son: $dhyper(x, k, N-k, n), phyper(x, k, N-k, n)$, ó $qhyper(q, k, N-k, n)$, según si se quiere conocer la función de probabilidad en x o el valor de la función de distribución acumulada en x o un cuantil de la distribución hypergeométrica respectivamente.

a)

- $dhyper(x, k, N-k, n)$
- $dhyper(0, 5{,}15{,}2)$
- 0.5526316

$$P_1(X_1 = 0) = 0{,}5526316 \cong 0{,}55263$$

b)

- $dhyper(2{,}5{,}15{,}2)$
- 0.05263158

$$P_1(X_1 = 2) = 0{,}05263158 \cong 0{,}05263$$

Si se obtiene dos componentes defectuosos se rechaza el lote, con una probabilidad de 0,05263

c) Plan de muestreo doble:

1era etapa:

- dhyper (1, 5, 15, 2)
- 0.3947368

$$P_1(X_1 = 1) = 0{,}3947368 \cong 0{,}39474$$

2da etapa

- dhyper (0, 4, 14, 1)
- 0.7777778

$$P_2(X_2 = 0) = 0{,}7777778 \cong 0{,}77778$$

- dhyper(0,5,15,2)+dhyper(1,5,15,2)*dhyper(0,4,14,1)
- 0.8596491

$$P(Aceptar) \cong 0{,}8596491 \cong 0{,}85964$$

4-2-5 MODELO GEOMÉTRICO

La distribución geométrica es un modelo adecuado para aquellos procesos en los que se repiten los ensayos hasta la obtención del resultado deseado.

El experimento tiene sólo dos resultados mutuamente excluyentes, es decir los resultados del ensayo son dicotómicos. Los ensayos son independientes entre sí. Llámese éxito a la ocurrencia del suceso y fracaso a la no ocurrencia. Sea p la probabilidad de éxito del evento y $(q = 1 - p)$ la probabilidad de fracaso. Se realizan los ensayos idénticos e independientes hasta que ocurre el primer éxito. Las probabilidades p y q son constantes en todas las pruebas, por lo tanto las pruebas son independientes. Por lo que el parámetro en la distribución geométrica es p, (probabilidad de éxito).

Este proceso consta de un número no definido de pruebas o ensayos. Este proceso concluye cuando se obtiene el primer éxito.

La variable aleatoria es el número de ensayos realizados hasta obtener el primer éxito.

Se observa que esta variable es muy parecida a la distribución binomial, la diferencia esta que en vez de fijar el número de ensayos, n, y observar el número de éxitos, en esta distribución geométrica se realizan las pruebas hasta que se obtiene el primer éxito.

$$X \sim G(x, p)$$

El experimento podría realizarse una vez y obtener un éxito, y terminar aquí; o podría seguir indefinidamente.

Características de una variable con distribución Geométrica:

1. Es una variable aleatoria discreta.
2. Cada ensayo es una variable Bernoulli, es decir tiene dos posibles resultados, mutuamente excluyentes, "éxito" o "fracaso", con probabilidades p y $(q = 1 - p)$ respectivamente.
3. Se realizan ensayos estadísticamente independientes. Como los ensayos son independientes entonces la probabilidad de éxito p, permanece constante durante los distintos ensayos.
4. Se realizan los ensayos hasta obtener un éxito.
5. La variable aleatoria geométrica X, representa el número de pruebas hasta obtener el primer éxito inclusive.

Obtención de la función de probabilidad o función de cuantía

El espacio muestral Ω para el experimento contiene el siguiente conjunto infinito numerable de elementos:

$$\Omega = \left\{E_1, E_2, \ldots, E_k\right\}$$

Siendo:[1]

E_1: E (éxito en la primer prueba)
E_2: (F,E) (fracaso en la primer prueba y éxito en la segunda)
E_3: (F,F,E) (fracasos en la primera y segunda prueba, éxito en la tercera)
E_4: (F,F,F,E) (fracasos en la primera, segunda y tercera prueba, éxito en la cuarta)
....

$E_k : (\underbrace{F, F, F, \ldots., F}_{k-1}, E)$ (fracasos en las k-1 pruebas y éxito en la prueba k)

[1] "Estadística. Notas de clases". Universidad Nacional de Cuyo. Facultad de Ciencias Económicas. Año 2012. Pág. 55.

La variable geométrica definida como: X: “El número de pruebas hasta obtener el primer éxito inclusive”

Entonces si X=1, significa que en la primera prueba se obtuvo el éxito, si X=2, significa que se realizó dos pruebas, un fracaso y el éxito, si X=3 hay tres pruebas, las dos primeras fracaso y la tercera éxito, en general para X=x significa que se realizaron x pruebas, (x-1) fracasos y el éxito. Entonces:

$$P(X=x)=P(E_x)=P(\underbrace{F\cap F\cap F\cap\ldots\cap F}_{x-1}\cap E)$$

Como los ensayos son independientes tenemos:

$$P(X=x)=P(E_x)=\underbrace{P(F).P(F)\ldots.P(F)}_{x-1}P(E)$$

$$Como \qquad P(E)=p, \quad P(F)=1-p$$

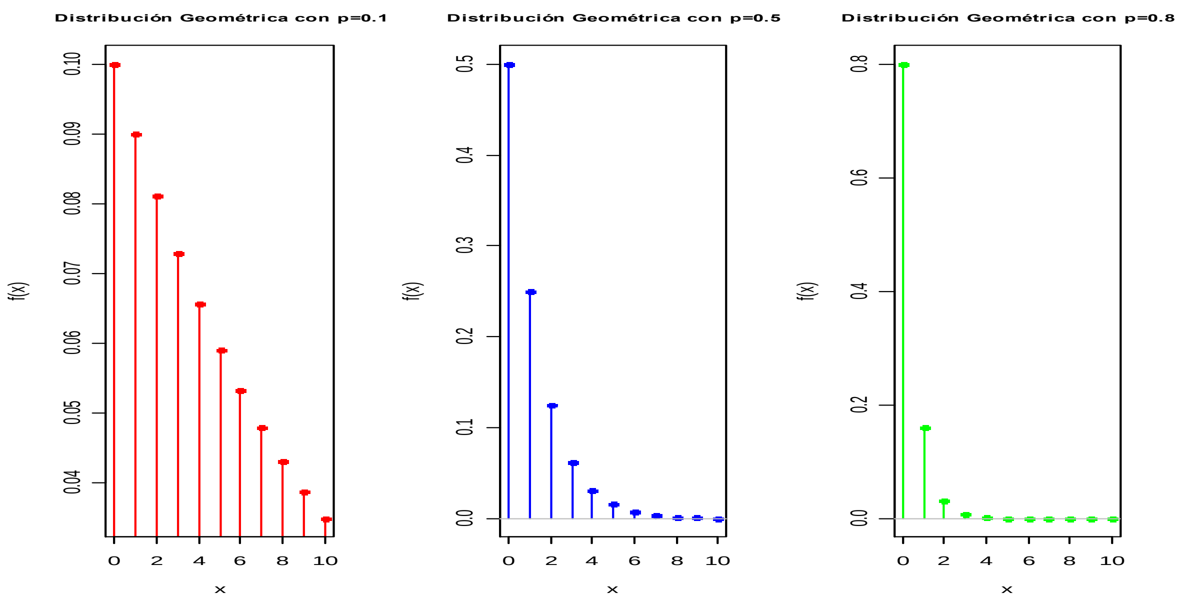
$$P(X=x)=P(E_x)=\underbrace{(1-p)\ldots(1-p)}_{x-1}p$$

$$f_X(x)=P(X=x)=P(E_x)=\underbrace{(1-p)\ldots(1-p)}_{x-1}p$$

La probabilidad de la intersección de x eventos independientes da lugar a la distribución geométrica con función de densidad de probabilidad dada por:

$$f_X(x)=\begin{cases}(1-p)^{x-1}p & siendo\ x=1,2,3,\ldots;0\le p\le 1\\ 0 & en\ cualquier\ otro\ caso\end{cases}$$

Representación gráfica de la función de probabilidad de una variable aleatoria con distribución geométrica:



Función de la distribución acumulada de probabilidad

En base a la función de probabilidad se obtiene la función de distribución acumulada.

$$F(k) = \sum_{x=1}^{k} q^{x-1}p\,, \quad para\; x = 1, 2, 3, \dots$$

Desarrollando la expresión tenemos:

$$F(k) = P(X \leq k) = p\left(1 + q + q^2 + \cdots + q^{k-1}\right) = p.\frac{1 - q^k}{1 - q}$$

Dado que $p = 1 - q$, tenemos:

$$F(k) = P(X \leq k) = \begin{cases} 0 \qquad , & si\; x \leq 0 \\ 1 - q^k, & si\; x \geq 1 \end{cases}$$

4-2-5-1 FUNCIÓN GENERADORA DE MOMENTO DE UNA VARIABLE GEOMÉTRICA

$$m_X(t) = E(e^{tX})$$

$$m_X(t) = \sum_{x \epsilon X(\Omega)} e^{tx} f_x(x)$$

$$m_X(t) = \sum_{x \epsilon X(\Omega)} e^{tx} . p . q^{x-1}$$

$$m_X(t) = \frac{p}{q} \sum_{x=1}^{\infty} (e^t q)^x$$

$$m_X(t) = \lim_{x \to \infty} \frac{p}{q} (e^t q + (e^t q)^2 + (e^t q)^3 + \cdots + (e^t q)^x)$$

$$m_X(t) = \frac{p}{q} \lim_{x \to \infty} e^t q . (1 + (e^t q)^1 + (e^t q)^2 + \cdots + (e^t q)^{x-1})$$

$$m_X(t) = \frac{p}{q} \lim_{x \to \infty} \frac{(e^t q - (e^t q)^{x+1})}{1 - e^t q}$$

$$Si \quad |e^t q| < 1, entonces\ (e^t q)^x \to 0, \text{luego}$$

$$m_X(t) = \frac{p}{q} . \frac{e^t q}{(1 - e^t q)}$$

Simplificando q del numerador, con el denominador, se obtiene la función generadora de momentos de una variable con distribución geométrica.

$$m_X(t) = \frac{pe^t}{1 - e^t q}$$

4-2-5-2 ESPERANZA DE UNA VARIABLE GEOMÉTRICA

A partir de la función generadora de momentos se puede obtener la esperanza haciendo la derivada primera de dicha función y luego evaluándola en $t = 0$.

Se simplifica la función generadora de momentos para derivar una expresión más simple.

$$m_X(t) = \frac{pe^t}{1-e^tq} = \frac{p}{e^{-t}(1-e^tq)} = \frac{p}{e^{-t}-q} = p(e^{-t}-q)^{-1}$$

$$m'_X(t) = \frac{d(m_X(t))}{dt} = \frac{d(p(e^{-t}-q)^{-1})}{dt}$$

$$m'_X(t) = (-1).p.(e^{-t}-q)^{-2}.(-e^{-t})$$

$$m'_X(t) = p.(e^{-t}-q)^{-2}.(e^{-t})$$

$$En\ t = 0 \quad entonces \quad m'_X(t=0) = p(1-q)^{-2} = p(p)^{-2} = \frac{p}{p^2} = \frac{1}{p}$$

Luego la esperanza de una variable con distribución geométrica es:

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

4-2-5-3 VARIANZA DE UNA VARIABLE GEOMÉTRICA

Considerando la fórmula de varianza de una variable aleatoria tenemos:

$$Var(X) = E(X^2) - [E(X)]^2$$

Se calcula primeramente la esperanza de la variable al cuadrado $E(X^2)$, la cual se obtiene a partir de la función generadora de momentos. Como se acaba de hacer con la esperanza, se calcula la derivada segunda de la función generadora de momentos y luego se evalúa en $t = 0$.

En símbolos: $m''_X(t=0) = \mathrm{E}(\mathrm{X}^2)$

Siendo la función generadora de momentos de una variable con distribución geométrica:

$$m_X(t) = \frac{pe^t}{1-e^tq} = p(e^{-t}-q)^{-1}$$

Su derivada primera es:

$$m'_X(t) = \frac{d(m_X(t))}{dt} = p.(e^{-t}-q)^{-2}.(e^{-t})$$

Se vuelve a derivar para obtener la derivada segunda:

$$m''_X(t) = p.(-2)(e^{-t}-q)^{-3}.(e^{-t})^2(-1) + p(e^{-t}-q)^{-2}(e^{-t})(-1)$$

$$m''_X(t) = 2p.(e^{-t}-q)^{-3}.(e^{-t})^2 - p(e^{-t}-q)^{-2}(e^{-t})$$

$$m''_X(t=0) = 2p.(1-q)^{-3} - p(1-q)^{-2}$$

$$m''_X(t=0) = 2p.p^{-3} - pp^{-2} = 2.p^{-2} - p^{-1}$$

$$m''_X(t=0) = \frac{2}{p^2} - \frac{1}{p}$$

$$m''_X(t=0) = \mathrm{E}(\mathrm{X}^2) = \frac{2}{p^2} - \frac{1}{p}\ ; \quad (4.5.1)$$

Recuerde la fórmula de varianza:

$$Var(X) = E(X^2) - [E(X)]^2\ ; \quad (4.5.2)$$

A continuación se calcula la $[E(X)]^2$

$$E(X) = \frac{1}{p}\ \rightarrow [E(X)]^2 = \frac{1}{p^2}\ \ (4.5.3)$$

Se reemplaza (4.5.1) y (4.5.3) en la fórmula de la varianza (4.5.2), queda:

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

$$\sigma_X^2 = Var(X) = \frac{q}{p^2}$$

Ejemplo 4-13:
La probabilidad de que un jugador de futbol marque en un penal es de 0,8.

a) ¿Cuál es la probabilidad de que un jugador de futbol marque por primera vez en su quinto penal?,
b) ¿Cuál es el número esperado de penales que necesita para marcar su primer gol?
c) ¿Cuál es la probabilidad de que un jugador de futbol marque por primera vez antes de su quinto penal?

Sea X: "Números de penales que necesita hasta marcar su primer gol"

$$X \sim G(x, p = 0{,}8)$$

a)

$$f(5) = P(X=5) = (1-p)^4.p$$

$$P(X=5) = (0{,}2)^4.0{,}8 = 1{,}28.10^{-3}$$

La probabilidad de que un jugador de futbol realice un gol por primera vez en su quinto penal es de $1{,}28.10^{-3}$.

b)

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

$$\mu = \frac{1}{0{,}8} = 1{,}25$$

Se espera que necesite en promedio 1,25 penales para marcar su primer gol.

c)

$$P(X<5)=P(X\leq 4)=F(4)$$
$$F(4)=P(X\leq 4)=P(X=1)+P(X=2)+P(X=3)+P(X=4)$$
$$F(4)=P(X\leq 4)=0{,}8+0{,}2.\,0{,}8+0{,}2^2.\,0{,}8+0{,}2^3.\,0{,}8=0{,}9984$$

La probabilidad de que un jugador de futbol marque por primera vez antes de su quinto penal es de 0,9984.

Resolución con el software R:

Cuando se trabaja con la distribución geométrica en "R", las sentencias a utilizar son: $dgeom(x-1,p), pgeom(x-1,p)$, ó $qgeom(q,p)$, según si se quiere conocer la función de probabilidad en x o el valor de la función de distribución acumulada en x o un cuantil de la distribución hypergeométrica respectivamente.

Entonces:

a) Sea X: "Números de penales que necesita hasta marcar su primer gol"

$$X \sim G(x,p=0{,}8)$$

```
>dgeom(5-1,0.8)
[1] 0.00128
```

$$f(5)=P(X=5)=0{,}00128=1{,}28.\,10^{-3}$$

b) Como

$$E(X)=\frac{1}{p}=\frac{1}{0{,}8}$$

```
> 1/0.8
[1] 1.25
```

c)

$$P(X<5)=P(X\leq 4)=F(4)$$

```
>pgeom(x-1, p)
> pgeom(3,0.8)
[1] 0.9984
```

$$F(4)=0{,}9984$$

Ejemplo 4-14:
Un Instituto que supervisa los laboratorios en base a los requisitos de la Norma IRAM 301, dedicado a calibrar instrumentos de medición, estima que la probabilidad de que una balanza de medición electrónica muestre una desviación excesiva es de 0,05. ¿Cuál es la probabilidad de que:

a) la sexta de estas balanzas de medición sometidas a prueba de calibración sea la primera en mostrar una desviación excesiva?
b) la quinta de estas balanzas de medición sometidas a prueba, sea la primera que no muestre una desviación excesiva?
Solución:

a) Sea: X: “Número de calibraciones de balanzas electrónicas hasta encontrar una que compute una con desviación excesiva”

$$X \sim G(x, p = 0{,}05)$$

$$f(6) = P(X = 6) = (1 - p)^5 . p$$

$$P(X = 6) = (1 - 0{,}05)^5 . 0{,}05 = 0{,}95^5 . 0{,}05$$

$$P(X = 6) = 0{,}0387$$

La probabilidad de que en la sexta medición se encuentre la primera balanza electrónica que mida una desviación excesiva es de 0,0387.

b) Sea Y: “El número de calibraciones hasta encontrar la primera balanza electrónica que no mida una desviación excesiva”

$$Y \sim G(y, p = 0{,}95)$$

$$f(5) = P(Y = 5) = (1 - p)^4 . p$$

$$f(5) = P(Y = 5) = (1 - 0{,}95)^4 . 0{,}95$$

$$f(5) = P(Y = 5) = (0{,}05)^4 . 0{,}95 = 0{,}0000059$$

La probabilidad de que en la quinta medición se encuentre la primera balanza que no muestre una desviación excesiva es de 0,0000059.

Resolución con el software R:

a)

- dgeom(5,0.05)
- 0.03868905

$$P(X = 6) = 0{,}03868905 \cong 0{,}0387$$

b)

- dgeom(4,0.95)
- 5.9375e-06

$$P(Y = 5) = 0{,}0000059375 \cong 0{,}0000059$$

4-2-6 MODELO BINOMIAL NEGATIVO

La distribución binomial negativa es un modelo adecuado para tratar aquellos procesos en los que se repite una determinada prueba hasta obtener un número fijo de resultados favorables.

El experimento tiene dos resultados mutuamente excluyentes, la ocurrencia o no de un suceso, es decir los resultados del ensayo son dicotómicos, e independientes entre sí. Llámese éxito a la ocurrencia del suceso y fracaso a la no ocurrencia. Sea p la probabilidad de éxito del evento y $(1 - p = q)$ la probabilidad de fracaso. Se realizan los ensayos idénticos e independientes hasta que ocurre el k-ésimo éxito. Las probabilidades p y q son constantes en todas las pruebas, por lo tanto las pruebas son independientes.

Este proceso consta de un número no definido de pruebas o ensayos. Este proceso concluye cuando se obtiene k- éxitos.

Si el número de resultados favorables buscados fuera 1 estaríamos en el caso de la distribución geométrica. Por lo que se dice que la distribución geométrica es un caso especial de la distribución binomial negativa.

La variable aleatoria es el número de ensayos realizados hasta obtener el k-ésimo éxito.

Anotamos: $X \sim Binomial\ Negativa(x, k, p)$

Siendo k y p los parámetros de la distribución, k representa la cantidad de éxitos buscados, y p representa la probabilidad de obtener un éxito.

Esta distribución es llamada también distribución de Pascal.

Función de probabilidad o función de cuantía:

Para obtener la fórmula general de este modelo consideremos que tenemos el evento E_k con k éxitos:

$$E_k: \underbrace{\underbrace{FFF \ldots F}_{x-k} \overbrace{\underbrace{EE \ldots E}_{k-1}\, E}^{k \text{ éxitos}}}_{x\ veces}$$

$$E_k: F^{x-1}E^{k-1}E$$

Si se aplica probabilidad tenemos:

$$P(E_k) = P(\underbrace{F \cap F \cap F \cap \ldots \cap F}_{x-k} \cap \underbrace{E \cap \ldots \cap E \cap E}_{k \text{ éxitos}})$$

Como los ensayos son independientes tenemos:

$$P(E_k) = \underbrace{P(F).P(F) \ldots .P(F)}_{x-k} . \underbrace{P(E) \ldots P(E).P(E)}_{k \text{ éxitos}}$$

$$Como \qquad P(E) = p, \quad P(F) = 1 - p = q$$

$$P(E_k) = \underbrace{\underbrace{(1-p) \ldots (1-p)}_{x-k} \underbrace{p \ldots p.p}_{k \text{ éxitos}}}_{x\ veces}$$

$$P(E_k) = (1-p)^{x-k} p^k \qquad (4.6.1)$$

Obtuvimos la $P(E_k)$, Qué es la probabilidad del suceso E_k, que sería la probabilidad de x, ($f(x) = P(X = x)$), si el suceso podría aparecer sólo en ese orden. Dado que puede aparecer en otros órdenes debemos considerar todas las formas posibles en que aparecen los (x-k) fracasos y los (k-1) éxitos, dado que el último siempre será un éxito.

La cantidad de órdenes distintos que pueden aparecer (x-k) fracasos con (k-1) éxitos, se obtiene calculando el número combinatorio:

$$\binom{x-1}{k-1}$$

Completando la relación (4.6.1) obtenemos la función de probabilidad o función de cuantía:

$$f(x) = \begin{cases} \binom{x-1}{k-1}(1-p)^{x-k}p^k; & si \quad x = k, k+1, k+2, \ldots \ldots \\ 0 & en\ cualquier\ otro\ caso \end{cases}$$

Características de una variable con distribución Binomial Negativa:

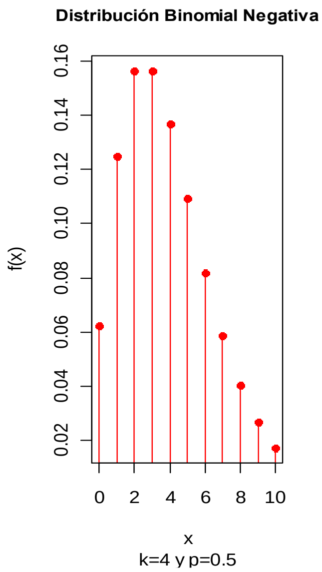
1. Es una variable aleatoria discreta.
2. Cada ensayo es una variable Bernoulli, es decir tiene dos posibles resultados, mutuamente excluyentes, "éxito" o "fracaso", con probabilidades p y (1-p) respectivamente.
3. Se realizan ensayos estadísticamente independientes. Como los ensayos son independientes entonces la probabilidad de éxito p, permanece constante durante los distintos ensayos.
4. Se realizan los ensayos hasta obtener k éxito.
5. La variable aleatoria geométrica X, representa el número de pruebas hasta obtener el k-ésimo éxito inclusive.
6. Su función de probabilidad o función de cuantía es:

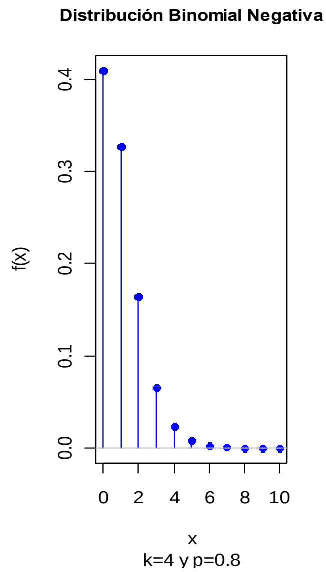
$$f(x) = \begin{cases} \binom{x-1}{k-1}(1-p)^{x-k}p^k; & si \quad x = k, k+1, k+2, \ldots \ldots \\ 0 & en\ cualquier\ otro\ caso \end{cases}$$

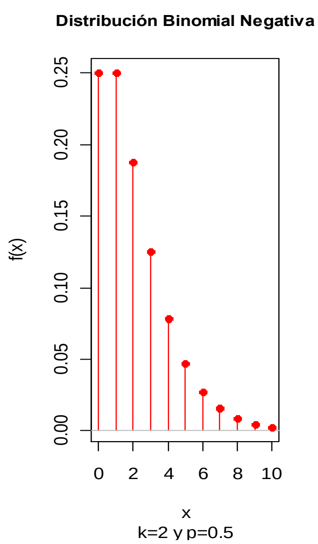
Si X es una variable que cumple con estas características decimos que se distribuye como una Binomial Negativa, con parámetros, p y k

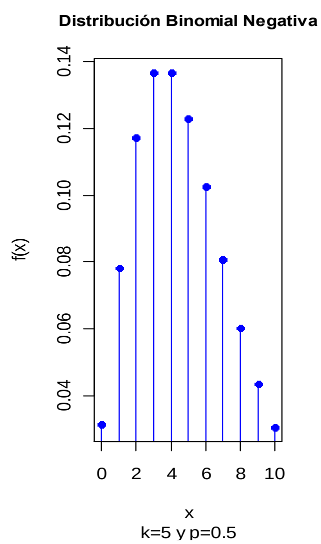
En símbolos $X \sim BN\ (p,\ k)$

Representación gráfico de la función de probabilidad de una variable con distribución binomial negativa:









4-2-6-1 FUNCIÓN GENERADORA DE MOMENTO DE UNA VARIABLE BINOMIAL NEGATIVA

$$m_X(t) = E(e^{tX})$$

$$m_X(t) = \sum_{x\epsilon X(\Omega)} e^{tx} f_x(x)$$

$$m_X(t) = \sum_{x\epsilon X(\Omega)} e^{tx}.\binom{x-1}{k-1}(1-p)^{x-k}p^k$$

$$m_X(t) = \frac{p^k}{q^k}\sum_{x=k}^{\infty}\binom{x-1}{k-1}(e^t q)^x$$

$$m_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^k$$

4-2-6-2 ESPERANZA DE UNA VARIABLE BINOMIAL NEGATIVA

A partir de la función generadora de momentos se puede obtener la esperanza haciendo la derivada primera de dicha función y luego evaluándola en $t = 0$.

$$m_X(t) = \left(\frac{pe^t}{1-qe^t}\right)^k = \frac{p^k.e^{tk}}{(1-qe^t)^k}$$

$$m'_X(t) = \frac{d(m_X(t))}{dt} = \frac{p^k.e^{tk}.k.(1-qe^t)^k - p^k.e^{tk}.k.(1-qe^t)^{k-1}.(-qe^t)}{(1-qe^t)^{2k}}$$

$$m'_X(t) = \frac{p^k.e^{tk}.k.(1-qe^t)^{k-1}.[(1-qe^t)+qe^t]}{(1-qe^t)^{2k}}$$

$$m'_X(t) = \frac{p^k.e^{tk}.k.[(1-qe^t)+qe^t]}{(1-qe^t)^{k+1}} = \frac{p^k.e^{tk}.k}{(1-qe^t)^{k+1}}$$

$$m'_X(t=0) = \frac{p^k.k}{(1-q)^{k+1}} = \frac{p^k.k}{p^{k+1}} = \frac{k}{p}$$

$$E(X) = \frac{k}{p}$$

Obtuvimos la esperanza de la variable X con parámetros k y p.

4-2-6-3 VARIANZA DE UNA VARIABLE BINOMIAL NEGATIVA

Considerando la fórmula de varianza de una variable aleatoria tenemos:

$$Var(X) = E(X^2) - [E(X)]^2$$

Se calcula primeramente la esperanza de la variable al cuadrado $E(X^2)$, la cual se obtiene a partir de la función generadora de momentos. Como se acaba de hacer con la esperanza, se calcula la derivada segunda de la función generadora de momentos y luego se evalúa en $t = 0$.

En símbolos: $m''_X(t = 0) = \mathrm{E}(\mathrm{X}^2)$

Siendo la función generadora de momentos de una variable con distribución binomial negativa:

$$m_X(t) = \left(\frac{pe^t}{1 - qe^t}\right)^k = \frac{p^k . e^{tk}}{(1 - qe^t)^k}$$

Su derivada primera es:

$$m'_X(t) = \frac{d(m_X(t))}{dt} = \frac{p^k . e^{tk} . k}{(1 - qe^t)^{k+1}}$$

Se vuelve a derivar para obtener la derivada segunda:

$$m''_X(t) = \frac{k^2 . p^k e^{tk} . (1 - qe^t)^{k+1} - k . p^k e^{tk} . (k + 1) . (1 - qe^t)^k . (-qe^t)}{(1 - qe^t)^{2k+2}}$$

$$m''_X(t) = \frac{k . p^k e^{tk} . (1 - qe^t)^k . [k(1 - qe^t) - (k + 1) . (-qe^t)]}{(1 - qe^t)^{2k+2}}$$

$$m''_X(t) = \frac{k . p^k e^{tk} . [k - kqe^t + kqe^t + qe^t]}{(1 - qe^t)^{k+2}}$$

$$m''_X(t) = \frac{k . p^k e^{tk} . [k + qe^t]}{(1 - qe^t)^{k+2}}$$

Evaluamos la derivada segunda en t=0

$$m''_X(t = 0) = \frac{k . p^k . [k + q]}{(1 - q)^{k+2}}$$

$$m''_X(t = 0) = \frac{k^2 . p^k + k . p^k . q]}{(1 - q)^{k+2}} = \frac{p^k(k^2 + k . q)}{p^{k+2}} = \frac{k^2 + k . q}{p^2}$$

$$m''_X(t = 0) = E(X^2) = \frac{k^2 + k . q}{p^2}$$

Recuerde ahora la fórmula de varianza:

$$Var(X) = E(X^2) - [E(X)]^2$$

A continuación se calcula la $[E(X)]^2$, dado que:

$$E(X) = \frac{k}{p}$$

Entonces tenemos:

$$[E(X)]^2 = \left[\frac{k}{p}\right]^2$$

Luego reemplazando en la fórmula de Varianza:

$$Var(X) = E(X^2) - [E(X)]^2$$

$$Var(X) = \frac{k^2 + k.q}{p^2} - \left[\frac{k}{p}\right]^2 = \frac{k^2 + k.q - k^2}{p^2} = \frac{k.q}{p^2}$$

$$Var(X) = \frac{k.q}{p^2}$$

Ejemplo 4-15:

Consideremos la empresa dedicada a calibrar instrumentos de medición del ejercicio 4-14, esta estima que la probabilidad de que una balanza de medición electrónica muestre una desviación excesiva es de 0,05.

1. ¿Cuál es la probabilidad de que tenga que calibrar ocho de estas balanzas electrónicas, para encontrar tres que midan una desviación excesiva?
2. ¿Cuántas balanzas se espera que sean necesarias calibrar en promedio hasta obtener tres mediciones que muestren una desviación excesiva? ¿Con qué desviación estándar?
3. ¿Cuál es la probabilidad de que la quinta calibración de estas balanzas electrónicas sometidas a prueba, sea la segunda que no muestre una desviación excesiva?
4. ¿Cuál es la probabilidad de que tenga que calibrar menos de cuatro balanzas electrónicas para encontrar dos con una desviación excesiva?

Solución:

1. Sea: X: "Número de calibraciones de balanzas electrónicas necesarias hasta obtener tres con desviaciones excesivas"

$$X \sim BN(p;\ k)$$

$$X \sim BN(p = 0{,}05; k = 3)$$

$$f(x) = P(X = x) = \binom{x-1}{k-1}(1-p)^{x-k}p^k$$

$$f(8) = P(X = 8) = \binom{8-1}{3-1}(1-0{,}05)^{8-3}0{,}05^3$$

$$f(8) = P(X = 8) = \binom{7}{2}(0{,}95)^5 0{,}05^3$$

$$f(8) = P(X = 8) = \binom{7}{2}(0{,}95)^5 0{,}05^3$$
$$f(8) = P(X = 8) = 0{,}00203$$

La probabilidad de que la tercera desviación excesiva se obtenga al calibrar la octava balanza electrónica es de 0,00203.

2. Sea: X: “Número de calibraciones de balanzas electrónica necesarias hasta obtener tres con desviaciones excesivas”

$$X \sim BN(p = 0{,}05; k = 3)$$

$$E(X) = \frac{k}{p}$$
$$E(X) = \frac{3}{0{,}05} = 60$$

Se espera necesitar calibrar en promedio 60 balanzas electrónicas para encontrar tres balanzas con desviación excesiva.
Calculemos ahora la varianza y desviación estándar de la variable aleatoria X

$$Var(X) = \frac{k.q}{p^2}$$
$$Var(X) = \frac{k.(1-p)}{p^2} = \frac{3.0{,}95}{0{,}05^2} = 1.140$$

$$DES(X) = \sqrt{Var(X)} = \sqrt{1140} = 33{,}7639$$

Se espera necesitar calibrar en promedio 60 balanzas electrónicas para encontrar tres balanzas con desviación excesiva, con una desviación estándar de aproximadamente 33.76 calibraciones.

3.

Sea: X: “Número de calibraciones de balanzas electrónicas necesarios hasta obtener dos desviaciones no excesivas”

$$X \sim BN(p = 0{,}95; k = 2)$$

$$f(x) = P(X = x) = \binom{x-1}{k-1}(1-p)^{x-k}p^k$$
$$f(5) = P(X = 5) = \binom{4}{1}(0{,}05)^3 0{,}95^2$$

$$f(5) = P(X = 5) = 4.(0{,}05)^3 0{,}95^2 = 0{,}000451$$

La probabilidad de que la segunda desviación no excesiva se obtenga al calibrar la quinta balanza electrónica es de 0,0086.

4. Sea X: “ Número de calibraciones de balanzas de mediciones electrónica necesarias para obtener dos con desviaciones excesivas”

$$X \sim BN(p = 0{,}05; k = 2)$$

$$P(X < 4) = F(3) = P(X \leq 3)$$

Como la variable puede tomar como mínimo el valor k, en este caso tenemos $x = k, k+1, es\ decir\ x = 2{,}3$

$$P(X < 4) = F(3) = P(X \leq 3)$$

$$P(X < 4) = P(X = 2) + P(X = 3)$$

$$P(X < 4) = \binom{x-1}{k-1}(1-p)^0 p^2 + \binom{x-1}{k-1}(1-p)^{3-2} p^2$$

$$P(X < 4) = (0{,}95)^0 . 0{,}05^2 + \binom{2}{1}(0{,}95)^{3-2} 0{,}05^2$$

$$P(X < 4) = 0{,}0025 + 0{,}0048 = 0{,}0073$$

Interpretación: La probabilidad de tener que calibrar menos de cuatro balanzas electrónicas para encontrar dos balanzas con una desviación excesiva es de 0,0073.

Resolución con el software R:

Si resolvemos el ejercicio utilizando el software “R” tenemos que utilizar el comando “nbinom”, cuando trabajamos con una variable con distribución binomial negativa. Siendo las sentencias a utilizar:

- dnbinom(x-k,k,p), la función de probabilidad en el punto x
- pnbinom(x-k,k,p), la función de distribución acumulada en el punto x

1. Sea: X: “Número de calibraciones de balanzas electrónicas necesarias hasta obtener tres con desviaciones excesivas”

$$X \sim BN(p;\ k)$$
$$X \sim BN(p = 0{,}05; k = 3)$$

- dnbinom(x-k, k, p)

```
> dnbinom(5, 3, 0.05)
[1] 0.002031175
```

Luego

$$f(8) = P(X = 8) = 0{,}002031175 \cong 0{,}00203$$

2. Sea: X: “Número de calibraciones de balanzas electrónica necesarias hasta obtener tres con desviaciones excesivas”

$$X \sim BN(p = 0{,}05; k = 3)$$

Cálculo de Esperanza con R:

- k/p

>3/0.05
[1] 60

Cálculo de Varianza con R:

- *k*q/p^2*

> 3*0.95/0.05^2
[1] 1140

Cálculo de Desviación Estándar con R:
> sqrt(1140)
[1] 33.76389

3. Sea: X: “Número de calibraciones de balanzas electrónicas necesarios hasta obtener dos desviaciones no excesivas”

$$X \sim BN(p = 0,95; k = 2)$$

- dnbinom(x-k, k, p)

> dnbinom(3, 2, 0.95)
[1] 0.00045125
Luego:

$$f(5) = 0,00045125 \cong 0,000451$$

4. Sea X: “Número de calibraciones de balanzas de mediciones electrónica necesarias para obtener dos con desviaciones excesivas”

$$X \sim BN(p = 0,05; k = 2)$$

$$P(X < 4) = F(3) = P(X \leq 3)$$

- pnbinom(x-k, k, p)

> pnbinom(1,2,0.05)
[1] 0.00725
Luego:

$$P(X \leq 3) = 0,00725 \cong 0,0073$$

4-2-7 MODELO MULTINOMIAL

Es una generalización de la distribución binomial, que surge cuando cada ensayo tiene más de dos resultados posibles. Por ejemplo cuando se estudia la calidad de un producto manufacturado, el cual se puede clasificar en superior, normal e inferior.

Se realizan n ensayos independientes, cada uno admite k resultados mutuamente excluyentes, $E_1, E_2, \ldots, E_k$ con probabilidad $p_1,\ p_2\ , \ldots, p_k$ respectivamente. La distribución multinomial dará la probabilidad de que los sucesos E_i ocurran x_i veces, en n pruebas independientes, con $i = 1,2, \ldots, k$, de donde se obtiene que:

$$\sum_{i=1}^{k} p_i = 1\,;\ \ \sum_{i=1}^{k} x_i = n$$

Siendo n como se aclaró anteriormente es el número de ensayos realizados independientes.

Características de una variable con distribución Multinomial:

1. Es una variable aleatoria discreta.
2. La variable tiene más de dos tipos de resultados, tiene k resultados mutuamente excluyentes posibles.
3. Se realizan n ensayos estadísticamente independientes. Como los ensayos son independientes entonces las probabilidades de éxito p_i, con $i = 1,2, \ldots, k$ permanecen constantes durante los distintos ensayos. Evidentemente se cumple que: $\sum_{i=1}^{k} p_i = 1$.
4. El número de repeticiones de los ensayos es constante, n.
5. La distribución Multinomial tiene entonces como parámetros, la cantidad de ensayos que se realizan n, y las probabilidades de los distintos sucesos, p_i con $i = 1,2, \ldots, k-1$ dado que se debe cumplir la relación $\sum_{i=1}^{k} p_i = 1$ entonces p_k,

La variable X_i se define como el número de veces que se ha presentado el evento E_i en las n pruebas, para $i = 1,2, \ldots, k$.

Nótese que si consideramos cada una de estas variables X_i, para $i = 1,2, \ldots, k$, por separado, entonces cada una de ellas tiene distribución binomial con parámetros n, p_i.

Para obtener la función de probabilidad se analizará igual que en la distribución binomial. Considerando la independencia de los ensayos, cualquier orden especificado que se produzca x_i resultados para el evento E_i, para $i = 1,2, \ldots, k$, ocurrirá con una probabilidad $p_1^{x_1}.p_2^{x_2} \ldots p_k^{x_k}$.

Para determinar el número total de órdenes en el experimento que tiene x_i éxitos para el evento E_i, para $i = 1,2, \ldots, k,$ es igual al número de particiones de n ensayos en k grupos con x_1 éxitos en el primer grupo, x_2 éxitos en el segundo grupo,…., x_k éxitos en el k-ésimo grupo. Esto se puede realizar de

$$\binom{n}{x_1, x_2, \ldots, x_k} = \frac{n!}{x_1!\,x_2! \ldots x_k!}$$

maneras, como todas las particiones son mutuamente excluyentes y ocurren con la misma probabilidad para un orden específico, se obtiene la distribución multinomial al multiplicar la probabilidad para un orden específico por el número total de particiones.

Entonces la función de probabilidad de la distribución multinomial está dada por:

$$f(x) = \begin{cases} \dfrac{n!}{x_1!\,x_2!\dots x_k!} p_1^{x_1}.p_2^{x_2}\dots p_k^{x_k}\,; & si \displaystyle\sum_{i=1}^{k} x_i = n; \quad con\; x_i = 0,1,..,n \\ 0 & en\ cualquier\ otro\ caso \end{cases}$$

Decimos que el vector aleatorio $(X_1, X_2, \dots\dots X_{k-1}) \sim Multinomial\ (n, p_1 p_2, \dots, p_{k-1})$
En donde $x_k = n - x_1 - x_2 - \cdots - x_{k-1}$

Ejemplo 4-16:
A unas elecciones de presidente se presentaron cuatro partidos políticos, A, B, C, D. El A obtuvo el 37% de los votos, el B obtuvo el 31% de los votos, el C obtuvo el 20% de los votos y el D obtuvo el 12%. ¿Cuál es la probabilidad de que al elegir 6 ciudadanos aleatoriamente 3 hayan elegido el partido A, 2 al partido B, 1 al partido C?

Siendo los eventos: A: Ciudadano que votaron al partido político A
B: Ciudadano que votaron al partido político B
C: Ciudadano que votaron al partido político C
D: Ciudadano que votaron al partido político D

Siendo X_1: "Número de ciudadanos que votaron al partido A"
X_2: "Número de ciudadanos que votaron el partido B"
X_3: "Número de ciudadanos que votaron al partido C"
X_4: "Número de ciudadanos que votaron al partido D"
Siendo n: el número de veces que se ha realizado el ensayo independientemente, en este caso 6.

Siendo p_1: probabilidad de que un ciudadano vote por el partido A
p_2: probabilidad de que un ciudadano vote por el partido B
p_3: probabilidad de que un ciudadano vote por el partido C
p_4: probabilidad de que un ciudadano vote por el partido D

Siendo $p_1 = 0{,}37;\ p_2 = 0{,}31;\ p_3 = 0{,}20;\ p_4 = 0{,}12$
Debemos comprobar que se cumpla la condición:

$$\sum_{i=1}^{k} p_i = 1$$

$$0{,}37 + 0{,}31 + 0{,}20 + 0{,}12 = 1$$

Además se debe cumplir:

$$\sum_{i=1}^{k} x_i = n = 6$$

$$\sum_{i=1}^{4} x_i = 3 + 2 + 1 + 0 = 6$$

La función de probabilidad conjunta de la distribución multinomial es:

$$P(X_1 = x_1 \cap \ X_2 = x_2 \cap X_3 = x_3 \cap X_4 = x_4) = f(x_1, x_2, x_3, x_4) = \frac{n!}{x_1!\, x_2! \dots x_k!} p_1^{x_1}. p_2^{x_2} \dots p_k^{x_k}$$

Reemplazando por los valores obtenidos:

$$f(3,2,1,0) = \frac{6!}{3!\, 2!\, 1!\, 0!} 0{,}37^3. 0{,}31^2. 0{,}20^1. 0{,}12^0$$

$$f(3,2,1,0) = \frac{6!}{3!\, 2!\, 1!\, 0!} 0{,}37^3. 0{,}31^2. 0{,}20^1. 0{,}12^0$$

$$f(3,2,1,0) \cong 0{,}05841$$

Nota: Recuerde que $0! = 1$ por convención
Interpretación: La probabilidad de que al tomar 6 ciudadanos aleatoriamente, 3 hayan votado al partido A, 2 al partido B, 1 al partido C y ninguno al partido D es aproximadamente del 0,05841

4-2-7-1 ESPERANZA Y VARIANZA DE UNA VARIABLE MULTINOMIAL

Suponiendo que X_i, tiene distribución binomial con probabilidad de éxito p_i, tenemos:

$$E(X_i) = n. p_i, \quad i = 1,2, \dots, k$$

$$Var(X_i) = n. p_i(1 - p_i), \ \ i = 1,2, \dots, k$$

4-2-7-2 FUNCIÓN GENERADORA DE MOMENTOS DE UNA VARIABLE MULTINOMIAL

$$m_X(t) = E(e^{tX})$$

$$m_X(t) = E(e^{\sum_{i=1}^{k} t_i x_i})$$

$$m_X(t) = (p_1 e^{t_1} + p_2 e^{t_2} + \dots + p_k e^{t_k})^n$$

$$m_X(t) = \left(\sum_{i=1}^{k} p_i e^{t_i} \right)^n$$

4-8 PROBLEMAS PROPUESTOS SOBRE MODELOS DE DISTRIBUCIÓN DE VARIABLES ALEATORIAS DISCRETAS

4.1 En 14 experimentos que estudian las características de durabilidad de los circuitos integrados sometidos a esfuerzo, 10 utilizaban silicio en su fabricación y los otros 4 utilizan una aleación de silicio- germanio, (SiGe). Si dos de los experimentos al azar son suspendidos por razones económicas. ¿Cuál es la probabilidad de que:

a) ninguno de los experimentos suspendidos emplee silicio en su fabricación?
b) sólo uno de los experimentos suspendidos emplee silicio?
c) ninguno de los dos experimentos suspendidos utilicen una aleación de silicio - germanio?

4.2 Un empresario está interesado en estudiar la calidad de sus tubos LED, para lo cual toma como indicador la cantidad de tubos defectuosos encontrados en una muestra determinada. Una caja contiene 90 tubos LED de los cuales 18 son defectuosos. Se toma una muestra de 4 tubos sin reposición.

a) Enuncie la variable en estudio. Identifique la distribución de la variable y plantee sus parámetros.
b) ¿Cuál es la probabilidad de no obtener tubos LED defectuoso?
c) ¿Cuál es la probabilidad de obtener como máximo 3 tubos LED defectuosos?
d) ¿Cuál es la probabilidad de obtener más de 3 tubos LED defectuosos?
e) ¿Cuántos tubos LED defectuosos se espera obtener en la muestra?, ¿con qué desviación estándar?

4.3 Un jugador de baloncesto se dispone a tirar a un cesto hasta anotar. Si se supone que sus tiros son independientes y la probabilidad de anotar al cesto es de 0,7. ¿Cuál es la probabilidad de necesitar dos tiros?, ¿tres tiros?

4.4 Se sabe que el 3% de los alumnos que cursan ingeniería son muy inteligentes. Se toma una muestra de 100 alumnos de los 1300 que cursan regularmente, elegidos sin reposición. Se está interesado en conocer el número de alumnos que cursan ingeniería que son muy inteligentes.

a) Enuncie la variable en estudio. Identifique la distribución de la variable y plantee sus parámetros.
b) ¿Cuál es la probabilidad de elegir 2 alumnos muy inteligentes?
c) ¿Cuál es la probabilidad de elegir como máximo 3 alumnos muy inteligentes?
d) ¿Cuántos alumnos que cursan ingeniería se espera encontrar muy inteligentes?, ¿Con qué desviación estándar?

4.5 En un determinado banco se reciben en promedio 4 cheques sin fondo por día.

1- Determine el modelo de distribución de la variable aleatoria, justifique su elección. Obtenga el valor de los parámetros que la caracterizan.

2- ¿Cuál es la probabilidad de que se reciban:

a) 5 cheques sin fondo en un día determinado?
b) por lo menos 4 cheques sin fondo en un día determinado?
c) como máximo 6 cheques sin fondo en dos días consecutivos?
d) menos de 18 cheques sin fondo en una semana? (Considera que la semana consta de 5 días)

3- La probabilidad de que cierto tipo de batería para celulares falle en una prueba de 2 meses es del 10%.

a) Determine el modelo de distribución de la variable aleatoria, justifique su elección. Obtenga el valor de los parámetros que la caracterizan.
b) Encuentre la función de probabilidad. Grafíquela.
c) Encuentre la función de distribución de probabilidad acumulada. Grafíquela.
d) Hallar la probabilidad de que en una muestra de 12 baterías permanezcan encendidas durante 2 meses:
   1) Por lo menos 8
   2) A lo más 5
   3) Exactamente 9 lámparas
e) ¿Cuántas baterías para celulares se espera que permanezcan encendidas al cabo de 2 meses? ¿Con que desviación estándar?
f) Encuentre el coeficiente de asimetría.

4- La probabilidad de que un conductor apruebe el examen escrito para obtener su licencia de conducir vehículos pesados es de 0,6.

1- Determine el modelo de distribución de la variable aleatoria, justifique su elección. Obtenga el valor de los parámetros que la caracterizan.
2- Encuentre la función de probabilidad.
3- Encuentre la probabilidad de que un conductor elegido al azar:
   a) Apruebe el examen escrito en el tercer intento.
   b) Apruebe el examen antes del cuarto intento.
   c) Necesite más de un intento para aprobar el examen escrito.
4- ¿Cuántos exámenes escritos se espera que necesite realizar para aprobar y obtener su licencia de conducir vehículos pesados?
5- Considere la siguiente expresión en R: pgeom(4, 0.6), exprese que probabilidad representa y encuéntrela.

4.8 Una aseguradora de autos después de un estudio de sus datos plantea que la tasa de reclamos por indemnización es de 2,25 por semana.
Encuentre la probabilidad de que:

a) El número de reclamos sea de 2 en una semana determinada.
b) El número de reclamos sea de menos de 10 en una semana determinada.
c) El número de reclamos sea por lo menos de 7 en un período de cuatro semanas.
d) El número de reclamos este entre 4 y 6 (inclusivos) en un período de cuatro semanas.
e) ¿Cuántos reclamos se esperan recibir por semana?, ¿con qué desviación estándar?

4.9 En una cierta área del conurbano de la provincia de Buenos Aires, se da como causa del 70 % de los robos, la necesidad de dinero para comprar narcóticos. Si tomamos los próximos 20 robos reportados en esta área.

a) Enuncie la variable en estudio. Indique qué distribución de probabilidad empleará que exprese la situación. ¿Cuánto valen los parámetros de la distribución aplicada?

b) ¿Cuál es la probabilidad de que:

1. Exactamente 5 robos se debieran a la necesidad de dinero para comprar narcóticos?
2. A lo sumo 5 se debieran a la necesidad de dinero para comprar narcóticos?
3. El número de robos debido a la necesidad de dinero para comprar narcóticos se encuentre comprendido en el intervalo $E(X) \pm DE(X)$.

4.10 El esquema de aceptación para comprar grandes lotes de motores eléctricos es probar no más de 75 motores seleccionados al azar, y rechazar el lote si falla un motor eléctrico. Se conoce que la probabilidad de falla es 0,002. En cada uno de los ítems siguientes enuncie la variable aleatoria, su distribución y sus parámetros.

¿Cuál es la probabilidad de que:

a) se acepte un lote? (Enuncie la variable en estudio, su distribución y su o sus parámetros)

b) Se rechace el lote en la prueba 15? (Enuncie la variable en estudio, su distribución y su o sus parámetro)

c) Se rechace a lo sumo en la prueba 12?

4.11 La probabilidad de que una persona que vive en una ciudad cosmopolita tenga una notebook se estima en 0,4. Se selecciona al azar distintas personas de esta ciudad para consultarles si tienen notebook. Encontrar la probabilidad de que:

a) La décima persona consultada sea la quinta en tener una notebook.

b) La primera persona con notebook sea la quinta consultada.

c) La tercera persona con notebook sea la séptima consultada.

4.12 Entre los métodos de lucha contra el granizo en fruticultura, se utiliza un método directo de lucha contra este problema, que son los proyectiles. Los componentes de un proyectil se envían en lotes de 25 componentes. La inspección de control de calidad consiste en seleccionar 3 componentes de cada lote y si ninguno de ellos tienen defectos se acepta el lote.

a) Determine el modelo de distribución de la variable aleatoria, justifique su elección. Obtenga el valor de los parámetros que la caracterizan.

b) Encuentre la función de probabilidad y grafíquela.

c) ¿Cuál es la probabilidad de que se acepte el lote si este contiene 5 componentes defectuosos?

d) ¿Cuántos lotes se espera aceptar si se mantiene el número de componentes defectuosos?, ¿con que desviación estándar?

4.13 En el departamento de control de calidad de una industria metalúrgica se inspeccionan las unidades que provienen de una línea de ensamble. Se sabe que la proporción de defectuosas es del 5%.

a) ¿Cuál es la probabilidad de que la vigésima unidad inspeccionada sea la segunda que se encuentre defectuosa?
b) ¿Cuál es la probabilidad de que la tercera unidad inspeccionada sea la segunda que se encuentre defectuosa?
c) ¿Cuántas unidades se espera inspeccionar hasta encontrar la segunda defectuosa?, ¿con qué desviación estándar?

4.14 Una empresa exportadora de motores para autos, los embarca en lotes de 50. Antes de que tal cargamento sea aceptado, un inspector de control de calidad selecciona 5 motores y revisa estrictamente sus especificaciones. Si ninguno de los motores probados es defectuoso, el lote es aceptado. Si se encuentra uno o más defectuosos, se inspecciona el cargamento completo. Suponga que en realidad hay tres motores defectuosos en el lote. ¿Cuál es la probabilidad de que sea necesaria una inspección del 100%?

4.15 En un cierto cruce de dos accesos automovilísticos, se producen en promedio 3 accidentes por mes. Considere que el número de accidentes que ocurren en dicho cruce por mes es una variable aleatoria con distribución de Poisson.

a) Determine la variable bajo estudio y su función de probabilidad.
b) ¿Cuál es la probabilidad de que en un mes ocurran por lo menos 3 accidentes?
c) Indique el valor de la esperanza y el de la desviación estándar de la variable bajo estudio considerando como unidad de observación el mes.
d) Halle la probabilidad de que la variable difiera de la media en más de una unidad
e) ¿Cuál es la probabilidad de que en dos meses ocurran menos de 6 accidentes?
f) ¿Cuál es la probabilidad de que ocurra un accidente por semana?

4.16 En un cierto proceso de manufactura se sabe que, la probabilidad de obtener un producto defectuoso es 0,1. El departamento de control de calidad inspeccionan los productos terminados.

a) ¿Cuál es la probabilidad de que el quinto inspeccionado sea el primer defectuoso?
b) ¿Cuál es la probabilidad de que el décimo producto inspeccionado sea el segundo defectuoso?

4.17 Una empresa comercial analizó el número de pedidos recibidos por día para un cierto artículo de su stock y encontró que el número promedio es de 3,6 pedidos por día.

a) ¿Cuál es la probabilidad de que no se reciban más de 5 pedidos para dicho artículo en un día?
b) ¿Cuál es la probabilidad de que no se reciban pedidos en un determinado día para dicho artículo?
c) Si el número promedio de pedidos por día fuera 5, ¿sería mayor o menor la variabilidad de la distribución del número de pedidos por día, que si la media fuera 3,6 pedidos por día?

4.18 Un auditor considera que los documentos de una sociedad son satisfactorios solamente si no existen documentos con errores en una muestra obtenida de un lote de 100. Considere que la muestra la toma sin reposición.

a) Si toma una muestra de 25 documentos sin reposición, ¿cuál es la probabilidad de que el auditor concluya que los documentos de la sociedad son satisfactorios si en el lote existe un documento con error?
b) Si toma una muestra de 25 documentos sin reposición, ¿cuál es la probabilidad de que el auditor concluya que los documentos de la sociedad son satisfactorios si en el lote existen 10 documentos con errores?
c) Si toma una muestra de 35 documentos, ¿cuál es la probabilidad de que el auditor concluya que los documentos son satisfactorios si en el lote existen 10 documentos con error?
d) Es muy pequeña la probabilidad de que no existan documentos con errores en una muestra de tamaño 25, si existen 10 documentos con errores, ¿sería esta probabilidad menor con una muestra de tamaño mayor?

4.19 Un distribuidor de balanzas digitales ha determinado a partir de numerosos ensayos que el 5% de ellas no están bien calibradas, por lo que dice que no cumplen las especificaciones mínimas establecidas. Vende dichas balanzas en cajas de 50, garantizando que el 90% cumple las especificaciones mínimas. Se está interesado en estudiar la cantidad de balanzas digitales que se distribuyen que no cumplen con las especificaciones mínimas establecidas.

a) Enuncie la variable en estudio. La distribución de la variable y sus parámetros.
b) Calcular la probabilidad de que una caja determinada no cumpla con la garantía.
c) ¿Cuantas balanzas digitales se espera que no cumplan con las especificaciones mínimas?, ¿con qué desviación estándar?

4.20 El número de accidentes de trabajo que se producen en una fábrica por semana, siguen una ley de Poisson, tal que la probabilidad de que haya 5 accidentes es 5/4 de que haya 3.

a) Enuncie la variable en estudio.
b) Encuentre el valor del parámetro de la distribución de Poisson.
c) Sabiendo que la función de distribución acumulada de probabilidad en k es 0,87, encontrar el valor máximo de accidentes semanales, k.

4.21 El número de colisiones que se producen en una determinada esquina de la ciudad de Mendoza, sigue una distribución de Poisson por parámetro $\lambda = 0{,}5$, colisiones por mes.

a) ¿Cuál es la probabilidad de que se produzca una colisión en un determinado mes?
b) ¿Cuál es la probabilidad de que se produzca más de dos colisiones en un determinado mes?
c) ¿Cuál es la probabilidad de que se produzca más de dos colisiones en dos meses consecutivos?

4.22 En una central telefónica automática, el número de llamadas erróneas que se conectan en promedio es de 2 llamadas por día.

a) Defina la variable aleatoria en estudio. Indique que modelo utilizaría para la para la presente situación.
b) ¿Cuál es el número esperado y la desviación estándar de la variable aleatoria X?
c) Compruebe los resultados obtenidos en el inciso anterior utilizando la función generadora de momentos.
d) Para un día determinado ¿cuál es la probabilidad que se efectúen 4 conexiones erróneas?
e) ¿Cuál es el número mínimo de llamadas independientes que se requieren para asegurar con probabilidad de 0,9 que por lo menos una de las llamadas sea conectada erróneamente?
f) ¿Cuál es la probabilidad que se efectúen 6 conexiones erróneas en una semana determinada?

4.23 Un comerciante vende arandelas en paquetes de 200 y garantiza que a lo más 10% son defectuosas. Un comprador mayorista controla cada paquete extrayendo 10 arandelas sin reposición. Si la muestra no contiene arandelas defectuosas acepta el paquete, de otra manera lo rechaza. Hallar la probabilidad de que en este proceso cualquier paquete dado se rechace aun cuando satisfaga la garantía.

4.24 En una fábrica de baterías de automóviles, el director del sector de calidad afirma que la proporción de unidades defectuosas de su producción es del 6%. Un comprador toma una muestra de 16 unidades, las prueba y encuentra cuatro defectuosas. Aceptando lo que afirma el director de calidad de la empresa, ¿cuál es la probabilidad de este suceso? A partir de dicho valor, ¿puede concluirse que la empresa está equivocada?

4.25 Un vendedor de programas de computación tienen stock: 5 windows, 4 antivirus y 3 programas para realizar gráficos nivel profesional. Una empresa del medio decide comprar 6 programas para sus computadoras. Hallar la probabilidad de que compre 3 windows, 2 antivirus y 1 graficador. Suponga independencia entre los sucesos.

4.26 Los registros de una empresa que produce fuentes de alimentación conmutada, conoce por experiencia previa que la probabilidad de que una de sus fuentes de alimentación nueva requiera de revisión por mal funcionamiento es de 0,20.

a) El departamento de control de calidad revisa las fuentes de alimentación. ¿Cuál es la probabilidad de que la sexta fuente de alimentación controlada en el mes, sea la segunda en necesitar revisión?
b) ¿Cuántas fuentes de alimentación se espera que tengan que controlar en promedio para encontrar dos fuentes de alimentación con problemas de funcionamiento?, ¿con qué desviación estándar?
c) ¿Cuál es la probabilidad de que la octava fuente de alimentación producida por esta empresa en el mes, sea la tercera en necesitar revisión por mal funcionamiento?

d) ¿Cuántas fuentes de alimentación se espera que tengan que controlar en promedio para encontrar tres fuentes de alimentación con problemas de funcionamiento? ¿con qué desviación estándar?

4.27 Una empresa industrial utiliza un sistema de control para la verificación de la calidad de sus productos antes de ser enviados a sus compradores. El método utilizado es de doble etapa. Se preparan cajas de 30 productos para su embarque y se prueba una muestra tomada al azar de 3, para controlar si hay productos que no cumplen con los estándares de calidad de la empresa. Si el departamento de control, encuentra algún producto defectuoso en la muestra, la caja completa se regresa para su control. Si no se encuentra ningún producto defectuoso, la carga se envía a su destino.

a) Enuncie la variable en estudio, su distribución y sus parámetros.
b) ¿Cuál es la probabilidad de que sea enviada una caja que contiene 3 productos defectuosos?
c) ¿Cuál es la probabilidad de que una caja que contiene sólo un producto defectuoso sea devuelta para su control total?

4.28 Suponga que la empresa industrial del ejercicio anterior 4-27, decide cambiar el método de control de la calidad de sus productos antes de ser enviados. Según el nuevo procedimiento, un inspector toma un artículo al azar, lo inspecciona y luego lo repone en la caja, un segundo inspector realiza el procedimiento igualmente y por último un tercer inspector realiza el mismo procedimiento. La caja no se envía a destino si alguno de los tres productos es defectuoso.

a) Enuncie la variable en estudio, su distribución y sus parámetros.
b) ¿Cuál es la probabilidad de que sea enviada una caja que contiene 3 productos defectuosos?
c) ¿Cuál es la probabilidad de que una caja que contiene sólo un producto defectuoso sea devuelta para su control total?

4.29 En una pequeña ciudad se estimó que 800 de los 2000 votantes residentes, están en contra del impuesto a las ganancias para empleados en dependencia. Si se toma una muestra aleatoria de 85 votantes y se le pide su opinión con relación a este impuesto. ¿Cuál es la probabilidad de que por lo menos 20 estén a favor del nuevo impuesto?

4.30 Un cargamento con lotes de 45 calculadoras científicas cada uno, se consideran aceptables si contienen no más de dos defectuosos. El método para muestrear el lote consiste en seleccionar una muestra aleatoria de 4 calculadoras científicas. Se rechaza el lote si se encuentra una defectuosa. ¿Cuál es la probabilidad de que se encuentre exactamente una defectuosa en la muestra, si existen 3 defectuosas en el lote?

4.31 La probabilidad de que un estudiante de ingeniería apruebe su última materia es de 0,6.

a) Encuentre la probabilidad de que un estudiante apruebe su última materia:
   aa) En el tercer intento

ab) Antes del cuarto intento.
ac) En el primer intento

b) ¿Cuántos intentos se espera realizar hasta que apruebe su última materia?, ¿con que desviación estándar?

4.32 En la producción de cámaras digitales se sabe que, en promedio 1 de cada 10 tiene algún tipo de problemas en su óptica.

a) ¿Cuál es la probabilidad de que la quinta cámara digital controlada sea la primera con problemas en su óptica?
b) ¿Cuál es la probabilidad de que la octava cámara digital controlada sea la segunda con problemas en su óptica?

4.33 Una empresa que vende únicamente por internet, envía un email a sus clientes para comunicarle una nueva forma de pago. La probabilidad de que un cliente elegido al azar conteste este email es de 0,1. Encuentre:

a) La distribución de probabilidad de la variable aleatoria X: número de emails que debe enviar hasta obtener una respuesta.
b) La esperanza y desviación estándar de la variable aleatoria X, e interprete en términos del problema.
c) La distribución de probabilidad de la variable Y: número de emails que recibe respuesta de 15 enviados.
d) La esperanza y desviación estándar de la variable aleatoria Y, e interprete en términos del problema.
e) La distribución de probabilidad de la variable aleatoria W: número de emails que debe enviar hasta obtener exactamente k respuestas.
f) La probabilidad de que deba enviar 6 emails antes de obtener dos respuestas.
g) La probabilidad de que en el octavo emails enviado sea el tercero en ser contestado.

4.34 Una empresa vitivinícola realiza una exportación de vinos con 100 cajas de vino malbec, 80 cajas de vino torrontés y 120 cajas de vino cabernet sauvignon. Se extrae una muestra aleatoria de 20 cajas para su inspección. ¿Encuentre la probabilidad de que este conformada por 8 cajas de malbec, 5 cajas de torrontés y de 7 cajas de cabernet sauvignon?

4.35 Una prueba de preselección tiene tres posibles resultados: positivos, negativos e inciertos. Se conoce que en una población el 12 % de los sujetos son positivos, el 68 % negativos y el resto son inciertos. ¿Qué probabilidad hay que en una muestra de 9 individuos tomada de dicha población, obtener exactamente 4 individuos positivos, 3 negativos y 2 inciertos?

4.36 Considere una evaluación que contiene 15 preguntas del tipo verdadero/falso. La evaluación se aprueba contestando correctamente por lo menos 9 preguntas. Si se contesta al azar, lanzando una moneda para decidir el valor de verdad de cada pregunta.

a) Determine el modelo de distribución de la variable aleatoria, justifique su elección. Obtenga el valor de los parámetros que la caracterizan.

b) ¿Cuál es la probabilidad de aprobar la evaluación?
c) ¿Cuántas preguntas se espera contestar correctamente utilizando este método?

4.37 Una compañía de radio taxis y remises de la provincia de Córdoba, recibe en promedio un pedido vía telefónica de tres por minuto.

a) Enuncie la variable aleatoria. Indique su parámetro.
b) Determine la función de probabilidad de la variable aleatoria.
c) Hallar la probabilidad de que reciba un pedido de un automóvil en un minuto.
d) Hallar la probabilidad de que reciba más de tres pedidos en un minuto.
e) Hallar la probabilidad de que reciba cuatro pedidos en dos minutos.

4.38 Pruebe que si $X_1 \, y \, X_2$ son dos variables aleatorias con distribución de Poisson $P(\lambda_1) \, y \, P(\lambda_2)$, la variable aleatoria suma de estas dos variables tienen también distribución de Poisson $P(\lambda_1 + \lambda_2)$. (Ayuda: Utilice función generadora de momentos)

4.39 Una corporación de exploración petrolera desea comprobar la existencia de hidrocarburos en una determinada región de la cuenca neuquina, por lo que se debe recurrir a la perforación de pozos exploratorios. Se planea realizar 8 perforaciones independientes. Cada perforación tiene una probabilidad de 0,125 de obtener petróleo en forma productiva. El presupuesto de los costos de la compañía considerando todos los costos de equipos e instalaciones de apoyo, es de U$S 60.000 por cada pozo. Un pozo productivo extrae petróleo por valor 10 veces más de lo que les costó perforar dicho pozo.

a) Defina la variable aleatoria de interés. Indique el modelo que refleja la situación.
b) Encuentre la distribución de probabilidad de la variable en estudio.
c) Encuentre la probabilidad de que exactamente cuatro pozos sean productivos.
d) Encuentre la probabilidad de que al menos dos pozos sean productivos.
e) Encuentre la probabilidad de que la mayoría de los pozos sean productivos.
f) Encuentre la probabilidad de que el número de pozos exploratorios productivos se encuentre entre $(\mu \pm \sigma)$
g) ¿Cuántos pozos espera la compañía que sean productivos? ¿Con qué desviación estándar?
h) La ganancia que espera obtener la compañía por los 8 pozos.
i) La desviación estándar de las ganancias de la corporación.