Unidad II

Definición de Variable Aleatoria. Variable Discreta. Distribución de probabilidad. Función de distribución acumulada. Valor esperado, Varianza y Desvío estándar de una Variable discreta. Distribución Binomial, Geométrica, Binomial negativa, Hipergeométrica y Poisson.

Pistonesi Silvina

Modelos de Variables

Aleatorias Discretas

Modelos de Variables Aleatorias Discretas

- *Distribución Binomial*.
- *Distribución Geométrica*.
- *Distribución Binomial negativa*.
- *Distribución Hipergeométrica.*
- *Distribución de Poisson*.

Def. Modelo Binomial

Un experimento se denomina Binomial si satisface las siguientes condiciones:

- **El experimento consta de *n* ensayos idénticas, siendo *n* fijo.**

- **Cada ensayo tiene sólo dos resultados posibles, que se denominarán Éxito (E) y Fracaso (F). La probabilidad de tener un solo éxito en un sólo ensayo es igual a p, y la de tener un solo fracaso es $q = 1 - p$. Un ensayo de este tipo se denomina *ensayo de Bernoulli*.**

- **Los ensayos son independientes, es decir, que el resultado de un ensayo no influye sobre el de los otros.**

- **La probabilidad p y q permanecen constantes de ensayo a ensayo.**

Ejemplos

1. **Se arroja un dado equilibrado 2 veces y se llama Éxito al suceso "se obtiene un cinco".**

2. **Se extraen 4 bolillas *con reposición* de una urna que contiene 5 bolillas blancas y 3 rojas y se denomina Éxito al suceso "la bolilla extraída es blanca".**

3. **Un fabricante de computadoras afirma que el 10% de sus computadoras necesitan un servicio posventa, antes del año de compra. Una empresa compra 3 computadoras. Se denomina Éxito a la computadoras que requieren service antes del año de compra.**

2. **Se extraen 4 bolillas *con reposición* de una urna que contiene 5 bolillas blancas y 3 rojas y se denomina Éxito al suceso “la bolilla extraída es blanca”.**

N = 8 bolillas

5 B

3 **R**

n =4
Con repos.

B = «la bolilla es blanca»
R=«la bolilla es roja»

- **El experimento consta de**

- **Cada ensayo tiene sólo dos resultados posibles:**

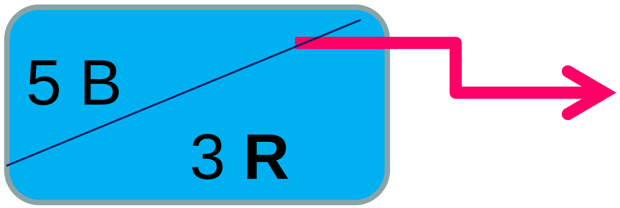
Éxito= **y Fracaso** =

$$p = P(\quad) =$$

$$q = 1 - p = P(\quad) =$$

2. Se extraen 4 bolillas ***con reposición*** de una urna que contiene 5 bolillas blancas y 3 rojas y se denomina **Éxito** al suceso "la bolilla extraída es blanca".

N = 8 bolillas



n =4
Con repos.

B = «la bolilla es blanca»
R=«la bolilla es roja»

p **= P(   ) =   ,**

q **= 1 –** ***p*** **= P(   )=**

- **¿Los ensayos son independientes?**

P((B,B,B,R)) =

3. **Un fabricante de computadoras afirma que el 10% de sus computadoras necesitan un servicio posventa, antes del año de compra. Una empresa compra 3 computadoras. Se denomina Éxito a la computadoras que requieren service antes del año de compra.**

N = ?

10% **S**

90% **S**^{c}

S = «la computadora necesita service postventa»

n =3

Sin repos.

- **El experimento consta**

- **Cada ensayo tiene sólo dos resultados posibles:**

  Éxito =

  Fracaso =

p **= P(   ) =**

***q* = 1 – *p* = P(   ) =**

3. **Un fabricante de computadoras afirma que el 10% de sus computadoras necesitan un servicio posventa, antes del año de compra. Una empresa compra 3 computadoras. Se denomina Éxito a la computadoras que requieren service antes del año de compra.**

N = **?**

10% **S**

90% $\mathbf{S^c}$

S = «la computadora necesita service postventa»

n =3

Sin repos.

p **= P( ) =**

q* = 1 – *p **= P( ) =**

- **¿Los ensayos son independientes?**

P((S,S,S^c)) =

Variable aleatoria binomial

Dado un experimento binomial que consta de *n* ensayos y en el cual P(E) = *p*, se denomina v.a. binomial a la variable

X = "número de éxitos en las *n* ensayos".

Distribución de probabilidad binomial

La distribución de probabilidad para una v.a. binomial está dada por:

$$p_X(x) = P(X = x) = \binom{n}{x} p^x \, q^{n-x} \qquad \text{con } x = 0, 1, 2, \ldots, n$$

Donde

p* =** probabilidad de éxito en un solo ensayo **y *q* = 1- *p

***n* =** número de ensayos

x = número de éxitos en *n* ensayos

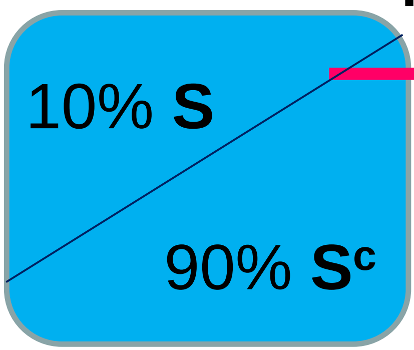
$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Ejemplo 3 **Un fabricante de computadoras afirma que el 10% de sus computadoras necesitan un servicio posventa, antes del año de compra. Una empresa compra 3 computadoras. Se denomina Éxito a la computadoras que requieren service antes del año de compra.**

N = **?**

S = «la computadora necesita service postventa»

p = P(S) = **0.1**



n =3

Sin repos.

X =

R_X ={ },   X ~ B (n = , p =)

$$\mathbf{P(X} = \mathbf{1)} = P(\qquad\qquad\qquad)$$

Pistonesi Silvina

Notación

X ~ B(*n*, *p*) indica que la v.a. X tiene distribución binomial con parámetros *n* y *p*.

Observar

Para cada x, P(X = x) constituye un término del desarrollo del *binomio de Newton* (p + q)^n, por este motivo X recibe el nombre de v.a. binomial.

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

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \dots + \binom{n}{n-2}a^2b^{n-2} + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n$$

Valor Esperado y Varianza de una v.a. binomial

X = "*número de éxitos en las n ensayos*".

X ~ B(*n, p*) $\qquad$ **R_X = { 0,1,…,n },**

$$\boxed{E(X) = n\ p}$$

E(X) *representa el número esperado de éxitos logrados en los en n ensayos*.

$$\boxed{V(X) = n\ p\ q}$$

V(X) *es la variabilidad del número de éxitos respecto del número promedio de éxitos logrados en n ensayos.*

Ejemplo 1

Un buscador de Internet intenta encontrar cierta palabra clave en una secuencia de páginas web independientes. Se cree que el 15% de los sitios web contienen esta palabra clave. Si visita 10 páginas web elegidas al azar,

a. ¿Cuál es la probabilidad de que en todos los sitios web visitados contengan esta palabra clave?

b. ¿Cuál es la probabilidad de que por lo menos 8 de los 10 sitios web visitados no contengan esta palabra clave?

c. ¿Cuál es el número promedio de sitios web que contienen esta palabra clave entre los visitados?

a. ¿Cuál es la probabilidad de que en todos los sitios web visitados contengan esta palabra clave?

X = "*N° de páginas web que contienen esta palabra clave entre las 10 visitadas***"**

$\mathbf{R_X = \{ \qquad \},}$ $\quad \mathbf{X \sim B\,(}\boldsymbol{n} \mathbf{= \quad ,}\ \boldsymbol{p} \mathbf{= \qquad)}$

$$\mathbf{p_X(\quad) = P(X = \qquad)}$$

b. ¿Cuál es la probabilidad de que por lo menos 8 de los 10 sitios web visitados no contengan esta palabra clave?

Y =

$R_Y = \{ \qquad \},\qquad Y \sim B\,(n = \quad ,\ p = \qquad)$

$P(Y \qquad) =$

e) ¿Cuál es el número promedio de sitios web que contienen esta palabra clave entre los visitados?

X =

$R_X = \{ \qquad \}$, $\qquad X \sim B(n = \qquad, p = \qquad)$

Ejemplo 2

En unas pruebas de alcoholemia se ha observado que el **5%** de los conductores controlados dan positivo en la prueba y que el **10%** de los conductores controlados no llevan puesto el cinturón de seguridad. También se ha observado que las dos infracciones son independientes. Una noche cualquiera un agente de tránsito para cinco conductores al azar,

a) determinar la probabilidad de que exactamente 3 conductores no lleven puesto el cinturón de seguridad.

b) determinar la probabilidad de que a lo sumo a 2 conductores controlados le de positivo en la prueba de alcoholemia.

c) determinar la probabilidad de que al menos uno de los conductores controlados haya cometido alguna de las dos infracciones

d) ¿Cuál es el número promedio de conductores que no llevan cinturón de seguridad puesto en 5 controlados?

Distribución Geométrica

Def.

- **El experimento consiste en una sucesión de ensayos idénticos, hasta obtener el primer éxito.**
- **Cada ensayo tiene dos resultados posibles que llamamos éxito (E) y fracaso (F), con P (E) = p, P (F)= q =1-p.**
- **En cada ensayo la probabilidad de éxito p y de fracaso q no varían.**
- **Los ensayos son independientes.**

Ejemplos*:*

1. *Número de lanzamientos de una moneda hasta obtener la primer cara.*

2. *Número de artículos inspeccionados hasta obtener el primer artículo defectuoso.*

3. *Número de niveles jugados hasta perder una vida en un juego de playstation.*

Distribución Geométrica

Experimento: *Se cuentan los lanzamientos de una moneda hasta obtener la primer cara.*

- **El experimento consiste en una sucesión de ensayos idénticos, hasta obtener el primer éxito?**

  Si, el Éxito=

- **Cada ensayo tiene dos resultados posibles:**

  E = y el fracaso **F** =

  P (E) = ***p*** = y **P (F)** = ***q*** =***1- p*** =

- **En cada ensayo la probabilidad de éxito p y de fracaso q, no varían. ¿es cierto?**

P(tener que hacer 3 lanzamientos para obtener la primer cara)

Variable aleatoria Geométrica

Dado un experimento Geométrico, en el cual P(E) = *p*, se denomina v.a. Geométrica a la variable:

X = "*número de ensayos hasta obtener el primer éxito*".

Distribución de probabilidad Geométrica

La distribución de probabilidad para una v.a. Geométrica está dada por:

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

Donde

p* =** probabilidad de éxito en un solo ensayo **y *q* = 1- *p

x = número de ensayos hasta lograr el primer éxito.

Podemos ver que **P(X=x)** es un sumando del desarrollo de la ***serie geométrica*** que da motivo al nombre de v.a **Geométrica**.

Notación

X ~ G(*p*) indica que la v.a. X tiene distribución Geométrica con parámetro *p*.

Valor Esperado y Varianza de una v.a. Geométrica

X = "*número de ensayos hasta obtener el primer éxito***".**

X ~ G(*p*) $\qquad$ R_X **={1, 2, …}**

$$E(X) = 1/\,p \quad y \quad V(X) = q\,/\,p^2$$

E(X) *representa el número esperado ó promedio de ensayos a realizar hasta obtener el primer éxito***.**

V(X) *es la variabilidad del número de ensayos a realizar hasta obtener el primer éxito***.**

Ejemplo 1

Un director de recursos humanos entrevista candidatos, uno por uno, para llenar una vacante, ie, hasta encontrar el candidato que reciba la oferta del puesto. Se sabe por datos anteriores que el **20%** de los entrevistados recibe una oferta.

a. ¿Cuál es la probabilidad de que el director deba entrevistar a 7 candidatos para obtener el primero que reciba la oferta?

b. ¿Cuál es la probabilidad de que de que se deba entrevistar a por lo menos 3 candidatos para obtener el primero que reciba la oferta?

c. ¿Cuál es el número promedio de entrevistas que debe efectuar el director para encontrar el candidato que reciba la oferta?

Ejemplo 1

Un director de recursos humanos entrevista candidatos, uno por uno, para llenar una vacante. Se sabe por datos anteriores que el 20% de los entrevistados recibe una oferta.

a. ¿Cuál es la probabilidad de que el director deba entrevistar a 7 candidatos para obtener el primero que reciba la oferta?

Sea X =

$$\mathbf{X \sim G\,(\mathit{p} = \qquad) \qquad R_X = \{ \qquad \},}$$

$$\mathbf{P(X= \qquad) =}$$

b. ¿Cuál es la probabilidad de que de que se deba entrevistar a por lo menos 3 candidatos para obtener el primero que reciba la oferta?

$$\mathbf{P(X \qquad) =}$$

c. **¿Cuál es el número promedio de entrevistas que debe efectuar el director para encontrar el candidato que reciba la oferta?**

X = “.

$R_X = \{ \qquad \}$, $\qquad$ $X \sim G\,(p = \qquad)$

Teniendo en cuenta el Ejemplo 2 (en Dist. Binomial)

En unas pruebas de alcoholemia se ha observado que el **5%** de los conductores controlados dan positivo en la prueba y que el **10%** de los conductores controlados no llevan puesto el cinturón de seguridad. También se ha observado que las dos infracciones son independientes. Una noche cualquiera un agente de tránsito para cinco conductores al azar,

¿Cual es la probabilidad de que el agente de tránsito deba controlar a 10 personas para encontrar el primero que no lleva el cinturón de seguridad puesto?

X = "*número de conductores controlados hasta encontrar el primero que no lleva puesto el cinturón de seguridad***",**

X ~ G (*p* = *0.1*) $\qquad$ $\mathbf{R_X = \{1, 2, \ldots\}}$**,**

$$P(X=10) = 0.1 \cdot 0.9^{10-1} = 0.039$$

Distribución Binomial Negativa (Pascal)

Def.

- **El experimento consiste en una sucesión de ensayos idénticos, hasta obtener *k* éxitos.**

- **Cada ensayo con dos resultados posibles que llamamos éxito (E) y fracaso (F), con P(E) = *p*, P(F) = *q* =*1-p*.**

- **Al repetir el experimento la probabilidad de éxito *p* y de fracaso *q* no varían.**

- **Los intentos son independientes.**

Ejemplos:

1. *Número lanzamientos de una moneda, hasta obtener cinco* caras.

2. *Número de artículos inspeccionados hasta obtener 3 defectuosos.*

Variable aleatoria Binomial Negativa

Dado un experimento Binomial Negativa, en el cual P(E) = p, y k es el número de éxitos a obtener, se denomina v.a. Binomial Negativa a la variable:

X = "*número de ensayos hasta obtener el k-ésimo éxito*".

Si el número de ensayos necesarios es x, una realización sería:

$$\underbrace{\underbrace{F\ \ F\ \ F\ \ F\ldots F}_{x-\,k \text{ fracasos}}\ \ \underbrace{E\ \ E\ \ E\ldots.\ \ E}_{k-\,1 \text{ éxitos}}\ \ \underbrace{E}_{\text{último éxito}}}_{x \text{ repetciones}}$$

Con probabilidad $\longrightarrow$ $p^k(1-p)^{x-k}$

Ensayos son independientes

Multiplicamos esta probabilidad por todos los ordenamientos posibles de (x- k) fracasos y (k -1) éxitos:

$$p_X(x) = P(X = x) = \binom{x-1}{k-1} p^k\, q^{x-k} \qquad \text{con } x = k, k+1, \ldots$$

Notación

X ~ BN(*k*, *p*) indica que la v.a. X tiene distribución Binomial Negativa con parámetros k y *p*.

Valor Esperado y Varianza de una v.a. Binomial Negativa

X = "*número de ensayos hasta obtener el k-ésimo éxito*".

R_X **={ k,k+1, …},**

$$E(X) = k/p \quad y \quad V(X) = qk/p^2$$

E(X) *representa el número esperado ó promedio de ensayos a realizar hasta obtener el* ***k-ésimo éxito*****.**

V(X) *es la variabilidad del número de ensayos a realizar hasta obtener el k-ésimo éxito*.

Ejemplo 1

Después que un virus informático entró en el sistema, un sistema gestor de datos comprueba la condición de todos los archivos importantes. Se sabe que cada archivo tiene una probabilidad de 0.2 de ser dañado por el virus, independientemente de otros archivos.

a. ¿Cuál es la probabilidad de que se deban analizar 5 archivos para encontrar el cuarto archivo dañado por el virus?

b. ¿Cuál es la probabilidad de que de que se deban analizar entre 5 y 7 archivos para obtener el tercer archivo dañado por el virus?

Ejemplo 1

Después que un virus informático entró en el sistema, un sistema gestor de datos comprueba la condición de todos los archivos importantes. Se sabe que cada archivo tiene una probabilidad de 0.2 de ser dañado por el virus, independientemente de otros archivos.

a. ¿Cuál es la probabilidad de que se deban analizar 5 archivos para encontrar el cuarto archivo dañado por el virus?

X = "*N° de archivos analizados hasta obtener el cuarto dañado por el virus*",

$$X \sim BN(k = \quad , p = \quad) \qquad R_X = \{ \qquad \},$$

$$- \quad - \quad - \quad - \quad D \quad \left\{ \begin{array}{lllll} D & D & D & D^c & D \\ D & D & D^c & D & D \\ D & D^c & D & D & D \\ D^c & D & D & D & D \end{array} \right.$$

Cada **4-upla** tiene la misma probabilidad

$P((D,D,D,D^c,D)) =$

Ejemplo 1

Después que un virus informático entró en el sistema, un sistema gestor de datos comprueba la condición de todos los archivos importantes. Se sabe que cada archivo tiene una probabilidad de 0.2 de ser dañado por el virus, independientemente de otros archivos.

b. ¿Cuál es la probabilidad de que se deban analizar entre 5 y 7 archivos para obtener el tercer archivo dañado por el virus?

X =

X ~ BN (k = , *p* =) $R_X = \{ \quad \}$

Teniendo en cuenta el Ejemplo 2 (en Dist. Binomial) Determinar:

En unas pruebas de alcoholemia se ha observado que el **5%** de los conductores controlados dan positivo en la prueba y que el **10%** de los conductores controlados no llevan puesto el cinturón de seguridad. También se ha observado que las dos infracciones son independientes.

a. ¿Cual es la probabilidad de que el agente de tránsito deba controlar a lo sumo a 4 personas para encontrar el segundo conductor que le da positivo la prueba de alcoholemia?

b.¿Cual es el número promedio de conductores deba controlar el agente de tránsito para encontrar el décimo que le da positivo la prueba de alcoholemia?

Teniendo en cuenta el Ejemplo 2 (en Dist. Binomial) Determinar:

a. ¿Cual es la probabilidad de que el agente de tránsito deba controlar a lo sumo a 4 personas para encontrar el segundo conductor que le da positivo la prueba de alcoholemia?

X = "N° *de conductores a controlar hasta encontrar el segundo que le da positivo la prueba de alcoholemia***",**

X ~ BN (k= 2, *p* = *0.05*) R_X ={2,3, …},

$$P(X \leq 4) = P(X=2) + P(X=3) + P(X=4) =$$

$$= \binom{2-1}{2-1} 0.05^2\, 0.95^{2-2} + \binom{3-1}{2-1} 0.05^2\, 0.95^{3-2} + \binom{4-1}{2-1} 0.05^2\, 0.95^{4-2} =$$

$$= 0.0025 + 0.00475 + 0.0068 = 0.014$$

b. ¿Cual es el número promedio de conductores deba controlar el agente de tránsito para encontrar el décimo que le da positivo la prueba de alcoholemia?

X **= “***N° de conductores a controlar hasta encontrar el décimo que le da positivo la prueba de alcoholemia***”,**

X ~ BN (k= 10, *p* = *0.05*) $\qquad$ **R_X ={10,11, …},**

$$E(X) = \frac{k}{p} = \frac{10}{0.05} = 200$$

En promedio, el agente de tránsito debe para 200 conductores para encontrar el décimo que le da positivo la prueba de alcoholemia.

IMPORTANTE !!!!

Modelo	Definición de la v.a.	Lo aleatorio es el
Binomial	***N° de éxitos en n ensayos***	**Número de éxitos**
Binomial Negativa	***N° de ensayos necesarios para obtener el k-ésimo éxito***	**Número de ensayos**

Obs.: La distribución **Geométrica** es un **caso particular** de la distribución **Binomial Negativa.**

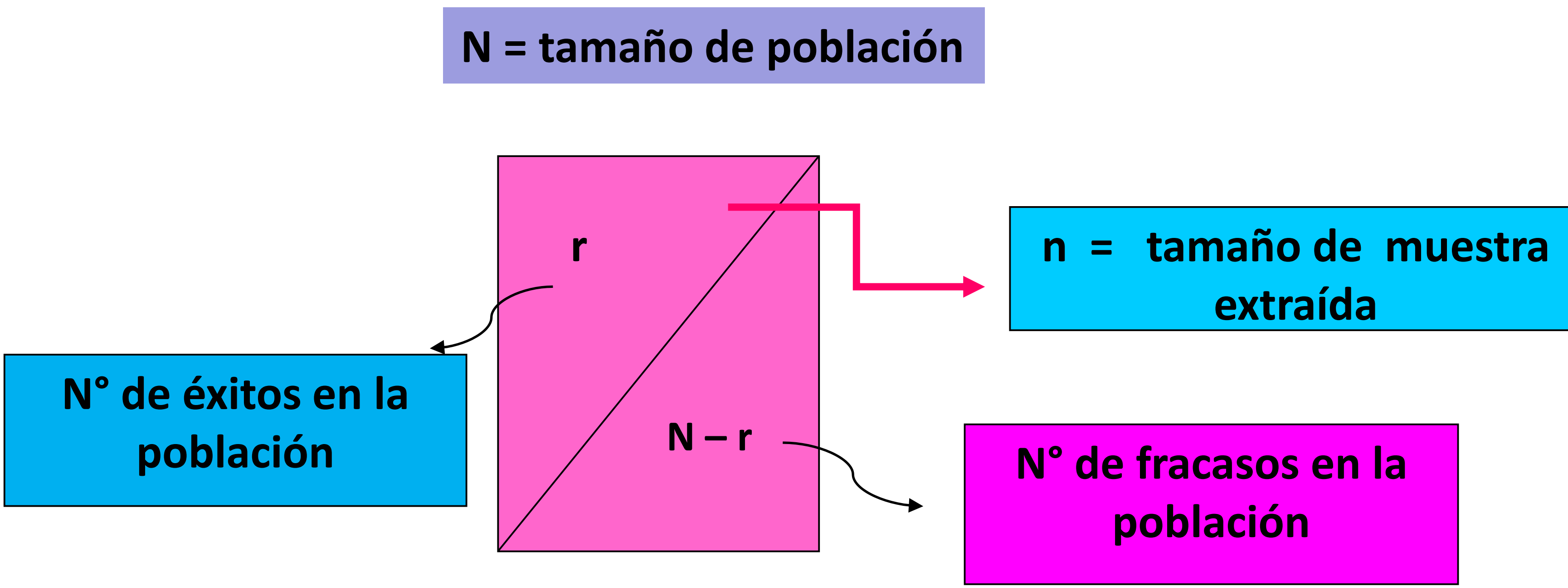
Modelo Hipergeométrico

Def.

Un experimento se denomina Hipergeométrico si satisface las siguientes condiciones:

El experimento consiste en extraer al azar y sin sustitución *n* elementos de un población finita de *N* elementos, cada uno de los cuales puede ser clasificado como éxito (E) y fracaso (F).
Los *N* elementos de la población se dividen en *r* éxitos y *N - r* fracasos (F).

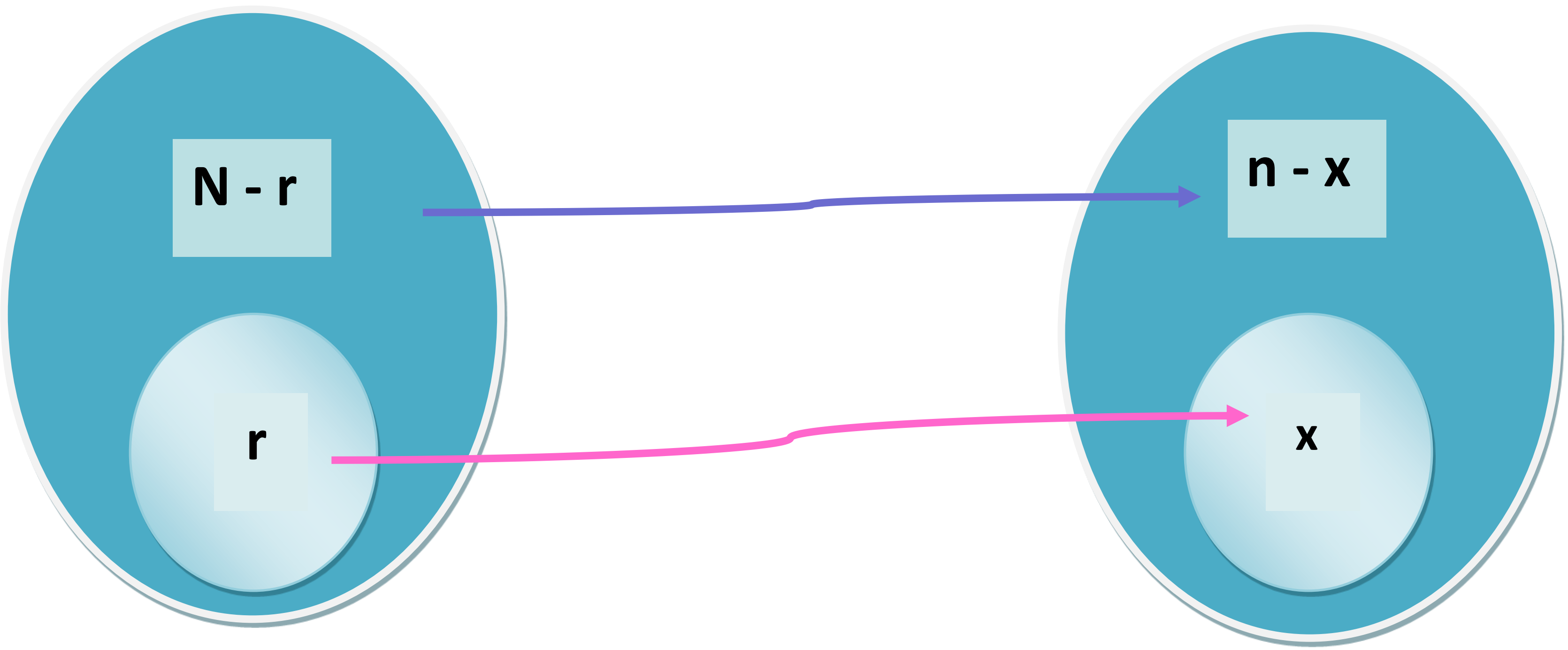
Modelo Hipergeométrico



Modelo Hipergeométrico

N = tamaño de población

n = tamaño de muestra



Ejemplos

1. De una urna que contiene 3 bolillas blancas y 7 negras se extraen 4 bolillas *sin reposición*.

2. A una convención política nacional asisten 30 representantes de cierto estado, de los cuales 20 apoyan al candidato A y 10 al candidato B. Se eligen al azar 4 representantes para formar una comisión evaluadora de un proyecto sobre contaminación ambiental.

3. De un grupo de 20 empresas, de las cuales 15 son nacionales y las restantes extranjeras, la Corporación del Comercio y la Industria elige al azar 3 empresas para adquirir ciertos suministros.

Variable Aleatoria Hipergeométrica:

Dado un experimento hipergeométrico, se denomina v.a. hipergeométrica a la variable:

X = "número de éxitos en la muestra de tamaño *n*".

Distribución de probabilidad hipergeométrica:

La distribución de probabilidad para una v.a. hipergeométrica está dada por:

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

donde ***N*** **=** tamaño de población
n **=** tamaño de muestra
x = número de éxitos en la muestra

Notación:

X ~ H(*N*, *n*, *r*) indica que la v.a. X tiene distribución hipergeométrica con parámetros *N*, *n* y *r*.

Valor Esperado y la Varianza de una v.a. Hipergeométrica

X = "*número de éxitos en la muestra de tamaño n*".

X ~ H(*N*, *n*, *r*)

$$E(X) = n\frac{r}{N} \quad y \quad V(X) = n\frac{r}{N}\,\frac{N-r}{N}\left(\frac{N-n}{N-1}\right)$$

donde $\left(\frac{N-n}{N-1}\right)$ **es el factor de corrección por población finita.**

E(X) *representa el número esperado de éxitos logrados en los en n ensayos*.

V(X) *es la variabilidad del número de éxitos respecto del número promedio de éxitos logrados en n ensayos.*

Ejemplo 1

Un equipo de trabajo establecido por la Agencia de Protección Ambiental de los Estados Unidos programó visitas a empresas industriales para investigar la posibilidad de violaciones a los reglamentos para el control de la contaminación. De un total de 18 empresas industriales sólo se investigarán 3 de ellas elegidas al azar. Se sabe que 6 de las empresas están operando realmente sin cumplir con los reglamentos.

Considérese el "*N° de empresas industriales que operan bajo violaciones a los reglamentos para el control de la contaminación en la muestra de 3*".

N = 18 empresas

12 **C**

6 $\mathbf{C}^c$

n =3
Sin repos.

Se trata de un Experimento

a. ¿Se trata de un experimento hipergeométrico?
¿Cuál es la v.a. X de interés? ¿cuál es su recorrido? ¿y cuáles son sus parámetros?

X

$R_X = \{ \qquad \},$ $\qquad X \sim H\,(N = \quad , n = \quad , r = \quad)$

b. ¿Cuál es la probabilidad de que a lo sumo una empresa industrial seleccionada opere violando los reglamentos?

$P(X \qquad) =$

c. ¿Cuál es la probabilidad de que por los menos dos empresas seleccionadas operen violando los reglamentos?

d. ¿Cuál es el número promedio de empresas entre las seleccionadas que operen sin violar los reglamentos?

Y

$Y \sim H\,(N = \quad, \mathit{n} = \quad, \mathit{r} = \quad)$ $\qquad R_Y = \{ \qquad \},$

Ejercicio:

La mayoría de las personas conoce el póker con manos de 5 cartas. Con 52 cartas, ¿cuál es la probabilidad de que la mano de cinco cartas contenga:

1.Un par de ases.

2. Exactamente un as.

3. Ningún as.

4. Cuando menos 1 as.

5. ¿Cuál es el número promedio de ases entre las 5 cartas repartidas?

Observación!!!

La **distribución de probabilidad Hipergeométrica** está estrechamente relacionada con la **distribución de probabilidad Binomial**.

La **diferencia** entre ambas radica en la *independencia de los ensayos,* que necesita la **distribución Binomial** y en que la probabilidad de *éxito cambia de un ensayo a otro* en la **distribución Hipergeométrica** (ensayos no son independientes).

Modelo de Poisson

Es el modelo que se caracteriza por evaluar el número de sucesos que ocurren en un intervalo especificado de tiempo o espacio.

Ejemplos

1. **El número de mails que llegan a un servidor de correo durante una hora.**

2. **N° de clientes que ingresan a un banco por día.**

3. **N° de fallas en un producto manufacturado.**

4. **N° de solicitudes de seguro procesadas por una compañía en un período de un año.**

5. **N° de bacterias en 5 cm^3 de agua.**

Modelo de Poisson

Las propiedades de un experimento de Poisson son:

- El experimento consiste en contar el número de veces que ocurre un determinado suceso durante un intervalo de tiempo o espacio específico (área, volumen, peso, distancia, etc.).

- La probabilidad de una ocurrencia del suceso es igual en dos intervalos disjuntos de igual longitud.

- La ocurrencia del suceso en cualquier intervalo es independiente de la ocurrencia del suceso en cualquier otro intervalo disjunto. Proceso de Poisson no tiene *memoria*.

- La probabilidad de que se presente un suceso, es proporcional a la longitud del intervalo de tiempo o espacio considerado.

Variable aleatoria Poisson:

Dado un experimento Poisson, se denomina v.a. Poisson a la variable:

X **= "***número de sucesos que ocurren en un intervalo específico de tiempo o espacio***".**

Distribución de probabilidad Poisson:

La distribución de probabilidad para una v.a. Poisson está dada por:

$$\mathbf{p_X(x) = P(X = x)} = \frac{e^{-\lambda} \quad \lambda^{x}}{x!} \qquad \textbf{con} \;\; \mathbf{x} = 0,1,2,\ldots$$

Donde

λ= número promedio de eventos que ocurren en el intervalo de tiempo o espacio.

x = número de eventos que ocurren en el intervalo de tiempo o espacio.

Notación

$X \sim P(\lambda)$ indica que la v.a. X tiene distribución de Poisson de parámetro λ.

Valor Esperado y Varianza de una v.a. Poisson

X = "*número de sucesos que ocurren en un intervalo específico de tiempo o espacio*".

$X \sim P(\lambda)$ $\quad$ $R_X = \{0, 1, 2, \ldots\}$,

$$E(X) = \lambda = V(X)$$

Ambos valores coinciden.

Ejemplo 1

Los clientes de un proveedor de servicios de Internet abren nuevas cuentas a una tasa promedio de 10 cuentas por día Calcular:

a. **La probabilidad que durante el próximo día se abran tres cuentas nuevas.**

b. **La probabilidad de que en un período de 6 días se abran al menos dos nuevas cuentas.**

c. **El número promedio de cuentas nuevas abiertas en un período de un mes.**

d. **La probabilidad que el número de cuentas nuevas abiertas el próximo día se aleje del número medio en a lo sumo 1 desvío estándar.**

a. La probabilidad que durante el próximo día abran tres cuentas nuevas.

X = "*N° de cuentas nuevas abiertas por día*".

X ~ *P*(*λ* = *10 cuentas*)

$$p_X(3) = P(X = 3) = \frac{e^{-10}\ 10^3}{3!} = 0.0076$$

b. La probabilidad de que en un período de 6 días se abran al menos dos nuevas cuentas.

X = "*N° de cuentas nuevas abiertas por un cliente en un período de 6 días*".

$$si \quad 1\,\text{día} \rightarrow 10\,\text{cuentas}$$

$$6\,\text{días} \rightarrow \lambda = \frac{6.10}{1} = 60 \quad \text{cuentas}$$

X ~ *P*(*λ* = *60 cuentas*)

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1)$$

$$= 1 - \frac{e^{-60}\ 60^0}{0!} - \frac{e^{-60}\ 60^1}{1!} \cong 1$$

c. El número promedio de cuentas nuevas abiertas en un período de un mes.

X **= "***N° de cuentas nuevas abiertas en un mes***".**

$$si \quad 1\,\text{día} \rightarrow 10 \;\; \text{cuentas}$$

$$30\,\text{días} \rightarrow \lambda = \frac{30\,.\,10}{1} = 300 \quad \text{cuentas}$$

X ~ *P*(*λ* = *300 cuentas*)

$$E(X) = 300$$

*Por mes, en promedio se abren 300 cuentas nuevas***.**

d. **La probabilidad que el número de cuentas nuevas abiertas el próximo día se aleje del número medio en a lo sumo 1/2 desvío estándar.**

X **= “***N° de cuentas nuevas abiertas por día***”.**

$$X \sim P(\lambda = 10 \text{ cuentas})$$

$$desvío \quad estándar \quad = \quad \sigma_X = \sqrt{V(X)} = \sqrt{10} = 3.16$$

$$P(|X - 10| \leq 1/2 \,.\, \sigma_X) = P(|X - 10| \leq 3.16/2)$$

$$= P(-1.58 \leq X - 10 \leq 1.58) = P(8.42 \leq X \leq 11.58) =$$

$$= \frac{e^{-10} \quad 10^{9}}{9!} + \frac{e^{-10} \quad 10^{10}}{10!} + \frac{e^{-10} \quad 10^{11}}{11!} =$$

$$= 0.1251 + 0.1251 + 0.1137 = 0.3639$$

e. **Si se consideran los próximos 6 días, ¿cuál es la probabilidad que en uno de ellos se abran tres cuentas nuevas?**

Sea **Y** = "*Nº de días en los que se abran tres cuentas nuevas entre los 6 días considerados al azar.*"

Y ~ B(**n = 6, *p* = ?)**

$$\mathbf{P(Y = 1)} = \binom{6}{1} \; p^{1}\, q^{6-1} = \mathbf{?}$$

p = P (de que en un día se abran tres cuentas nuevas)

X = "*número de cuentas nuevas abiertas por día*".

X ~ *P*(λ = 10 cuentas)

$$p = \mathbf{P(X = 3)} = \frac{e^{-10}\;\; 10^{3}}{3!} = \mathbf{0.0076}$$

Luego,

$$\mathbf{P(Y=1)} = \binom{6}{1}\ \mathbf{0.0076^{1}\,(1-0.0076)^{6-1} = 0.0439}$$

El 4.39% de las veces que se eligen al azar 6 días, en uno de ellos **se abren tres cuentas nuevas.**

Ejemplo 2

Una prisión de máxima seguridad reporta que el número promedio de intentos de escape por mes es de 1,5. Calcular:

a. La probabilidad que durante el próximo mes se reporten tres intentos de escape.

b. La probabilidad de que el próximo semestre se reporten al menos un intento de escape.

c. El número promedio de intentos de escape el próximo año.

d. La probabilidad que el número de intentos de escape reportados el próximo mes se aleje del número medio en a lo sumo 1 desvío estándar.

a. La probabilidad que durante el próximo mes se reporten tres intentos de escape

X = "número de intentos de escape por mes".

X ~ *P*(λ *= 1,5 intentos*)

$$p_X(3) = P(X = 3) = \frac{e^{-1,5} \quad 1,5^3}{3!} = 0.126$$

b. La probabilidad de que el próximo semestre se reporten al menos un intento de escape.

X = "número de intentos de escape por semestre".

$$si \quad 1\,\text{mes} \rightarrow 1,5 \ \text{intentos}$$

$$6\,\text{meses} \rightarrow \lambda = \frac{6\,.\,1,5}{1} = 9 \quad \text{intentos}$$

X ~ *P*(λ *= 9 intentos*)

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) =$$

$$= 1 - \frac{e^{-9} \quad 9^0}{0!} = 0.999$$

c. El número promedio de intentos de escape el próximo año.

X = "número de intentos de escape por año".

$$si \quad 1\,\text{mes} \quad \rightarrow \quad 1{,}5 \ \text{intentos}$$

$$12\,\text{meses} \rightarrow \lambda = \frac{12 \,.\, 1{,}5}{1} = 18 \quad \text{intentos}$$

X ~ *P*(λ *= 18 intentos*)

$$E(X) = 18$$

Por año, en promedio se pronostican 18 intentos de escape en dicha prisión de máxima seguridad.

d. **La probabilidad que el número de intentos de escape reportados el próximo mes se aleje del número medio en a lo sumo 1 desvío estándar.**

X = "número de intentos de escape por mes".

X ~ *P*(λ *= 1,5 intentos*)

$$desvío \quad estándar \quad = \quad \sqrt{V(X)} = \sqrt{1,5} = 1,22$$

$$\mathbf{P}(\ |X - 1,5| \leq 1.\sigma_X) = \mathbf{P}(\ |X - 1,5| \leq 1,22) =$$

$$= \mathbf{P}(-1,22 \leq X - 1,5 \leq 1,22) = P(0,28 \leq X \leq 2,72) =$$

$$= \frac{e^{-1,5} \quad 1,5^1}{1!} + \frac{e^{-1,5} \quad 1,5^2}{2!} =$$

$$= 0.335 + 0.251 =$$

e. Si se consideran los próximos 6 meses, ¿cuál es la probabilidad que en dos de ellos se reporten tres intentos de escape?

Sea **Y** = "Nº de meses en los que se reportan tres intentos de escape entre los 6 considerados al azar."

Y ~ B(n = 6, *p* = ?)

$$\mathbf{P(Y = 2)} = \binom{6}{2} \; \boldsymbol{p^{2}\, q^{6-2}} = \mathbf{?}$$

p = P (de que en un mes se reportan tres intentos de escape)

X = "número de intentos de escape por mes".

X ~ *P(λ = 1,5 intentos)*

$$\mathbf{p} = \mathbf{P(X = 3)} = \frac{e^{-1,5}\; 1,5^{3}}{3!} = 0.126$$

Luego,

$$\mathbf{P(Y = 2)} = \binom{\mathbf{6}}{\mathbf{2}} \; \mathbf{0}.126^{\mathbf{2}} \; \mathbf{(1- 0.126)}^{\mathbf{6\text{-}2}} = \mathbf{0.1390}$$

El 13.9% de las veces que se eligen al azar 6 meses, en dos de ellos **se reporten tres intentos de escape.**

Ejercicio:

Un estudio del número de ventas de automóviles en una consecionaria del centro de Bahía reveló que, en promedio por día, se venden 3 automóviles.

1. ¿Cuál es la probabilidad de que que no venda ninguno en un día elegido al azar?

2. ¿Cuál es la probabilidad de que durante 5 días consecutivos se venda al menos un auto?

3. ¿Cuál es el número esperado de autos a vender el próximo año?

5. ¿Cuál es el desvío estándar del número de autos vendidos el próximo bimestre?

Aproximación de Binomial a Poisson

Si hacemos tender el número de repeticiones ***n*** a infinito y la probabilidad ***p*** de éxito a cero de tal forma que **λ = *n.p*** se mantenga constante, podemos aproximar la **P(*X* = *x*)** en la **distribución Binomial** por la **distribución de Poisson***:*

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

$$=\frac{\lambda^{x}}{x!}1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$=\frac{\lambda^{x}}{x!}1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$\xrightarrow[n\to\infty]{}\frac{\lambda^{x}}{x!}e^{-\lambda}$$

$$\boxed{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}\xrightarrow[n\to\infty]{}1$$

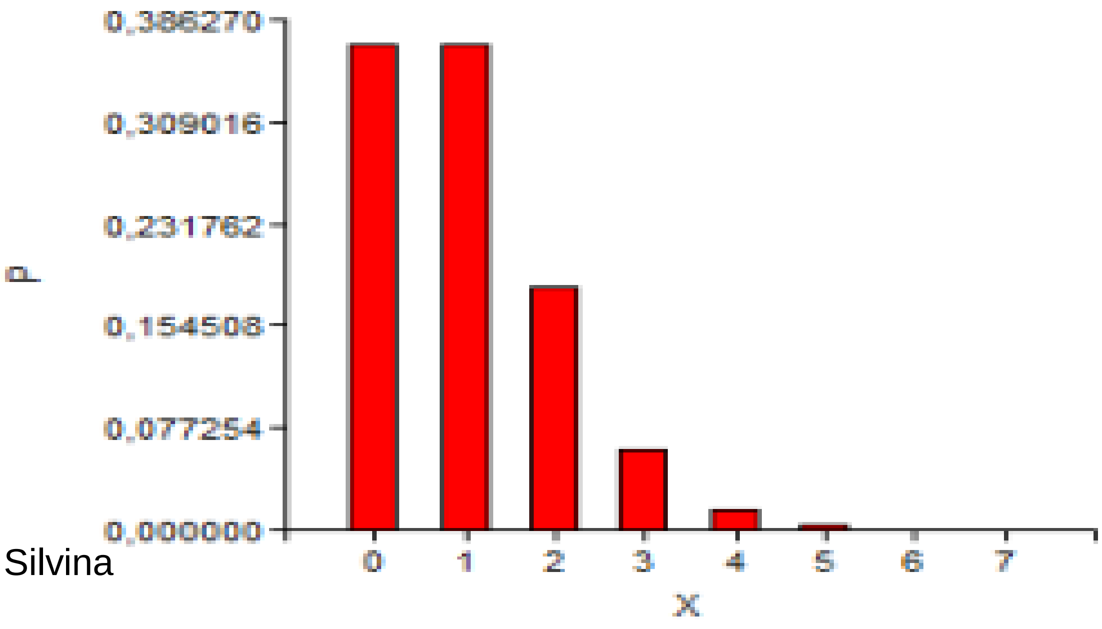
$$\boxed{\left(1-\frac{\lambda}{n}\right)^{n}}\xrightarrow[n\to\infty]{}e^{-\lambda}$$

Aproximación de Binomial a Poisson

Ejemplo Una muestra de sangre debidamente homogeneizada es vertida sobre una hemacitómetro para proceder al recuento de glóbulos blancos. Si la homogeneización fue correcta los glóbulos estarán distribuídos al azar en el hemacitómetro.

X : número de glóbulos blancos por cuadrícula
Se sabe que el número esperado de glóbulos por cuadrícula es 1.

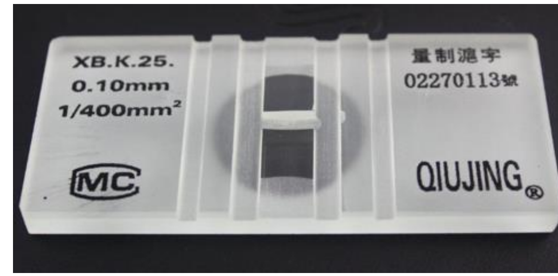
Gráficamente la distribución de probalidades es

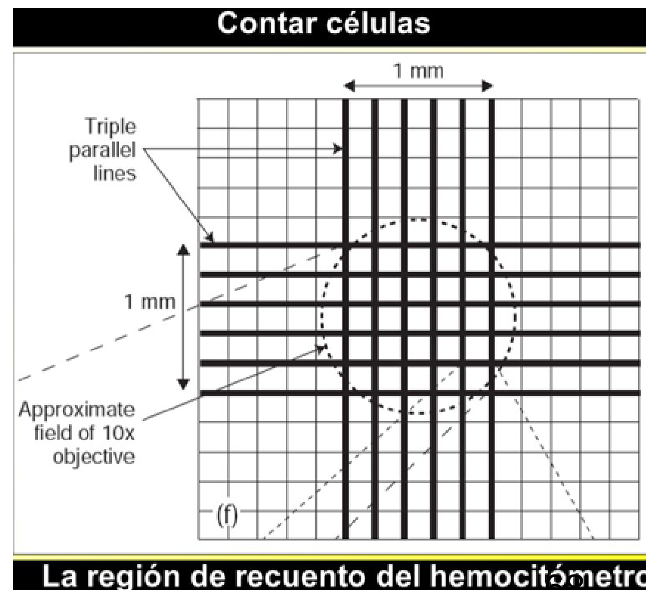


Hemacitómetro:

es un aparato utilizado para estimar la concentración de células de una muestra de sangre.

Consiste en una placa de cristal con una depresión central dividida en campos por una fina cuadrícula, cada uno de ellos de unos 0,05 mm cuadrados. El cristal se coloca bajo el microscopio y se mide la cantidad de células por cada cuadrícula, lo que permite calcular la concentración de células hemáticas (glóbulos blancos) en la muestra estudiada.





Aproximación de Binomial a Poisson

Si $X \sim B(n, p)$ con $n \geq 30$ y $np < 5$, entonces X se puede aproximar por una distribución Poisson con parámetros $\lambda = np$.

Gráficamente

- $B(50, 0.02)$ ($n = 50$ y $p = 0.02$) (azul)
- $P(1)$ (rojo)