

# Distribuciones de probabilidad discretas

Jhon Jairo Padilla A., PhD.

# Introducción

- A menudo, las observaciones de diferentes experimentos aleatorios tienen el mismo tipo general de comportamiento.
- Las v.a. discretas asociadas con estos experimentos se pueden describir esencialmente con la misma distribución de probabilidad (la misma fórmula básica).
- En otras palabras, existen diferentes distribuciones de probabilidad típicas que pueden usarse para modelar el comportamiento de variables aleatorias en diferentes experimentos aleatorios.

# Distribuciones de probabilidad discretas típicas y sus aplicaciones

Distribución	Variable Aleatoria
Binomial	Número de pacientes curados con un fármaco
Hipergeométrica	Número de artículos defectuosos en un lote de producción
Geométrica	Número de muestras requeridas para generar una falsa alarma en un control estadístico
Poisson	Número de leucocitos de una cantidad fija de una muestra de sangre
Geométrica	Búsqueda secuencial en una tabla
Poisson	Número de llegadas de llamadas a un conmutador telefónico

# Distribución discreta uniforme

- Es la más simple de todas las distribuciones.
- La v.a. toma cada uno de sus valores con una probabilidad idéntica.

- Definición:

- Si la v.a.  $X$  toma los valores  $x_1, x_2, \dots, x_k$  con idénticas probabilidades, entonces la distribución uniforme discreta está dada por

$$f(x;k) = 1/k, \quad x = x_1, x_2, \dots, x_k$$

Se utiliza la notación  $f(x;k)$  para indicar que la distribución uniforme depende del parámetro  $k$ .

# Ejemplo

- Se selecciona una bombilla al azar de una caja que contiene 4 bombillas, una de 40watts, una de 60w, una de 75w y una de 100w. Todas las bombillas tienen igual probabilidad de ser escogidas ( $1/4$ ). Por tanto, se tiene una distribución uniforme con

$$f(x;4)=1/4; \quad x=40, 60, 75, 100$$

# Media y Varianza de la distribución uniforme discreta

- La media y la varianza de la distribución uniforme discreta  $f(x;k)$  son:

$$\mu = \frac{1}{k} \sum_{i=1}^k x_i$$

$$\sigma^2 = \frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2$$

# Ejemplo

- Para el ejemplo de las bombillas se tiene que:

$$\mu = \frac{1}{k} \sum_{i=1}^k x_i = \frac{1}{4} (40 + 60 + 75 + 100) = 68,75$$

$$\sigma^2 = \frac{1}{k} \sum_{i=1}^k (x_i - \mu)^2 = \frac{1}{4} [(40 - 68,75)^2 + (60 - 68,75)^2 + (75 - 68,75)^2 + (100 - 68,75)^2] = 639,58$$

# **DISTRIBUCION BINOMIAL**



# Tipo de experimento

- Se realizan pruebas repetidas
- Cada prueba tiene dos resultados posibles (éxito o fracaso, 1 o cero, correcto o incorrecto, etc)
- Este proceso se denomina **Proceso de Bernoulli**
- Ejemplos:
  - Línea de ensamble donde se prueba cada uno de los productos y se determina si está defectuoso o no.
  - Lanzamiento de una moneda 10 veces. El resultado puede ser cara ó sello.
  - Se transmiten bits en serie en un canal de comunicaciones, cada bit puede llegar errado ó correcto

# Características de un proceso de Bernoulli

- El experimento consiste de  $n$  ensayos que se repiten
- Cada ensayo produce un resultado que se puede clasificar como éxito o fracaso (éxito y fracaso son sólo etiquetas, no tienen el significado de que algo esté bien o mal)
- La probabilidad de un éxito se denota como  $p$  y permanece constante de un ensayo a otro
- Los ensayos que se repiten son independientes

# Variable aleatoria binomial

- El número de éxitos de  $n$  experimentos de Bernoulli se denomina **variable aleatoria binomial**.
- La distribución de probabilidad de esta variable aleatoria discreta se llama **distribución binomial**.
- Distribución binomial:  $b(x; n, p)$ . Depende del número de ensayos y de la probabilidad de éxito en un ensayo dado.

# Distribución binomial

- Un experimento de Bernoulli puede tener como resultado un éxito con probabilidad  $p$  y un fracaso con probabilidad  $q=1-p$ . Entonces, la distribución de probabilidad de la v.a. binomial  $X$ , que es el número de éxitos en  $n$  ensayos independientes es,

$$b(x;n, p) = b(x) = \binom{n}{x} p^x q^{n-x} \quad , x=0,1,2,\dots, n$$

# Nombre “Binomial”

- La distribución binomial toma su nombre del hecho que la expresión matemática de la f.d.p. tiene una forma similar a la de la serie que describe el polinomio:

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

# Ejemplo

- La probabilidad de que cierta clase de componente sobreviva a una prueba de choque es  $\frac{3}{4}$ . Encuentre la probabilidad de que sobrevivan exactamente 2 de los siguientes 4 componentes que se prueben.

- Solución:

$$p = \frac{3}{4}$$

$$n = 4$$

$$x = 2$$

$$b(x; n, p) = b(x) = \binom{n}{x} p^x q^{n-x}$$

$$b(2) = \binom{4}{2} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^{4-2} = \frac{4!}{2!2!} \cdot \frac{3^2}{4^4} = \frac{27}{128}$$

# Ejemplo

- Las posibilidades de que un bit transmitido a través de un canal de transmisión digital se reciba con error es de 0.1. Suponga además que los ensayos de transmisión son independientes. Sea  $X$  el número de bits con error en los siguientes 4 bits transmitidos. Determine la probabilidad de que lleguen dos bits con error.
- Solución:

$$p=0.1$$

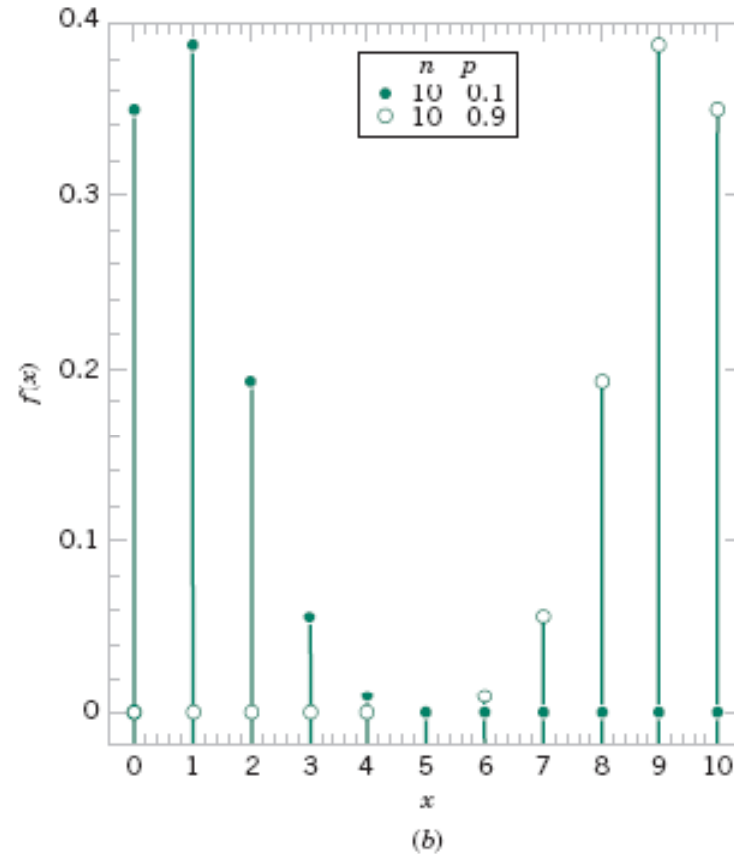
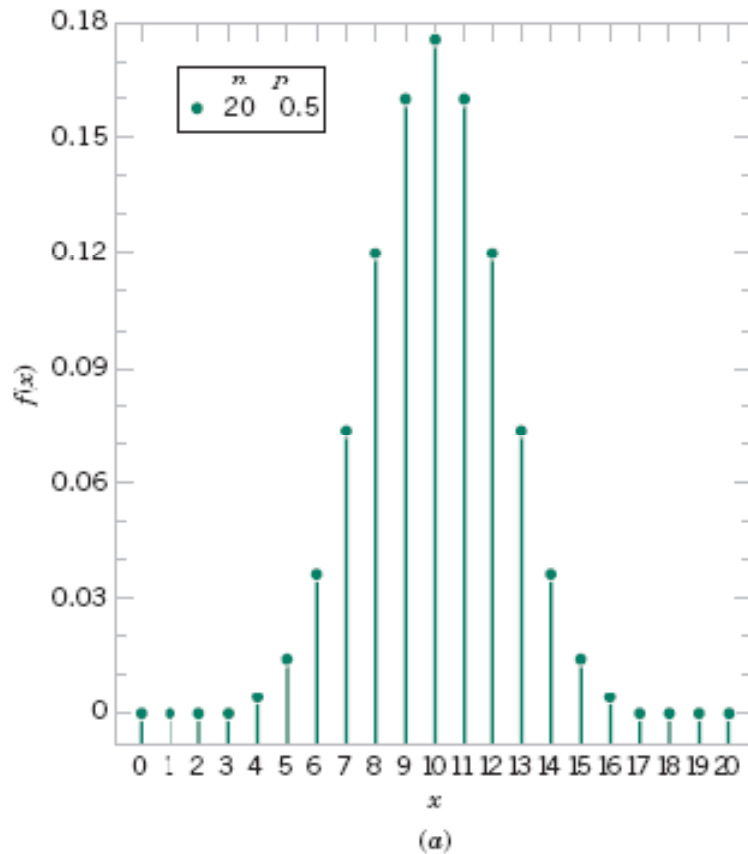
$$n=4$$

$$x=2$$

$$b(x; n, p) = b(x) = \binom{n}{x} p^x q^{n-x}$$

$$b(2) = \binom{4}{2} (0.1)^2 (0.9)^{4-2} = 0,0486$$

# Distribución binomial para diferentes valores de $n$ y $p$



**n** alto (Mayor densidad de puntos),  $p=0.5$  (Simetría)



# Aplicaciones

- Todo experimento que tenga resultados binarios (éxito/fracaso, defectuoso/no defectuoso, enfermo/sano, hombre/mujer, etc) y cuyos ensayos sean independientes.
- Ejemplos:
  - Medicina: Ensayos de fármaco que cura/no cura
  - Aplicaciones militares: Ensayos de envío de proyectiles que dan en el blanco/fallan
  - Comunicaciones: probabilidad de error de una cadena de bits
  - Calidad en Líneas de producción: Pruebas de productos Defectuosos/No defectuosos

# Media y Varianza

- La media y la varianza de la distribución binomial  $b(x;n,p)$  son

$$\mu = np$$

$$\sigma^2 = npq$$

- Ejemplo:
  - Para el ejemplo de la transmisión de 4 bits con  $p=0.1$ , se tiene

$$\mu = (4) \cdot (0,1) = 0,4$$

$$\sigma^2 = (4) \cdot (0.1) \cdot (0.9) = 0,36$$

# **DISTRIBUCION GEOMÉTRICA**

# Tipo de Experimento

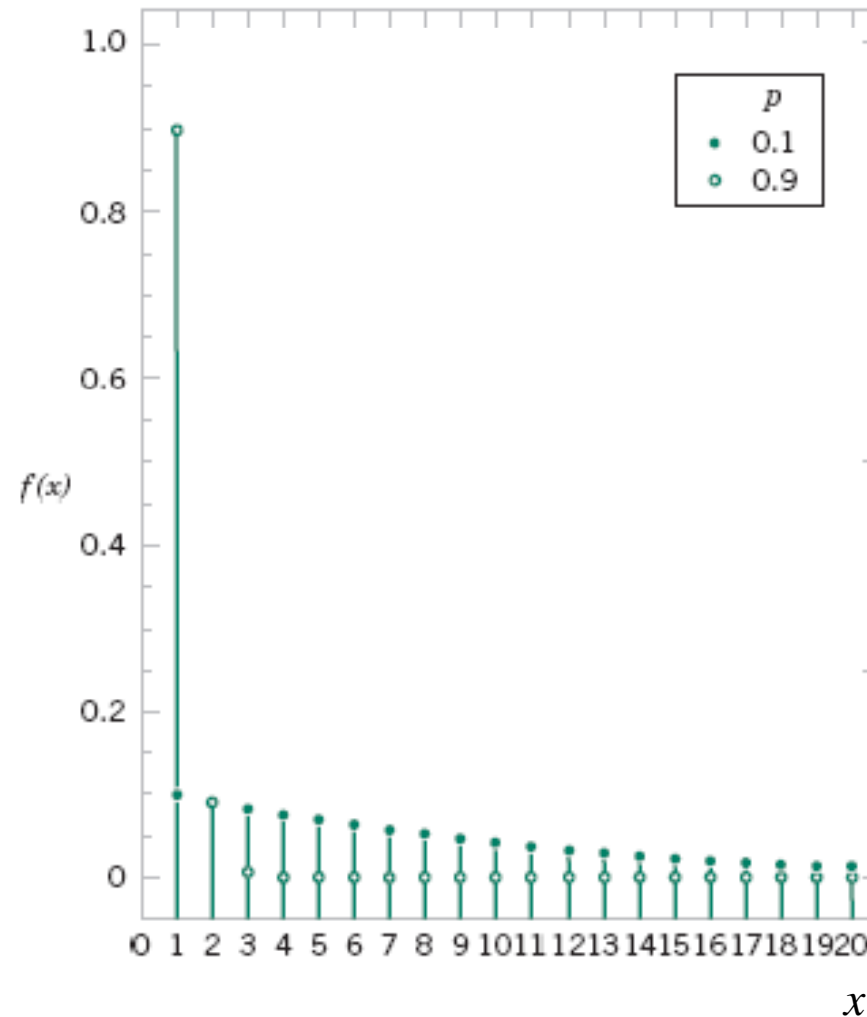
- Supóngase una serie de ensayos de Bernoulli independientes con probabilidad  $p$  de éxito en cada ensayo.
- Los ensayos se realizan hasta que se obtiene un éxito.
- La v.a.  $X$  denota el número de ensayos hasta que se obtiene el primer éxito.

# Definición

- En una serie de ensayos de Bernoulli independientes, con probabilidad constante  $p$  de un éxito, sea que la variable aleatoria  $X$  denote el número de ensayos hasta el primer éxito. Entonces  $X$  tiene una distribución geométrica con parámetro  $p$  y

$$f(x) = (1-p)^{x-1} p \quad , x=1,2,\dots$$

# Comportamiento de una distribución geométrica variando $p$



El decrecimiento de  $f(x)$  con cada valor de  $x$  tiene una progresión geométrica:  
 $f(x) = (1-p)f(x-1)$

# Ejemplo

- La probabilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error es de 0.1. Suponga que las transmisiones son eventos independientes y sea que la v.a.  $X$  denote el número de bits transmitidos hasta el primer error. Calcule la probabilidad de que se transmitan 4 bits correctamente y el quinto bit esté errado.

- *Solución:*

$$X=5$$

$$p=0.1$$

$$q=1-0,1=0,9$$

$$P(X = 5) = (0,9)^4 (0,1) = 0,066$$

# Ejemplo

- En cierto proceso de fabricación, en promedio, uno de cada 100 artículos está defectuoso. ¿Cuál es la probabilidad de que el quinto artículo que se inspecciona sea el primer artículo defectuoso que se encuentra?

- *Solución:*

$$p=1/100=0,01$$

$$x=5$$

$$P(X = 5) = (0,99)^4 (0,01) = 0,0096$$



# Media y varianza de una v.a. con distribución geométrica

- Si  $X$  es una v.a. geométrica con parámetro  $p$ , entonces la media y la varianza de  $X$  son

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

$$\sigma^2 = \frac{(1-p)}{p^2}$$

# Ejemplo

- En el ejemplo de la transmisión de bits con probabilidad 0.1 de que llegue un bit errado, calcule el número promedio de bits transmitidos incluyendo el primer bit errado. Calcule la desviación estándar de esta misma variable.

- *Solución:*  
 $p=0,1$   
$$\mu = E(X) = \frac{1}{0,1} = 10$$
$$\sigma = \left[ \frac{(1-p)}{p^2} \right]^{1/2} = \left[ \frac{(1-0,1)}{0,01^2} \right]^{1/2} = 9,49$$

# **DISTRIBUCION BINOMIAL NEGATIVA**

# Tipo de Experimento

- Supóngase una serie de ensayos de Bernoulli independientes con probabilidad  $p$  de éxito en cada ensayo.
- Los ensayos se realizan **hasta** obtener  $r$  éxitos
- Esta es una generalización de una distribución geométrica

# Definición

- En una serie de ensayos de Bernoulli independientes, con probabilidad constante  $p$  de éxito, sea que la v.a.  $X$  denote el número de ensayos hasta que ocurran  $r$  éxitos. Entonces  $X$  tiene una distribución binomial negativa

donde

$$f(x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r$$

- Para  $x=r, r+1, r+2, \dots$

# Aclaraciones

- Debido a que se necesitan al menos  $r$  ensayos para obtener  $r$  éxitos, el rango de  $X$  es de  $r$  a infinito.
- Si  $r=1$ , una v.a. binomial negativa se convierte en una v.a. geométrica

# Ejemplo

- Suponga que la probabilidad de que un bit transmitido a través de un canal de transmisión digital se reciba con error es 0,1. Suponga que las transmisiones son eventos independientes, y sea que la v.a.  $X$  denote el número de bits transmitidos hasta que se transmite el cuarto bit errado. Calcule la probabilidad de que el décimo bit sea el cuarto bit errado.

- Solución:

$r=4$  éxitos

$x=10$  ensayos

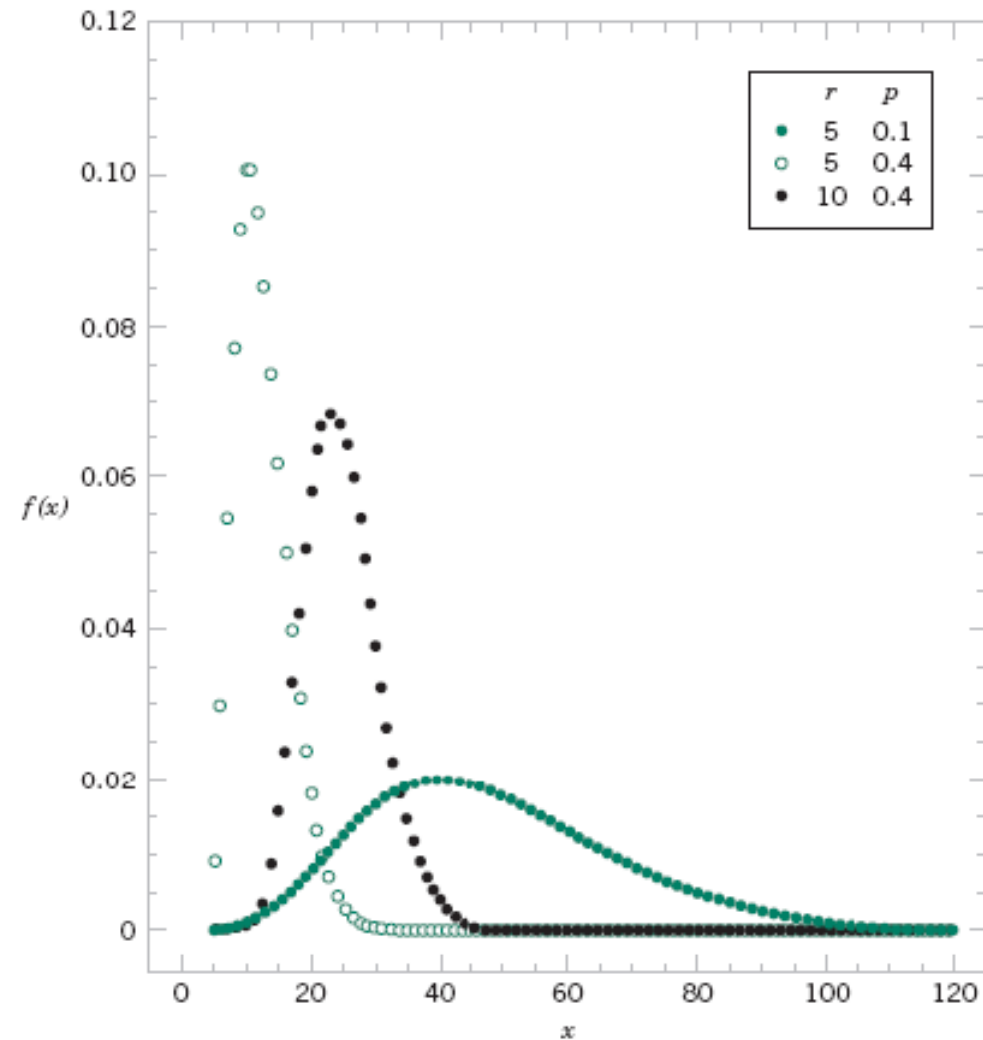
$$f(x) = \binom{10-1}{4-1} (1-0,1)^{10-4} 0,1^4$$

$$f(x) = \binom{9}{3} (0,9)^6 0,1^4 = \boxed{\binom{9}{3} (0,9)^6 0,1^3 0,1^1}$$

Probabilidad de que  
ocurran 3 errores en los  
nueve primeros ensayos  
(Distribución binomial)

Prob. Del  
cuarto bit  
errado

# Distribuciones binomiales negativas para diferentes valores de $r$ y $p$





# Media y Varianza

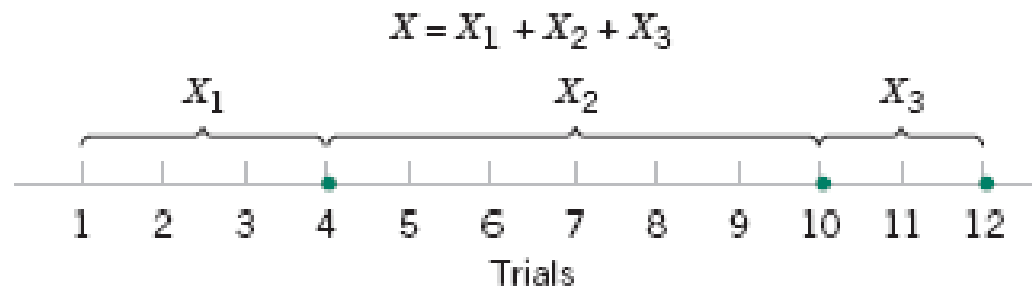
- Si  $X$  es una v.a. aleatoria binomial negativa con parámetros  $p$  y  $r$ , entonces la media y varianza de  $X$  son,

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

$$\sigma^2 = \frac{r(1-p)}{p^2}$$

# Relación entre las distribuciones geométrica y binomial negativa

- Se dice que una v.a. geométrica no tiene memoria.
- La **propiedad de falta de memoria** significa que cada vez que se obtiene un éxito se re-inicia el conteo de los resultados de los ensayos hasta el siguiente éxito.
- Sea  $X_1$  el número de ensayos hasta el primer éxito,  $X_2$  el número de ensayos hasta el segundo éxito, etc. Entonces, para obtener  $r$  éxitos  $X = X_1 + X_2 + \dots + X_r$



• indicates a trial that results in a "success".

v.a. binomial como suma de v.a. geométricas

# Aplicación de las distribuciones binomial negativa y geométrica

- Si la probabilidad de tener que hacer un gran número de intentos antes de obtener un éxito (distrib. geométrica) o  $r$  éxitos (distrib. Binomial negativa) es alta, esto podría significar que se emplearían muchos esfuerzos en lograr el objetivo (lo que podría significar altos costos en un proyecto).

# **DISTRIBUCIÓN HIPERGEOMÉTRICA**

# Tipo de experimento

- Se realizan experimentos de Bernoulli
- Interesa calcular la probabilidad para un número de observaciones que caen dentro de una categoría (éxito) en  $n$  ensayos.
- A diferencia de la distribución binomial, en que hay un supuesto de independencia (los elementos que se sacan se reemplazan o vuelven a aparecer en el espacio muestral), **en este caso no hay reemplazo.**
- **Aplicaciones:** Experimentos en que el artículo probado se destruye o desecha y por ello no se puede reemplazar en el espacio muestral.

# Ejemplo

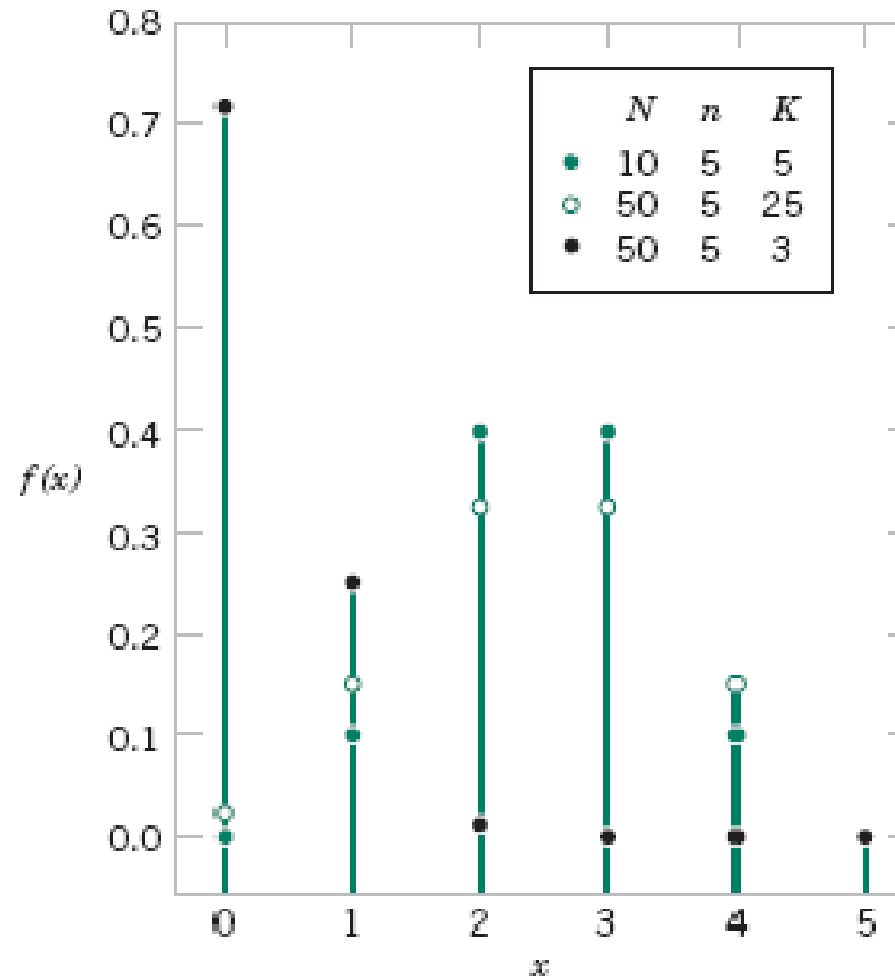
- Cuál es la probabilidad de tomar 3 cartas rojas en 5 extracciones de una baraja de 52 cartas, si hay 26 cartas rojas y 26 cartas negras?
- Solución:
  - No se puede aplicar la distribución binomial (a menos que cada carta que se saca se reemplace y que el paquete se revuelva para la siguiente extracción)
  - Nos interesamos en la posibilidad de sacar 3 cartas rojas de 26 posibles y 2 cartas negras de 26 posibles
  - Por tanto, hay  $\binom{26}{3}$  posibilidades de tomar 3 cartas rojas, y  $\binom{26}{2}$  posibilidades de tomar 2 cartas negras
  - El número total de formas de seleccionar 3 cartas rojas y 2 cartas negras es  $\binom{26}{3}\binom{26}{2}$
  - El número total de formas de tomar cualesquier 5 cartas de las 52 disponibles es  $\binom{52}{5}$
  - Por tanto, la probabilidad total será  $\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = 0,3251$

# Definición

- Un conjunto de  $N$  objetos contiene  $K$  objetos clasificados como éxitos y  $N-K$  objetos clasificados como fracasos
- Se selecciona una muestra con tamaño de  $n$  objetos al azar (sin reemplazo) de los  $N$  objetos, donde  $K \leq N$  y  $n \leq N$ .
- Sea que la v.a.  $X$  denote el número de éxitos en la muestra. Entonces  $X$  tiene una distribución hipergeométrica y

$$f(x) = \frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}} \quad \text{máx}\{0, n+K-N\} \leq x \leq \text{mín}\{K, n\}$$

# Distribución hipergeométrica para diferentes valores de $N$ , $n$ y $K$





# Media y varianza

- Si  $X$  es una v.a. con distribución hipergeométrica y con parámetros  $N$ ,  $K$  y  $n$ , la media y la varianza de  $X$  son

$$\mu = E(X) = np$$

$$\sigma^2 = np(1-p) \binom{N-n}{N-1}$$

- Donde  $p=K/N$

# Factor de corrección para poblaciones finitas

- La varianza de la distribución hipergeométrica es la varianza de la distribución binomial multiplicada por un factor.

$$\sigma^2 = np(1-p) \binom{N-n}{N-1}$$

Factor de corrección para poblaciones finitas. Se debe al hecho de no hacer reemplazos.

Varianza distribución binomial

# Relación con la distribución binomial

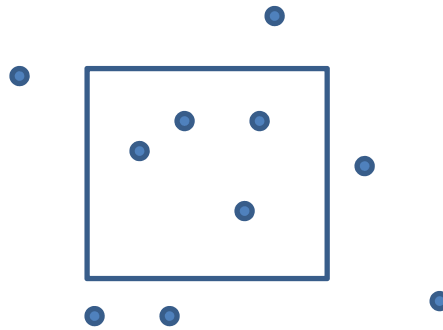
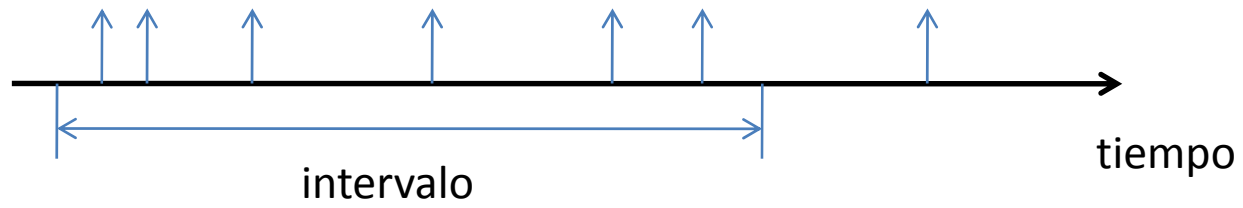
- Si  $n$  es pequeña comparada con  $N$ , la cantidad de artículos cambia en una proporción muy baja en cada prueba.
- La distribución binomial puede utilizarse como aproximación para la distribución geométrica cuando  $n$  es pequeña en comparación con  $N$
- Por regla general la aproximación es buena cuando  $(n/N) \leq 0.05$

# **DISTRIBUCIÓN DE POISSON**

# Tipo de Experimento

- Se desea describir el **número de resultados que ocurren** en:
  - Un intervalo de tiempo:
    - Minutos, días, semanas, meses, años
  - Una región específica:
    - Segmento de línea, un área, un volumen
- Aplicaciones:
  - Número de llamadas telefónicas por hora que recibe una oficina
  - Número de juegos suspendidos debido a la lluvia durante una temporada
  - Número de ratas de campo por acre
  - Número de bacterias en un cultivo dado
  - Número de errores mecanográficos por página

# Proceso de Poisson



Región

# Proceso de Poisson

- El **número de resultados** que ocurren en un intervalo o región es independiente del número que ocurre en cualquier otro intervalo o región disjunto (**No tiene memoria**).
- La **probabilidad** de que ocurra **un solo resultado** durante un intervalo muy corto o una región pequeña **es proporcional a la longitud del intervalo o al tamaño de la región**, y **no depende** del número de resultados que caen fuera de este intervalo o región.
- La **probabilidad de que ocurra más de un resultado** en tal intervalo corto o que caiga en tal región pequeña **es insignificante**.

# Distribución de Poisson

- El número  $X$  de resultados que ocurren durante un experimento de Poisson se llama **variable aleatoria de Poisson** y su distribución de probabilidad se llama **distribución de Poisson**



# Relación con la distribución binomial

- La probabilidad de una v.a con distribución binomial

tiene la forma: 
$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

- Si  $n$  se incrementa y  $p$  se hace muy pequeña, de manera que se mantenga siempre el producto  $pn = \lambda$  (**este producto es la media**), entonces

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

- Por tanto, cuando  $n$  es muy grande:

$$\lim_{n \rightarrow \infty} P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

# Definición

- La distribución de probabilidad de la v.a. de Poisson  $X$ , que representa el número de resultados que ocurren en un intervalo dado o región específicos ( $t$ ) y que se puede representar como  $\lambda t$ , es

$$f(x) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$$

- Donde  $\lambda$  es el número promedio de resultados por unidad de tiempo, distancia, área o volúmen.

# Media y Varianza

- Si  $X$  es una v.a. de Poisson con parámetro  $\lambda$ , entonces la media y la varianza de  $X$  son

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

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

- Consecuencia: Si la varianza de los conteos es mucho más grande que la media de los mismos, entonces la distribución de Poisson no es un buen modelo para la distribución de la v.a.

# Ejemplo

- Se presentan imperfecciones aleatoriamente a lo largo de un alambre delgado de cobre. Sea que  $X$  denote la v.a. que cuenta el número de imperfecciones en una longitud de  $L$  milímetros de alambre, y suponga que el número promedio de imperfecciones en 1 milímetros es  $\lambda=2.3$ . (a) Determine la probabilidad de exactamente 2 imperfecciones en 1mm de cable. (b) Determine la probabilidad de 10 imperfecciones en 5mm de alambre.

- **Solución:**

A)  $\lambda=2.3$ imperfecc/1mm  
t=1mm

$$f(2) = \frac{e^{-(2.3)(1)} ((2.3)(1))^2}{2!} = 0.265$$

B)  $\lambda=2.3$ imperfecc/1mm  
t=5mm

$$f(10) = \frac{e^{-(2.3)(5)} ((2.3)(5))^{10}}{10!} = 0.113$$

# Ejemplo

- El número promedio de camiones-tanque que llega cada día a cierta ciudad portuaria es 10. Las instalaciones en el puerto pueden manejar a lo más 15 camiones-tanque por día. ¿Cuál es la probabilidad de que en un día dado haya camiones que se tengan que regresar por no haber sido atendidos?
- Solución:
  - Se pregunta  $P(X > 15)$
  - Pero  $P(X > 15) = 1 - P(X \leq 15)$
  - Donde  $P(X \leq 15)$  es la función de distribución acumulada hasta  $X = 15$
  - Entonces:

$$1 - \sum_{x=0}^{15} f(x) = 1 - \sum_{x=0}^{15} \frac{e^{-(10)(1)} ((10)(1))^x}{x!} = 1 - 0.9513 = 0.0487$$

# Ejemplo

- En un proceso de fabricación donde se manufacturan productos de vidrio ocurren defectos o burbujas, lo cual ocasionalmente deja a la pieza indeseable para su venta. Se sabe que, en promedio, 1 de cada 1000 de estos artículos que se producen tiene una ó más burbujas. ¿Cuál es la probabilidad de que una muestra aleatoria de 8000 tenga menos de 7 artículos con burbujas?
- Solución:
  - Este es un experimento binomial con  $x=7$ ,  $n=8000$  y  $p=0.001$
  - Como  $p$  es cercana a cero y  $n$  es bastante grande, se puede usar la distribución de Poisson como aproximación, con  $\lambda=pn=8$ .
  - Entonces:

$$P(X < 7) = \sum_{x=0}^6 f(x) = \sum_{x=0}^6 \frac{e^{-(8)(1)} ((8)(1))^x}{x!} = 0.3134$$