TEMA IV.

"MODELOS PROBABILISTICOS DE FENOMENOS ALEATORIOS DISCRETOS".

DISTRIBUCION DE BERNOULLI

SI EN UN EXPERIMENTO SOLO APARECEN DOS POSIBLES RESULTADOS: "ÉXITO" O "FRACASO" , A DICHO EXPERIMENTO SE LE LLAMA DE BERNOULLI.

EJEMPLO:

SE LANZA UNA MONEDA, LOS POSIBLES RESULTADOS SON:

"CAE AGUILA" = ÉXITO

"CAE SOL" = FRACASO.

ASOCIANDO AL RESULTADO DEL EXPERIMENTO, UNA VARIABLE ALEATORIA (V.A.) x, ESTA VARIABLE TOMARA LOS VALORES DE 1 Y 0, ESTO ES, EL ESPECTRO DE x ES:

$$\{1, 0\}$$

DE ESTA MANERA SE PUEDE DEFINIR UNA FUNCION DE PROBABILIDAD:

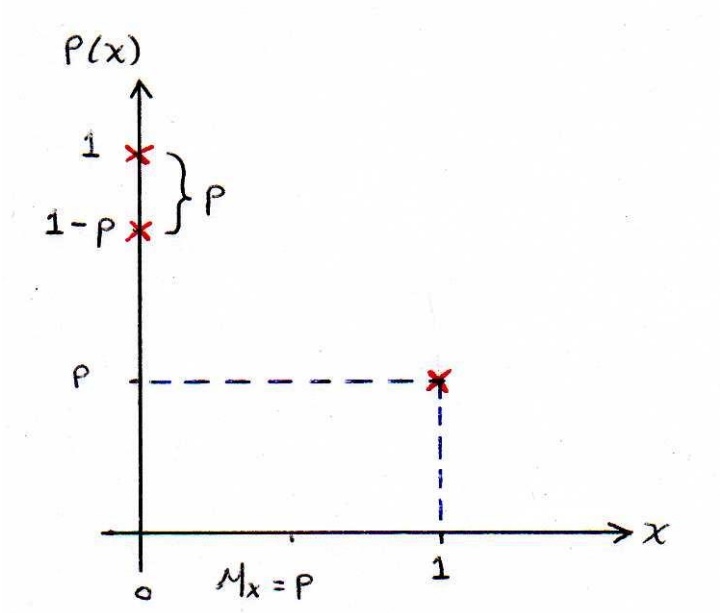
$$P(x) = \begin{cases} P & ; \quad x = 1 \quad \text{(ÉXITO)} \\ 1 - P & ; \quad x = 0 \quad \text{(FRACASO)} \end{cases}$$

ESTO ES $P(x = 1) = P$, $P(x = 0) = 1 - P = q$

DE TAL MANERA QUE:

$$\sum_{\forall x} P(x) = P(1) + P(0) = P + 1 - P = 1$$

GRAFICAMENTE:



LA MEDIA DE LA DISTRIBUCION.

$$\mu_x = E\{x\} = \sum_{\forall x} x\,P(x) = (0)\,P(0) + 1\,P(1)$$

$$= (0)(1-P) + (1)(P)$$

$$= p$$

LA VARIANCIA DE LA DISTRIBUCION:

$$\sigma_x^2 = E\{(x-\mu_x)^2\} = \sum_{\forall x} (x-\mu_x)^2\,P(x) = (0-P)^2\,P(0) + (1-P)^2\,P(1)$$

$$= P^2(1-P) + (1-P)^2\,P$$

$$= P(1-P)[P + (1-P)]$$

$$= P(1-P)$$

$$= pq$$

RECORDAR QUE:

$$\sigma_x^2 = E\{x^2\} - \mu_x^2$$

DISTRIBUCION BINOMIAL

SI EL EXPERIMENTO DE BERNOULLI SE LLEVA A CABO VARIAS VECES Y LAS PRUEBAS SON INDEPENDIENTES SE TIENE UN "PROCESO DE BERNOULLI".

SI LA VARIABLE ALEATORIA x REPRESENTA EL NUMERO TOTAL DE EXITOS Y EL EXPERIMENTO SE REPITE N VECES, ENTONCES SE BUSCARA CONOCER P(x)

EJEMPLO:

SE LANZA UNA MONEDA 3 VECES.

ÉXITO: CAE AGUILA ; P (x = 1) = P

FRACASO: CAE SOL ; P (x = 0) = 1 – P = q

P (x = 0) =?

$$P(x=0) = q \cdot q \cdot q = q^3 = {}_3C_0\, p^0\, q^3$$

P (x = 1) =?

$$P(x=1) = p \cdot q \cdot q + q \cdot p \cdot q + q \cdot q \cdot p$$

$$= 3\, p\, q^2$$

$$= {}_3C_1\, p\, q^2$$

P (x = 2) = ? ;

$$P(x=2) = p.\ p.\ q + q.\ p.\ p + p.\ q.\ p.$$

$$= 3\, p^2\, q$$

$$= {}_3C_2\, p^2\, q$$

P (x =3) = ?

$$P(x=3) = p.\ p.\ p$$

$$= p^3$$

$$= {}_3C_3\, p^3\, q^0$$

RELACION $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$

$$nCr = \frac{n!}{r!\,(n - r)\,¡}$$

N PRUEBAS EN UN PROCESO DE BERNOULLI

$$P(x = n) = ?$$

$$P(x = n) = {}_NC_n\, p^n\, q^{N-n}$$

EN LA DISTRIBUCION BINOMIAL PARA LA V. A x QUE REPRESENTA EL NUMERO DE EXITOS EN UN PROCESO DE BERNOULLI, SE DEFINE LA FUNCION DE PROBABILIDAD:

$$P(x) = {}_NC_n\, p^x\, q^{n-x}$$

$$P(X = n) = {}_NC_n\, p^n\, q^{N-n}$$

$$\mu = N\,p$$

$$\sigma^2 = N\,p\,q$$

EJEMPLO.

LA PROBABILIDAD DE QUE UN JUGADOR DE BASKET-BALL ANOTE UN TIRO LIBRE ES DE ¾. SUS TIROS SON INDEPENDIENTES, SI EN UN JUEGO PUEDE HACER 5 TIROS LIBRES, DETERMINAR:

a) LA PROBABILIDAD DE QUE ACIERTE EN TODOS SUS TIROS
b) LA PROBABILIDAD DE QUE FALLE TODOS SUS TIROS
c) LA PROBABILIDAD DE QUE ACIERTE POR LO MENOS LA MITAD DE SUS TIROS.

p = ¾ = 0.75 ; q = ¼ = 0.25 ; N = 5

Espectro = $\{0,1,2,3,4,5\}$

a) $P(x = 5) = {}_5C_5\, p^5\, q^0 = (5¡/(5¡\,(0)¡))\,(0.75)^5(0.25)^0$

$= 0.2373 = 23.73\%$

b) $P(x = 0) = {}_5C_0\, p^0\, q^5 = (5¡/(0¡\,(5)¡))\,(0.75)^0(0.25)^5$

$= 0.00097 = 0.097\%$

c) P (x > 3) = P (x=3) + P (x=4) + P (x=5) =

= 1 – (P(x=2) + P(x=1) + P(x=0))

DISTRIBUCION DE BERNOULLI.

$$P(x) = \begin{cases} P & ; x = 1 \\ 1 - P & ; x = 0 \end{cases} \qquad \mu_x = p \quad , \quad \sigma_x^2 = p(1-p) = pq$$

DISTRIBUCION BINOMIAL.

$$P(x) = {}_N C_x \, p^x \, q^{n-x} \qquad ; \; x = 0, 1, 2, \ldots, N \qquad \mu_x = Np$$

$$\sigma_x^2 = Npq$$

DISTRIBUCION GEOMETRICA

SI EN UN PROCESO DE BERNOULLI SE QUIERE CONOCER LA DISTRIBUCION DEL NUMERO DE PRUEBAS EFECTUADAS HASTA OBTENER EL PRIMER ÉXITO O SEA EL NUMERO DE FRACASOS CONSECUTIVOS QUE PRECEDEN AL PRIMER ÉXITO, A DICHA DISTRIBUCION SE LE LLAMA GEOMETRICA.

SEA x LA V. A. Y x = n EL NUMERO DE FRACASOS CONSECUTIVOS QUE PRECEDEN A UN ÉXITO, ENTONCES SE QUIERE CONOCER P (x =n)

EN UNA PRUEBA DE BERNOULLI
P (x = 1) = P
P (x = 0) = 1 – P = q

P (x = n) = q• q• q• ……. •q •p

n

P (x = n) = $q^n p$

P (x = n) = $q^n p$ FUNCION DE PROBABILIDAD

EJEMPLO:

UN JUGADOR DE BALONCESTO HACE 5 TIROS. OBTENER LA DISTRIBUCION GEOMETRICA.

p = ¾ ; q = ¼

SOLUCION:

Espectro = (0 , 1 , 2 , 3 , 4)

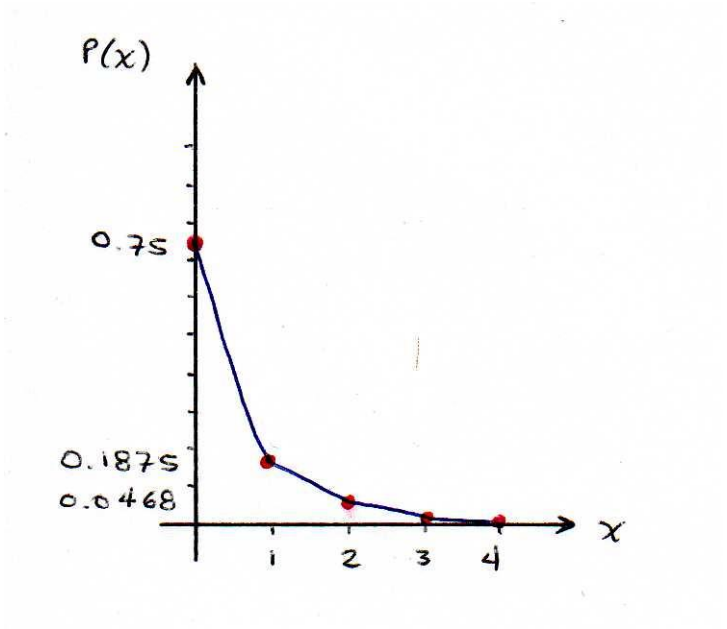
P (x=1) = qp = ¼ • ¾ = 3/16 = 0.1875

P (x=0) = q^0p = 1 x ¾ = 0.75

P (x=2) = q^2p = $(¼)^2$ • ¾ = 0.0468

P (x=3) = q^3p = $(¼)^3$ • ¾ = 0.0117

P (x=4) = q^4p = $(¼)^4$ • ¾ = 0.0029



CALCULO DE LA MEDIA Y DE LA VARIANCIA

μ_x **= E {x} =** $\sum_{\forall x}$ **x P (x)**

POR MEDIO DE LA TRANSFORMADA GEOMETRICA:

P^t (Z) = $\sum_{n=0}^{\infty}$ **q^np Z^n**

P(x)

$$= \sum_{n=0}^{\infty} (1 - q)\, q^n Z^n$$

$$= (1 - q) \sum_{n=0}^{\infty} q^n Z^n \quad \text{SERIE GEOMETRICA}$$

$$= (1 - q)\, \frac{1}{1 - qZ}$$

$$P^t(Z) = \frac{1 - q}{1 - qZ}$$

$$\mu_x = \frac{dP^t(Z)}{dZ}\Big|_{z=1} = \frac{-(1 - q)(-q)}{(+1 - q)(Z)^2}\Big|_{z=1} = \frac{q}{1 - q}$$

$$\sigma_x^2 = \frac{d^2 p^t(Z)}{dZ}\Big|_{z=1} + \mu_x - \mu_x^2 = \frac{q}{(1 - q)^2}$$

$$\mu_x = \frac{q}{1 - q} = \frac{¼}{¾} = ⅓$$

EJEMPLO:

SE SABE QUE EL 20% DE LOS TRANSISTORES QUE PRODUCE UNA INDUSTRIA SON DEFECTUOSOS:

i) CALCULAR LA PROBABILIDAD DE QUE 4 TANSISTORES TOMADOS AL AZAR:

a) **UNO RESULTA DEFECTUOSO**
b) **NINGUNO RESULTA DEFECTUOSO.**
c) **MAS DE 2 SEAN DEFECTUOSOS**
d) **POR LO MENOS 3 SEAN DEFECTUOSOS**

ii) CALCULAR LA PROBABILIDAD DE QUE 5 TRANSISTORES RESULTEN NO DEFECTUOSOS Y QUE EL SEXTO TRANSISTOR SEA DEFECTUOSO.

SOLUCION:

i) 4 TRANSISTORES.

SEA LA V.A. x QUE REPRESENTA EL NUMERO DE TRANSISTORES DEFECTUOSOS.

a) P (x = 1) =?

N = 4 **PROCESO DE BERNOULLI**

p = 0.2 **, q = 0.8**

TRABAJANDO UNA DISTRIBUCION BINOMIAL

$$\text{a) } P(x=1) = {}_4C_1\, p^1 q^3 = \frac{4!}{1!\,(3!)}(0.2)(0.8)^3 = 0.4096 = \underline{40.96\%}$$

$$\text{b) } P(x=0) = {}_4C_0 p^0 q^4 = 0.4096 = \underline{40.96\ \%}$$

$$\text{c) } P(x>2) = P(3) + P(4) = 1 - [P(0) + P(1) + P(2)] = 1-(0.4096+0.4096+0.1536)=0.0272$$
$$= \underline{2.72\%}$$

$$P(x=2) = {}_4C_2 p^2 q^2 = \frac{4!}{2!\,(4-2!)}(0.2)^2(0.8)^2 = 0.1536$$

$$\text{d) } P(3) + P(4) = 0.0256 + 0.0016 = 0.0272 = \underline{2.72\%}$$

$$P(x=3) = {}_4C_3 p^3 q = \frac{4!}{3!\,(4-3!)}(0.2)^3(0.8) = 0.0256$$

$$P(x=4) = {}_4C_4 p^4 q^0 = \frac{4!}{4!\,(4-4!)}(0.2)^4(0.8)^0 = 0.0016$$

ii) P (x = 5) =? **DISTRIBUCION GEOMÉTRICA**

x : TRANSISTOR DEFECTUOSO

P (x = 5) = q^5p = $(0.8)^5$ (0.2) = 0.065536 = 6.55%

ii') CALCULAR LA PROBABILIDAD DE SACAR 5 TRANSISTORES DEFECTUOSOS Y EL 6to. SEA NO DEFECTUOSO.

y : TRANSISTOR NO DEFECTUOSO

p = 0.8
q = 0.2

P (y = 5) = q^5p = $(0.2)^5$ (0.8) = 0.000256 = 0.0256%

DISTRIBUCIÓN BINOMIAL NEGATIVA E HIPERGEOMÉTRICA

LAS DISTRIBUCIONES HIPERGEOMETRICA Y BINOMIAL NEGATIVA ESTAN AMBAS CERCAMENTE RELACIONADAS CON LA DISTRIBUCION BINOMIAL. MIENTRAS QUE LA DISTRIBUCION BINOMIAL ES EL MODELO APROXIMADO DE PROBABILIDAD PARA MUESTREO SIN REEMPLAZO DE UNA POBLACION DICOTOMICA FINITA (S-F), LA DISTRIBUCION HIPERGEOMETRICA ES EL MODELO EXACTO DE PROBABILIDAD PARA EL NUMERO DE LAS *S* DE LA MUESTRA. LA VA X BINOMIAL ES EL NUMERO DE LAS *S* CUANDO EL NUMERO n DE INTENTOS ES FIJO, EN TANTO QUE LA DISTRIBUCION BINOMIAL NEGATIVA RESULTA DE FIJAR EL NUMERO DE LAS *S* Y HACER ALEATORIO EL NUMERO DE INTENTOS.

LAS SUPOSICIONES QUE LLEVAN A LA DISTRIBUCION HIPERGEOMETRICA SON:

1.- LA POBLACION O CONJUNTO DONDE DEBA HACERSE MUESTREO CONSTA DE *N* INDIVIDUOS, OBJETOS O ELEMENTOS (UNA POBLACION *FINITA*).

2.- CADA INDIVIDUO PUEDE SER CARACTERIZADO COMO UN ÉXITO (*S*) O FRACASO (F), Y HAY *M* EXITOS EN LA POBLACION.

3.- SE SACA UNA MUESTRA DE *N* INDIVIDUOS EN FORMA TAL QUE SEA IGUALMENTE PROBABLE QUE SE SELECCIONE CADA SUBCONJUNTO DE TAMAÑO *N*.

LA VARIABLE ALEATORIA DE INTERES ES X = NUMERO DE LAS *S* DE LA MUESTRA. LA DISTRIBUCION DE PROBABILIDAD DE X DEPENDE DE LOS PARAMETROS *n*, M Y N, ASI QUE DESEAMOS OBTENER
P (X = *x*) = h (*x* ; *n*, M, N).

EJEMPLO:

LA BIBLIOTECA DE UNA ESCUELA DE ESTUDIANTES NO GRADUADOS TIENE 20 EJEMPLARES DE CIERTO TEXTO DE INTRODUCCION A LA ECONOMIA, DE LOS CUALES 8 SON PRIMERAS IMPRESIONES Y 12 SON SEGUNDAS IMPRESIONES (QUE CONTIENEN CORRECCIONES DE ALGUNOS PEQUEÑOS ERRORES QUE APARECIERON EN LA PRIMERA EDICION). EL INSTRUCTOR DEL CURSO HA SOLICITADO QUE 5 EJEMPLARES SEAN PUESTOS EN RESERVA DE 2 HORAS. SI LOS EJEMPLARES SE SELECCIONAN EN UNA FORMA POR COMPLETO AL AZAR, DE MODO QUE CADA SUBCONJUNTO DE TAMAÑO 5 TENGA LA MISMA PROBABILIDAD DE SER SELECCIONADO, ¿Cuál ES LA PROBABILIDAD DE QUE *x* (*x* = 0, 1, 2, 3, 4 O 5) DE LOS SELECCIONADOS SEAN SEGUNDAS IMPRESIONES ?

EN ESTE EJEMPLO, EL TAMAÑO DE LA POBLACION ES N = 20, EL TAMAÑO MUESTRAL ES *n* = 5, Y EL NUMERO DE LAS *S* (SEGUNDA IMPRESIÓN = *S*) Y LAS *F* DE LA POBLACION SON M = 12 Y *N* – M = 8, RESPECTIVAMENTE. CONSIDERE EL VALOR x = 2. DEBIDO A QUE TODOS LOS RESULTADOS (CADA UNO FORMADO DE CINCO LIBROS PARTICULARES) SON IGUALMENTE PROBABLES,

$$P(X = 2) = h(2; 5, 12, 20) = \frac{\text{NUMERO DE RESULTADOS CON X = 2}}{\text{NUMERO DE POSIBLES RESULTADOS}}$$

EL NUMERO DE POSIBLES RESULTADOS DEL EXPERIMENTO ES IGUAL AL NUMERO DE FORMAS DE SELECCIONAR 5 DE LOS 20 OBJETOS, CUALQUIERA QUE SEA EL ORDEN, ESTO ES, $\binom{20}{5}$. PARA CONTAR EL NUMERO DE RESULTADOS QUE TENGAN X = 2, OBSERVE QUE HAY $\binom{12}{2}$ FORMAS DE SELECCIONAR 2 DE LAS SEGUNDAS IMPRESIONES, Y PARA CADA UNA DE ESTA FORMAS HAY $\binom{8}{3}$ FORMAS DE SELECCIONAR LAS 3 PRIMERAS IMPRESIONES PARA LLENAR LA MUESTRA. LA REGLA DEL PRODUCTO DA ENTONCES $\binom{12}{2}\binom{8}{3}$ COMO EL NUMERO DE RESULTADOS CON X = 2, ASI QUE:

$$h(2; 5, 12, 20) = \frac{\binom{12}{2}\binom{8}{3}}{\binom{20}{5}} = \frac{77}{323} = 0.238 = \underline{23.8\%}$$

EN GENERAL, SI EL TAMAÑO MUESTRAL *n* ES MENOR QUE EL NUMERO DE EXITOS DE LA POBLACION (M), ENTONCES EL VALOR MAXIMO POSIBLE DE X ES n. SIN EMBARGO, SI M<n (POR EJEMPLO, UN TAMAÑO MUESTRAL DE 25 Y SOLO 15 EXITOS DE LA POBLACION), ENTONCES X PUEDE SER A LO SUMO M. ANALOGAMENTE, SIEMPRE QUE EL NUMERO DE FRACASOS DE LA POBLACION (N - M) REBASE AL TAMAÑO MUESTRAL, EL VALO MINIMO POSIBLE DE X ES 0 (PORQUE TODOS LOS INDIVIDUOS A LOS QUE SE HIZO MUESTREO PODRIAN ENTONCES SER FRACASOS). SIN EMBARGO, SI N – M < n, EL VALOR MINIMO POSIBLE DE X ES n – (N – M). SI RESUMIMOS, LOS VALORES POSIBLES DE X SATISFACEN LA RESTRICCION MAX(0, n – (N – M)) ≤ x ≤ MIN(n, M). UN ARGUMENTO PARALELO AL DEL EJEMPLO ANTERIOR DA LA DISTRIBUCIÓN DE PROBABILIDAD DE X.

PROPOSICION: SI X ES EL NUMERO DE LAS S DE UNA MUESTRA COMPLETAMENTE ALEATORIA DE TAMAÑO n SACADA DE UNA POBLACION FORMADA POR M S y (N – M) F ENTONCES LA DISTRIBUCION DE PROBABILIDAD DE X, LLAMADA DISTRIBUCION HIPERGEOMETRICA, ESTA DADA POR

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

PARA x UN ENTERO QUE SATISFAGA MAX(0, n – N + M) ≤ x ≤ MIN(n, M)

EN EL EJEMPLO ANTERIOR n = 5, M = 12 Y N = 20, POR LO QUE h(x; 5, 12, 20) PARA X = 0, 1, 2, 3, 4, 5 SE PUEDE OBTENER AL SUSTITUIR ESTOS NUMEROS EN LA ECUACION ANTERIOR.

EJEMPLO:

CINCO EJEMPLARES DE UNA POBLACION ANIMAL CONSIDERADOS EN VIAS DE EXTICION EN CIERTA REGION HAN SIDO ATRAPADOS, MARCADOS Y PUESTOS EN LIBERTAS PARA QUE SE MEZCLEN EN LA POBLACION. DESPUES DE QUE HABIAN TENIDO UNA OPORTUNIDAD DE MEZCLARSE, SE SELECCIONO UNA MUESTRA ALEATORIA DE 10 DE ESTOS ANIMALES. SEA X = NUMERO DE ANIMALES MARCADOS DE LA

SEGUNDA MUESTRA. SI HAY EN REALIDAD 25 ANIMALES DE ESTE TIPO EN LA REGION, ¿ CUAL ES LA PROBABILIDAD DE QUE (a) X = 2? (b) X≤2?

LOS VALORES DE PARAMETRO SON n = 10, M = 5 (5 ANIMALES MARCADOS EN LA POBLACION), Y N = 25, POR LO QUE

$$h(x; 10, 5\ 25) = \frac{\binom{5}{x}\binom{20}{10-x}}{\binom{25}{10}} \qquad x = 0, 1, 2, 3, 4, 5$$

PARA (a)

$$P(X = 2) = h(2; 10, 5, 25) = \frac{\binom{5}{2}\binom{20}{8}}{\binom{25}{10}} = 0.385 = \underline{38.5\%}$$

PARA (b)

$$P(X \leq 2) = P(X = 0, 1, \text{o } 2) = \sum_{x=0}^{2} h(x; 10, 5, 25) =$$

$$= 0.057 + 0.257 + 0.385 = 0.699 = \underline{69.9\%}$$

SE DISPONE DE TABLAS COMPLETAS DE DISTRIBUCION HIPERGEOMETRICA, PERO DEBIDO A QUE LA DISTRIBUCION TIENE TRES PARAMETROS, ESTAS TABLAS REQUIEREN MUCHO MAS ESPACIO QUE LAS TABLAS PARA LA DISTRIBUCION BINOMIAL.

LA MEDIA Y VARIANZA DE X

COMO EN EL CASO BINOMIAL, HAY EXPRESIONES SENCILLAS PARA E(X) Y V(X) PARA VARIABLES ALEATORIAS (*va)* HIPERGEOMETRICAS.

PROPOSICION

LA MEDIA Y VARIANZA DE LA VARIABLE ALEATORIA (*va)* X HIPERGEOMETRICA QUE TIENE DISTRIBUCION DE PROBABILIDAD h (x; n, M, N) SON

$$E(X) = n\left[\frac{M}{N}\right], \qquad V(X) = \left[\frac{N-n}{N-1}\right] . n . \frac{M}{N} . \left[1 - \frac{M}{N}\right]$$

LA RAZON M/N ES LA PROPORCION DE LAS S DE LA POBLACION. SI SUSTITUIMOS M/N POR p EN E(X) Y V(X), OBTENEMOS.

$$E(X) = np$$

$$V(X) = \left[\frac{N-n}{N-1}\right] . np\,(1-p)$$

LA EXPRESION ANTERIOR, MUESTRA QUE LAS MEDIAS DE LAS VARIABLES ALEATORIAS (*va)* BINOMIALES E HIPERGEOMETRICAS SON IGUALES, EN TANTO QUE LAS VARIANZAS DE LAS DOS VARIABLES ALEATORIAS (*va)* DIFIEREN POR EL FACTOR (N – n)/(N – 1), A VECES LLAMADO *FACTOR FINITO DE CORRECCION DE POBLACION*. ESTE FACTOR ES MENOR DE 1, ASI QUE LA VARIABLE HIPERGEOMETRICA TIENE MENOR VARIANZA QUE LA VARIABLE ALEATORIA (*va)* BINOMIAL. EL FACTOR DE CORRECCION SE PUEDE ESCRIBIR COMO (1 – n/N)/(1 – 1/N), QUE ES APROXIMADAMENTE 1 CUANDO n ES PEQUEÑA EN RELACION CON N.

EJEMPLO

EN EL EJEMPLO DE MARCAR ANIMALES, n = 10, M = 5 Y N = 25, ASI QUE p = 5/25 = .2 Y

$$E(X) = 10\,(.2) = 2$$

$$V(X) = \frac{15}{24}(10)(.2)(.8) = (.625)(1.6) = 1$$

SI EL MUESTREO SE REALIZA CON REEMPLAZO, V(X) = 1. SUPONGAMOS QUE EL TAMAÑO DE LA POBLACION N NO SE CONOCE EN REALIDAD, POR LO QUE EL VALOR x ES OBSERVADO Y DESEAMOS ESTIMAR N. ES RAZONABLE IGUALAR LA PROPORCION MUESTRAL OBSERVADA DE LAS S, x/n, CON LA PROPORCION DE POBLACION, M/N, DANDO EL ESTIMADO

$$\tilde{N} = \frac{M \cdot n}{x}$$

SI M = 100, n = 40 Y x = 16, ENTONCES Ñ = 250.

DISTRIBUCION HIPERGEOMETRICA

LA FORMA MAS SIMPLE DE PERCATARSE DE LA DIFERENCIA ENTRE LA DISTRIBUCION BINOMIAL Y LA DISTRIBUCION HIPERGEOMETRICA, ES CONOCIENDO LA MANERA COMO SE LLEVA A CABO EL MUESTREO. LOS TIPOS DE APLICACIONES DE LA DISTRIBUCION HIPERGEOMETRICA SON MUY SIMILARES A AQUELLOS DE LA BINOMIAL. EL INTERES SE CENTRA EN EL CALCULO DE LAS PROBABILIDADES PARA EL NUMERO DE OBSERVACIONES QUE CAEN EN UNA CATEGORIA PARTICULAR. SOLO QUE EN EL CASO DE LA BINOMIAL, SE REQUIERE LA INDEPENDENCIA ENTRE INTENTOS. COMO RESULTADO, SI LA DISTRIBUCION BINOMIAL SE APLICA AL MUESTREO DE UN LOTE DE ARTICULOS (PAQUETE DE CARTAS, UNA CANTIDAD DE ARTICULOS DE UNA LINEA DE PRODUCCION), EL MUESTREO DEBE REALIZARSE CON REEMPLAZO DE CADA ARTICULO DESPUES DE OBSERVARSE. POR EL OTRO LADO, LA DISTRIBUCION HIPERGEOMETRICA NO REQUIERE INDENPENDENCIA Y SE BASA EN EL MUESTREO LLEVADO A CABO SIN REEMPLAZO.

LAS APLICACIONES DE LA DISTRIBUCION HIPERGEOMETRICA SE ENCUENTRAN EN MUCHAS AREAS, CON UN USO CONSIDERABLE EN EL MUESTREO DE ACEPTACION, LAS PRUEBAS ELECTRONICAS Y EL ASEGURAMIENTO DE LA CALIDAD. ES OBVIO QUE EN MUCHOS DE ESTOS CAMPOS LA PRUEBA SE REALIZA A EXPENSAS DE LA PIEZA QUE SE ESTA PROBANDO. ESTA SE DESTRUYE Y POR LO TANTO NO PUEDE REEMPLAZARSE EN LA MUESTRA. ENTONCES, ES NECESARIO EL MUESTREO SIN REEMPLAZO. EN LOS SIGUIENTES PARRAFOS SE UTILIZA COMO EJEMPLO UN CASO SIMPLE DE JUEGO DE CARTAS.

SE DESEA ENCONTRAR LA PROBABILIDAD DE OBTENER 3 CARTAS ROJAS EN 5 RETIROS DE UN PAQUETE COMUN DE 52 CARTAS. PARA RESOLVER LA SITUACION DE MUESTREAR SIN REEMPLAZO, EL PROBLEMA DEBE REPLANTEARSE. SI SE SACAN 5 CARTAS AL AZAR Y SE DESEA CONOCER LA PROBABILIDAD DE SELECCIONAR 3 CARTAS ROJAS DE LAS 26 DISPONIBLES Y 2 CARTAS NEGRAS DE LAS OTRAS 26 DISPONIBLES EN EL PAQUETE. SE TIENEN ($\binom{26}{3}$) MANERAS DE

SELECCIONAR 3 CARTAS ROJAS, Y POR CADA UNA DE ESTAS FORMAS SE PUEDEN SELECCIONAR 2 CARTAS NEGRAS EN ($^{26}_{2}$) MANERAS. POR LO TANTO, EL NUMERO TOTAL DE MODOS DE SELECCIONAR 3 CARTAS ROJAS Y 2 NEGRAS EN 5 INTENTOS ES EL PRODUCTO ($^{26}_{3}$)($^{26}_{2}$). EL NUMERO TOTAL DE MANERAS DE SELECCIONAR 5 CARTAS CUALQUIERA DE LAS 52 QUE ESTAN DISPONIBLES ES ($^{52}_{5}$). DE AQUÍ QUE LA PROBABILIDAD DE SELECCIONAR 5 CARTAS SIN REEMPLAZO DE LAS CUALES 3 SEAN ROJAS Y 2 NEGRAS ES:

$$\frac{\binom{26}{3}\binom{26}{2}}{\binom{52}{5}} = \frac{(26!/3!23!)(26!/2!24!)}{(52!/5!47!)} = 0.3251 = \underline{32.51\%}$$

EN GENERAL, EL INTERES QUE SE TIENE ES EN LA PROBABILIDAD DE SELECCIONAR x ÉXITOS DE LOS k POSIBLES RESULTADOS O ARTICULOS TAMBIEN CONSIDERADOS EXITOS Y n – x FRACASOS DE LOS N – k POSIBLES RESULTADOS O ARTICULOS TAMBIEN CONSIDERADOS FRACASOS, CUANDO UNA MUESTRA ALEATORIA DE TAMAÑO n SE SELECCIONA DE N RESULTADOS O ARTICULOS TOTALES. ESTO SE CONOCE COMO UN EXPERIMENTO HIPERGEOMETRICO.

UN EXPERIMENTO HIPERGEOMETRICO ES AQUEL QUE POSEE LAS DOS PROPIEDADES SIGUIENTES:

1. **UNA MUESTRA ALEATORIA DE TAMAÑO n SE SELEECIONA SIN REEMPLAZO DE UN TOTAL DE N RESULTADOS O ARTICULOS TOTALES.**

2. **k RESULTADOS O ARTICULOS DEL TOTAL N PUEDEN CLASIFICARSE COMO EXITOS Y N – k COMO FRACASOS**

EL NUMERO X DE EXITOS EN UN EXPERIMENTO HIPERGEOMETRICO RECIBE EL NOMBRE DE *VARIABLE ALEATORIA HIPERGEOMETRICA*. DE ACUERDO CON ESTO, LA DISTRIBUCION DE PROBABILIDAD DE LA VARIABLE HIPERGEOMETRICA SE LLAMA *DISTRIBUCION HIPERGEOMETRICA* Y SUS VALORES SON REPRESENTADOS POR h(x; N, n, k), DADO QUE DEPENDEN DEL NUMERO DE EXITOS k EN EL CONJUNTO N DEL CUAL SE SELECCIONAN n RESULTADO O ARTICULOS.

<u>DISTRIBUCIÓN BINOMIAL NEGATIVA (PASCAL).</u>

LA VARIABLE ALEATORIA BINOMIAL NEGATIVA Y SU DISTRIBUCIÓN ESTÁN BASADAS EN UN EXPERIMENTO QUE SATISFACE LAS SIGUIENTES CONDICIONES:

1. **EL EXPERIMENTO CONSTA DE UNA SECUENCIA DE INTENTOS INDEPENDIENTES.**
2. **CADA INTENTO PUEDE RESULTAR EN UN ÉXITO (S) O EN UN FRACASO (F).**
3. **LA PROBABILIDAD DE ÉXITO ES CONSTANTE DE UN INTENTO A OTRO, ASÍ QUE P(S EN EL INTENTO i) = p PARA i = 1,2,3,...**
4. **EL EXPERIMENTO CONTINÚA (LOS INTENTOS SE EJECUTAN) HASTA QUE UN TOTAL DE r ÉXITOS SE HAYA OBSERVADO, DONDE r ES UN ENTERO POSITIVO ESPECIFICADO.**

LA VARIABLE ALEATORIA DE INTERÉS ES x = NÚMERO DE FRACASOS QUE PRECEDEN AL r-ésimo ÉXITO; x SE LLAMA *VARIABLE ALEATORIA BINOMIAL NEGATIVA* PORQUE, EN CONTRASTE CON LA VARIABLE ALEATORIA BINOMIAL, EL NÚMERO DE ÉXITOS ES FIJO Y EL NÚMERO DE INTENTOS ES ALEATORIO.

LOS POSIBLES VALORES DE x SON 0, 1, 2,.... DENOTEMOS POR nb(x ; r , p) LA DISTRIBUCIÓN DE PROBABILIDAD DE x. EL EVENTO $\{X = x\}$ **ES EQUIVALENTE A (r – 1 S EN LOS PRIMEROS (x + r –1) INTENTOS Y UNA S EN EL (x + r)-ésimo INTENTO) (POR EJEMPLO, SI r = 5 Y x = 10, ENTONCES DEBE HABER CUATRO S EN LOS PRIMEROS 14 INTENTOS Y EL INTENTO 15 DEBE SER UNA S). COMO LOS INTENTOS SON INDEPENDIENTES:**

$$nb(x ; r , p) = P(X = x)$$
$$= P(\text{r – 1 S´s EN LOS PRIMEROS x + R – 1 intentos})\, P(S)$$

LA PRIMERA PROBABILIDAD DE LA EXTREMA DERECHA DE LA EXPRESIÓN ES LA PROBABILIDAD BINOMIAL

$$\binom{x + r - 1}{r - 1} p^{r-1} (1 - p)^{x}, \qquad \text{MIENTRAS } P(S) = p$$

LA DISTRIBUCIÓN DE PROBABILIDADA DE LA VARIABLE ALEATORIA BINOMIAL NEGATIVA CON PARÁMETROS r = NÚMERO DE LAS S Y p = P(S) ES:

$$nb(x ; r , p) = \binom{X + r - 1}{r - 1} p^{r} (1 - p)^{x}, \qquad X = 0, 1, 2, \ldots$$

EJEMPLO:

1. **UN PEDIATRA DESEA RECLUTAR 5 PAREJAS, CADA UNA DE LAS CUALES ESPERA TENER A SU PRIMER HIJO, PARA QUE PARTICIPEN EN UN NUEVO RÉGIMEN DE NACIMIENTO NATURAL. SEA p = P(UNA PAREJA SELECCIONADA AL AZAR ACCEDE A PARTICIPAR). SI p = 0.2 , ¿CUÁL ES LA PROBABILIDADA DE QUE SE PIDA A 15 PAREJAS QUE PARTICIPEN ANTES DE ENCONTRARA 5 QUE ACCEDAN? ESTO ES, CON S = (ACCEDE A PARTICIPAR), ¿CUÁL ES LA PROBABILIDADA DE QUE OCURRAN 10 F ANTES DE LA QUINTA S?**

p = 0.2

r = 5

x = 10

$$nb(10;5,0.2) = \binom{10+5-1}{5-1} (0.2)^5 (0.8)^{10} = 0.03439 = \underline{3.44\%}$$

LA PROBABILIDAD DE QUE AL MENOS 10 F SE OBSERVEN (A LO SUMO SE PIDA A 15 PAREJAS) ES:

$$P(x \leq 10) = \sum_{x=0}^{10} nb(x;5,0.2) = (0.2)^5 \sum_{x=0}^{10} \binom{x+4}{4} (0.8)^x = 0.1642 = 16.42\%$$

EN ALGUNAS PUBLICACIONES, LA VARIABLE ALEATORIA BINOMIAL NEGATIVA SE TOMA COMO EL NÚMERO DE INTENTOS X + r EN LUGAR DEL NÚMERO DE FRACASOS.

EN EL CASO ESPECIAL r = 1, LA DISTRIBUCIÓN DE PROBABILIDAD ES:

$$nb(x;1,p) = (1-p)^x p \quad , \qquad x = 0, 1, 2, \ldots..$$

EN UNA DISTRIBUCION DE PROBABILIDAD DE LA VARIABLE ALEATORIA PARA EL CASO DE UNA DISTRIBUCION BINOMIAL NEGATIVA, EL NÚMERO DE INTENTOS NECESARIOS PARA OBTENER LA PRIMERA S, Y LA DISTRIBUCIÓN DE PROBABILIDAD AHÍ ES SIMILAR A LA EXPRESIÓN ANTERIOR PARA EL CASO EN QUE r = 1.

TANTO x = NÚMERO DE LAS F COMO y = NÚMERO DE INTENTOS (= 1 + X) SE CONOCEN EN LA LITERATURA COMO VARIABLES ALEATORIAS

GEOMÉTRICAS, Y LA DISTRIBUCIÓN DE PROBABILIDADA SE LLAMA DISTRIBUCIÓN GEOMÉTRICA.

POR OTRO LADO EL NÚMERO ESPERADO DE INTENTOS HASTA LA PRIMERA S SE DEMUESTRA QUE ES 1/p DE MODOD QUE EL NÚMERO ESPERADO DE LAS F HASTA LA PRIMERA S ES (1/p) – 1 = (1 – p) / p . INTUITIVAMENTE, ESPERARÍAMOS VER r (1 – p)/ p F ANTES DE LA r-ésima S, Y ESTO ES CIERTAMENTE E(X). TAMBIÉN HAY UNA FÓRMULA SENCILLA PARA OBTENER V(X).

SI X ES UNA VARIABLE ALEATORIA BINOMIAL NEGATIVA CON DISTRIBUCIÓN DE PROBABILIDADA nb (x ; r , p), ENTONCES:

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

FINALMENTE, AL EXPANDIR EL COEFICIENTE BINOMIAL EN FRENTE DE $p^r (1 – p)^x$ Y HACER ALGUNA CANCELACIÓN, SE PUEDE VER QUE nb (x ; r , p) ESTA BIEN DEFINIDO AUN CUANDO r NO SEA UN ENTERO. SE HA ENCONTRADO QUE ESTA DISTRIBUCIÓN BINOMIAL NEGATIVA GENERALIZADA SE AJUSTA MUY BIEN A LOS DATOS OBSERVADOS EN UNA AMPLIA VARIEDAD DE APLICACIONES.

DISTRIBUCION DE POISSON

ESTA DISTRIBUCION SE PUEDE OBTENER A PARTIR DE LA HAY DISTRIBUCION BINOMIAL CONSIDERANDO QUE $N \to \infty$, MIENTRAS QUE $P \to 0$

$$P(x = n) = {}_NC_n\, P^n q^{N-n} \qquad ; \qquad x = 0, 1, 2, \ldots\ldots, N$$

$$P(x = n) = \frac{N!}{n!\,(N-n)!}\, P^N q^{n-N}$$

$$P(x = n) = \frac{N(N-1)(N-2)\ldots.[N-(n-1)][N-n]}{n![N-n]}\, P^n q^{N-n}$$

$$P(x = n) = \frac{N(N-1)(N-2)\ldots.[N-(n-1)}{n!}\, p^n q^{N-n}$$

$$= \frac{N}{N}\,\frac{N-1}{N}\,\frac{N-2}{N}\,\ldots\ldots\,\frac{N-(n-1)}{N}\,Nn\,\frac{P^n\,q^{N-n}}{n}$$

$$= (1)\,(1 - 1/N)\,1 - 2/N)\,\ldots..\,(1 - \frac{n-1}{N})\,\frac{(NP)^n\,q^{N-n}}{n!}$$

$$\frac{(PN)^n\,(1-P)^N}{n!\,(1-P)^N}$$

HACIENDO $\lambda = NP$ **;** $P = \frac{\lambda}{N}$ **, SE TIENE**

$$P\,(x = n) = (1 - 1/N)\,(1 - 2/N)\,\ldots..(1 - \frac{n-1}{N})\,\frac{\lambda^n\,(1 - \lambda/N)^N}{n!\,(1 - \lambda/N)^N}$$

$$P(x) = \lim_{N \to \infty}\,[\,(1 - \frac{1}{N})\,(1 - \frac{2}{N})\ldots.(1 - \frac{n-1}{N})\,\frac{\lambda^n\,(1 - \lambda/N)^N}{n!\,(1 - \lambda/N)^n}]$$

$$= \lim_{N \to \infty}\,\frac{\lambda^n\,(1 - \lambda/N)^N}{n!}$$

$$= \frac{\lambda^n}{n!}\,\lim_{N \to \infty}\,(1 - \frac{}{})^N$$

$$= \frac{\lambda^n}{n!}\,e^{-2}$$

$$P\,(x) = \frac{\lambda^n}{n!}\,e^{-2} \qquad ; \qquad n = 0, 1, 2, \ldots,$$

DISTRIBUCION DE POISSON.

LA FUNCION DE PROBABILIDAD QUE DEFINE ESTA DISTRIBUCION ES:

$$P(x = n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad ; \quad n = 0,1,2,\ldots,$$

O BIEN:

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad ; \quad x = 0,1,2,\ldots.$$

COMO YA SE OBSERVO, ESTA DISTRIBUCION ES UNA APROXIMACION A LA DISTRIBUCION BINOMIAL PARA LOS CASOS EN QUE N SEA GRANDE Y P SEA PEQUEÑA,

$$N \geq 50 \quad , \quad Np \leq 5$$

EJEMPLO:

SI EL 2% DE LOS TRANSISTORES PRODUCIDOS POR UNA FABRICA SALEN DEFECTUOSOS, ENCONTRAR LA PROBABILIDAD DE QUE EN UN LOTE DE 200 TRANSISTORES HAYA CUANDO MAS 5 TRANSISTORES DEFECTUOSOS.

EL PROBLEMA SE TRABAJA CON UNA DISTRIBUCION BINOMIAL, EN LA CUAL:

$$N=200 \quad , \quad p=0.02 \quad , \quad q=0.98$$

ÉXITO: EL TRANSISTOR ESTA DEFECTUOSO

$$P(x \leq 5) = ?$$

$$P(x \leq 5) = P(x=0) + P(x=1) + P(x=2) + \ldots\ldots + P(x=5)$$

$$= {}_{200}C_0\, p^0\, q^{200} + {}_{200}C_1\, p^1\, q^{199} + \ldots\ldots\ldots + {}_{200}C_5\, p^5\, q^{195}$$

COMO N ES GRANDE Y p ES PEQUEÑA PODEMOS APROXIMAR LA DISTRIBUCIION BINOMIAL POR MEDIO DE LA DE POISSON.

λ = **Np**

λ = **200 (.02)**

λ = **4**

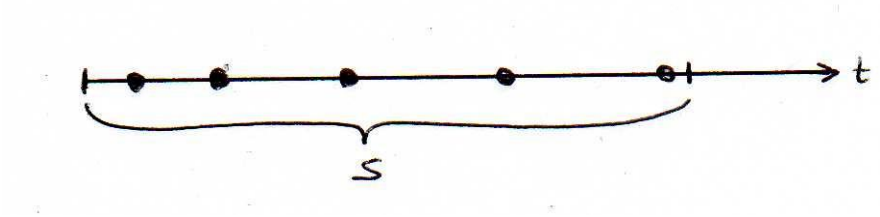
P(x≤5) = P(x=0) + P(x=1) + …… + P(x=5)

$$= \frac{4^0}{0!} e^{-4} + \frac{4}{1!} e^{-4} + \ldots\ldots + \frac{4^5}{5!} e^{-4}$$

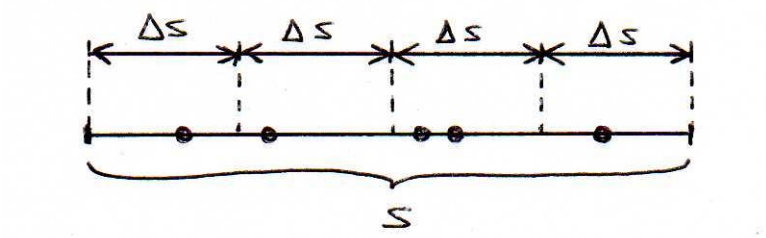
$$= e^{-4}\left(\frac{4^0}{0!} + \frac{4}{1!} + \frac{4^2}{2!} + \frac{4^3}{3!} + \frac{4^4}{4!} + \frac{4^5}{5!}\right)$$

= 0.7851

EN UN PROCESO DE POISSON LOS SUCESOS APARECEN EN INTERVALOS CONTINUOS (TIEMPO, VOLUMEN, LONGITUD, ETC.) DE MAGNITUD S.



SI SE DIVIDE EN INTERVALOS S (MAS PEQUEÑOS) EN SUBINTERVALOS DE MAGNITUD ΔS , DE TAL MANERA QUE LA PROBABILIDAD DE QUE SE PRESENTE UN PUNTO EN EL SUBINTERVALO SEA "p" Y LA PROBABILIDAD DE QUE NP SE PRESENTE "q".



DE ESTA MANERA LA PROBABILIDAD DE QUE SE PRESENTEN DOS O MAS PUNTOS EN UN SUBINTERVALO SE CONSIDERARA NULA O CASI CERO.

SI AL SEGMENTO s LO DIVIDIMOS EN UN NUMERO GRANDE N DE SUBINTERVALOS, ENTONCES:

$$\Delta S = \frac{S}{N} \qquad \ldots\ldots \text{(a)}$$

ADEMAS LA PROBABILIDAD DE ÉXITO P, ES DIRECTAMENTE PROPORCIONAL A ΔS, ESTO ES:

$$P = \lambda \Delta S \qquad \ldots\ldots \text{(b)}$$

DONDE λ ES UNA CONSTANTE DE PROPORCIONALIDAD.

SUSTITUYENDO (a) EN (b)

$$P = \lambda \frac{S}{N} \qquad \ldots\ldots \text{(c)}$$

CON LO CUAL

$$q = 1 - p = 1 - \lambda \frac{S}{N} \qquad \ldots\ldots \text{(d)}$$

EN CADA UNO DE LOS SUBINTERVALOS SE VERIFICA UNA PRUEBA DE BERNOILLI, EN DONDE:

ÉXITO = EN ΔS HAY UN PUNTO

FRACASO = EN ΔS NO HAY PUNTO

POR LO TANTO LA DISTRIBUCION BINOMIAL ES :

$$P(x=n) = {}_N C_n\, p^n\, q^{N-n} = {}_N C_n \left(\lambda \frac{S}{N}\right)^n \left(1 - \lambda \frac{S}{N}\right)^{N-n}$$

CONSIDERANDO QUE $N \to \infty$

$$P(x = n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad ; \quad n = 0,1,2,\ldots,$$

SI S = 1 :

$$P(x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad ; \quad x = 0,1,2,\ldots.$$

PROBLEMA:

1.- EL NUMERO DE MUERTES POR SUICIDIO EN UNA CIUDAD ES DE DOS POR DIA, SI LA VARIABLE ALEATORIA x REPRESENTA EL NUMERO DE MUERTES POR SUICIDIO EN UNA SEMANA, DETERMINAR LA PROBABILIDAD DE QUE HALLA:

a) **14 MUERTES POR SEMANA**
b) **AL MENOS 3 MUERTES POR SEMANA**

A) P (x = 14) = ¿?

λ **= 2 muertes / día = 14 muertes / semana**

$$P(x = 14) = \frac{14^{14}}{14!} e^{-14} = 0.1059 = \underline{10.59\%}$$

P (x >= 3) = P (x = 3) + P (x = 4) + …. + P (x = 14) =
= 1 - (P (x = 0) + P (x = 1) + P (x = 2)) =
= 1 - ($14^0/0! + 14^1/1! + 14^2/2!$) e^{-14} = 0.9999 = <u>100%</u>

2.- SUPONGAN QUE EN EL NUMERO DE LLAMADAS TELEFÓNICAS QUE LLEGAN A UN CONMUTADOR ES DE 120 POR HORA. DETERMINAR LA PROBABILIDAD DE QUE:

a) **EN UN MINUTO SE RECIBAN CERO LLAMADAS**
b) **EN UN MINUTO LLEGEN ENTRE UNA Y CINCO LLAMADAS**

DISTRIBUCIÓN POISSON

λ **= 120 LLAMADAS / HORA = 2 LLAMADAS / MINUTO**

A) P (x = 0) = $(2)^0/0! e^{-2}$ = 0.1353 = <u>13.53%</u>

B) P (1 <= x <= 5) = P (x = 1) + P (x = 2) + P (x = 3) + P (x = 4)
+ P (x = 5) = ($2^1/ 1! + 2^2/2! + 2^3/3! + 2^4/ 4! + 2^5/5!$) e^{-2}
=0.8481 = <u>84.81%</u>

DISTRIBUCIÓN POISSON

CALCULO DE LA MEDIA

$\mu = \lambda$

CALCULO DE LA VARIANCIA

$\sigma = \lambda$