Variables aleatorias discretas notables

Variable aleatoria Bernoulli

Caracteriza la ocurrencia o no de un suceso al realizar un experimento aleatorio.

Variable aleatoria Binomial

Cuenta el número de ocurrencias de un evento de interés en una serie de repeticiones independientes de un experimento aleatorio, como por ejemplo al extraer una muestra aleatoria con reposición de una población.

Variables aleatorias discretas notables

Variable aleatoria hipergeométrica

Cuenta el número de elementos de cierta clase en una muestra aleatoria extraída sin reposición de una población que contiene elementos de esa clase.

Variable aleatoria geométrica

Cuenta el número de repeticiones independientes de un experimento aleatorio hasta la primera ocurrencia de un evento de interés.

Variables aleatorias discretas notables

Variable aleatoria Poisson

Asociada a un especial proceso denominado de Poisson. Sirve para contar número de ocurrencias de un hecho en un tiempo dado.

Puede usarse como aproximación a la distribución binomial.

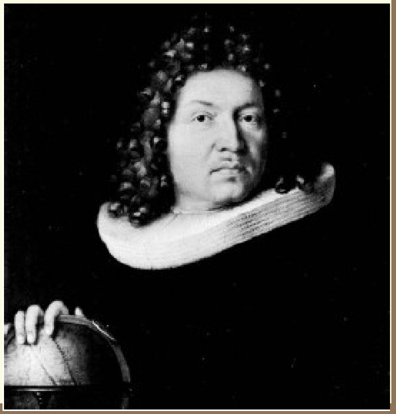
Variable aleatoria Bernoulli

Caracteriza la ocurrencia o no de un suceso al realizar un experimento aleatorio.

$$X \quad \mathit{Bernoulli}(p)$$

Recorrido:

- 0 = “fracaso”
- 1 = “éxito”



Variable aleatoria Bernoulli

$$X \sim Bernoulli(p)$$

$$p_X(1) = p, p_X(0) = 1 - p = q$$

$$\mu = E[X] = 0 \times (1-p) + 1 \times p = p$$

$$E[X^2] = 0^2 \times (1-p) + 1^2 \times p = p$$

$$\sigma^2 = E[X^2] - (E[X])^2 = p - p^2 = p(1-p) = pq$$

Variable aleatoria binomial

Cuenta el número de ocurrencias de un evento de interés en una serie de repeticiones independientes de un experimento aleatorio, como por ejemplo al extraer una muestra aleatoria con reposición de una población.

$$X \quad Binomial(n, p)$$

- Se realizan experimentos **idénticos e independientes** con igual probabilidad de éxito en cada experimento.
- número de éxitos en los experimentos

Recorrido:

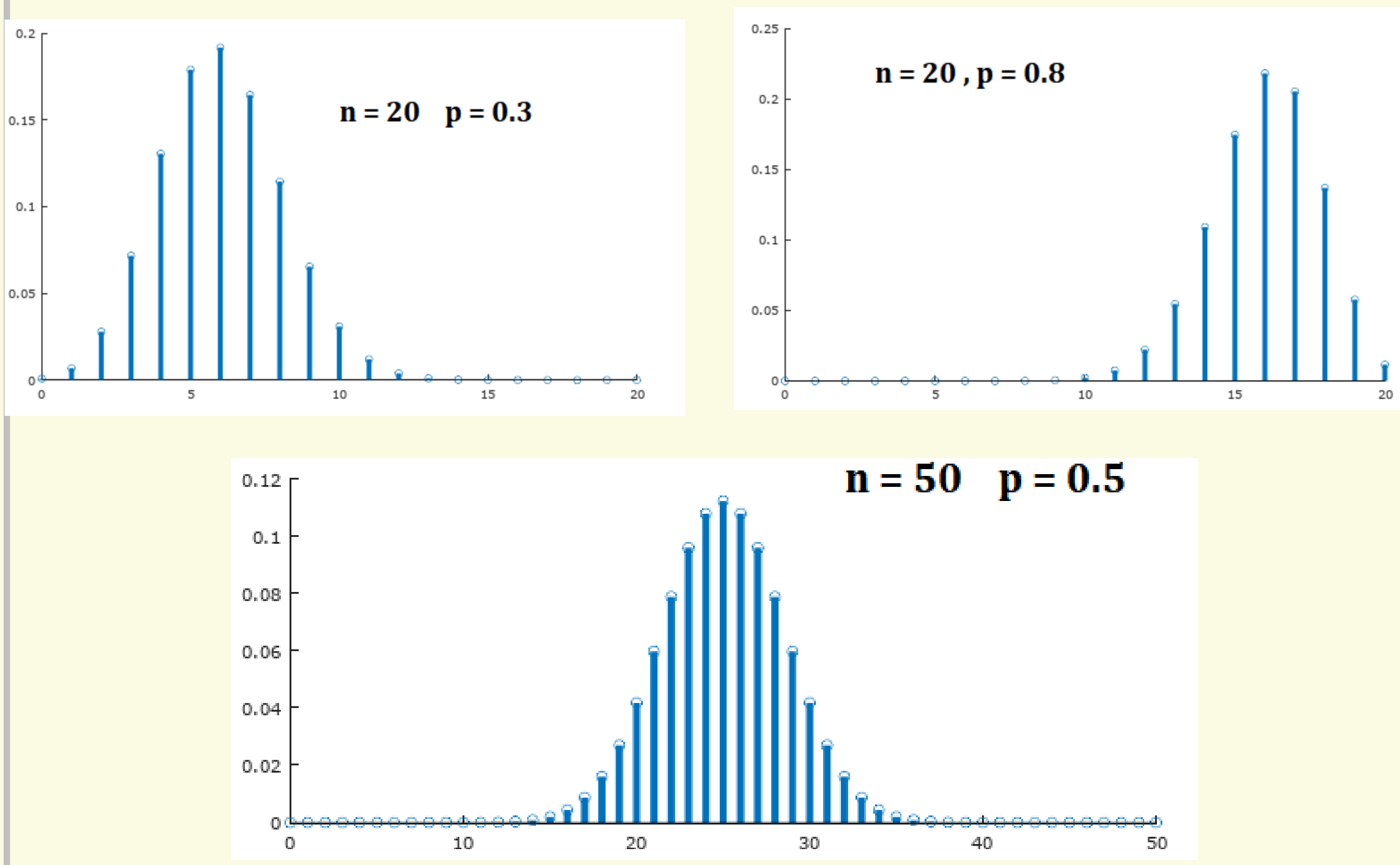
Variable aleatoria binomial

$$X \quad Binomial(n, p)$$

Recorrido:

$$p_X(k) = \binom{n}{k} \times p^k \times q^{n-k} \; \forall k \in R_X$$

Variable aleatoria binomial



Variable aleatoria binomial

$$X \sim Binomial(n, p)$$

Recorrido:

$$p_X(k) = \binom{n}{k} \times p^k \times q^{n-k} \; \forall k \in R_X$$

$$\mu = E[X] = \sum_{k=0}^{n} k \times \binom{n}{k} \times p^k \times q^{n-k}$$

Variable aleatoria binomial

$$\mu=\sum_{k=0}^{n} k\times\binom{n}{k}\times p^{k}\times q^{n-k}$$

$$\mu=\sum_{k=0}^{n} k\times\frac{n!}{k!\left(n-k\right)!}\times p^{k}\times q^{n-k}$$

$$\mu=\sum_{k=1}^{n}\frac{n\times\left(n-1\right)!}{\left(k-1\right)!\left(n-1-\left(k-1\right)\right)!}\times p^{k}\times q^{n-k}$$

$$\mu=n\times p\times\sum_{k=0}^{n-1}\frac{\left(n-1\right)!}{\left(k-1\right)!\left(n-1-\left(k-1\right)\right)!}\times p^{k-1}\times q^{n-k}$$

$$\mu=n\times p\times\left(p+q\right)^{n-1}=n\times p$$

Variable aleatoria binomial

$$X \sim Binomial(n, p)$$

Recorrido:

$$p_X(k) = \binom{n}{k} \times p^k \times q^{n-k} \; \forall k \in R_X$$

$$\mu = E[X] = n \times p$$

$$\sigma^2 = Var[X] = n \times p \times q$$

¡Demostrar!

Una fábrica de baterías entrega su producto en lotes de gran tamaño (1000 baterías). El comprador habitual desea rechazar los lotes que selecciona al azar y controla 20 baterías de cada lote. Si encuentra 4 o más baterías defectuosas rechaza el lote entero. ¿Cuál es la probabilidad de que el lote sea rechazado si el porcentaje de baterías defectuosas es:

a) 5 %,

b) 10 %,

c) ¿Qué sucedería con las anteriores probabilidades si el número crítico de rechazos fuera aumentado de 4 a 5?

: número de baterías defectuosas en la muestra de tamaño 20 (como el tamaño poblacional es muy grande se asume que se extrajeron con reposición)

: El lote se rechaza tras el control de calidad de las 20 baterías

$$P(A)=P(X\geq 4)=1-P(X\leq 4)=1-\sum_{k=0}^{3}\binom{20}{k}\times p^{k}\times q^{20-k}$$

Una fábrica de baterías entrega su producto en lotes de gran tamaño (1000 baterías). El comprador habitual desea rechazar los lotes que selecciona al azar y controla 20 baterías de cada lote. Si encuentra 4 o más baterías defectuosas rechaza el lote entero. ¿Cuál es la probabilidad de que el lote sea rechazado si el porcentaje de baterías defectuosas es:

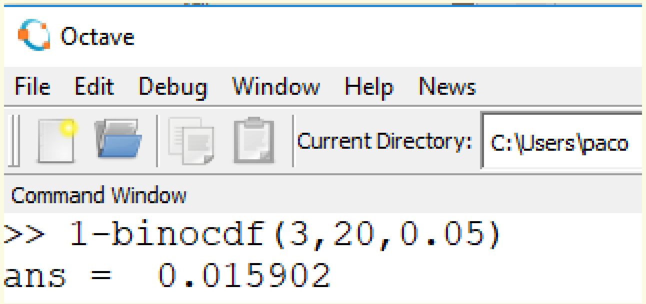
a) 5 %,

b) 10 %,

c) ¿Qué sucedería con las anteriores probabilidades si el número crítico de rechazos fuera aumentado de 4 a 5?

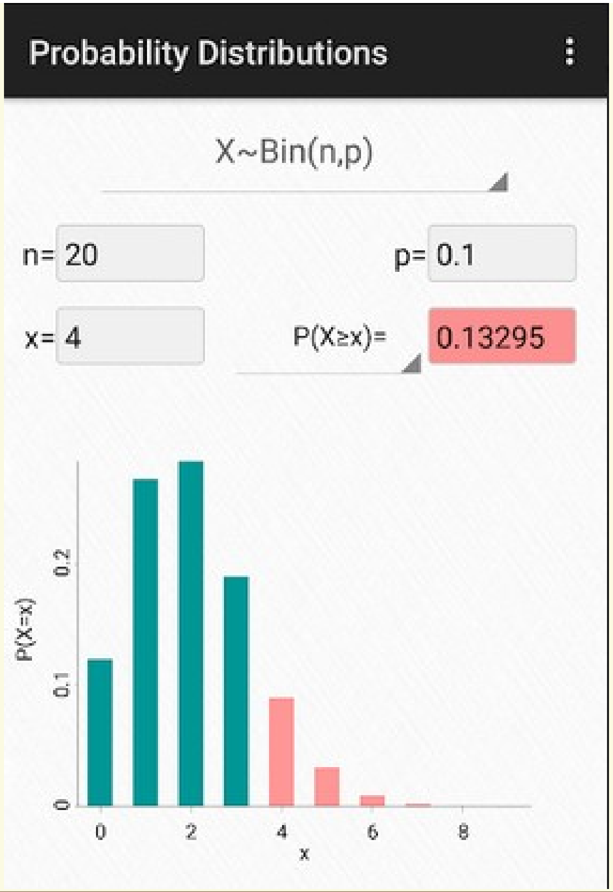
$$P(A)=P(X\geq 4)=1-P(X\leq 4)=1-\sum_{k=0}^{3}\binom{20}{k}\times p^k\times q^{20-k}$$

p	*P*(*A*) = *P*(*X* >= 4)	*P*(*A*) = *P*(*X* >= 5)
0.05	0.0159	0.0026
0.1	0.1330	0.0432



Variables aleatorias discretas notables

$$P(A) = P(X \geq 4) = 1 - P(X \leq 3) = 1 - \sum_{k=0}^{3} \binom{20}{k} p^k (1-p)^{20-k}$$



Esta aplicación permite calcular probabilidades puntuales de variables discretas y también acumuladas como en este ejemplo. Las celdas claras se completan con datos y la roja devuelve el resultado. Dispone de ayudas, fórmulas, así como el valor esperado y la varianza.

Variable aleatoria hipergeométrica

Cuenta el número de elementos de cierta clase en una muestra aleatoria extraída sin reposición de una población que contiene elementos de esa clase.

$$X \quad Hipergeom\acute{e}trica\left(N, M, n\right)$$

- Una población de *individuos*, divididos en 2 categorías, A y B.
- La categoría A tiene individuos, y la B tiene .
- Se toma una *muestra* de individuos de la población al azar **sin reposición.**
- número de individuos de la clase A que salen.

Variable aleatoria hipergeométrica

$$X \quad Hipergeométrica(N, M, n)$$

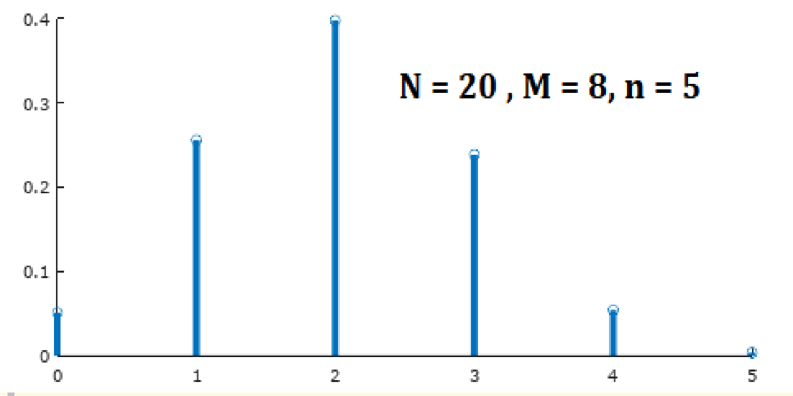
Recorrido:

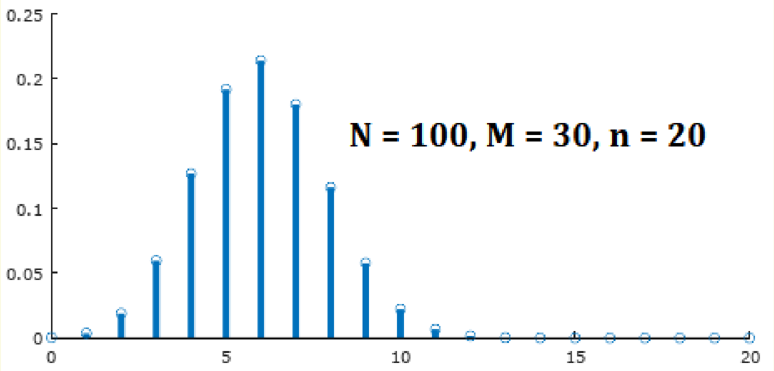
$$p_X(k) = \frac{\binom{M}{k} \times \binom{N-M}{n-k}}{\binom{N}{n}} \; \forall k \in R_X$$

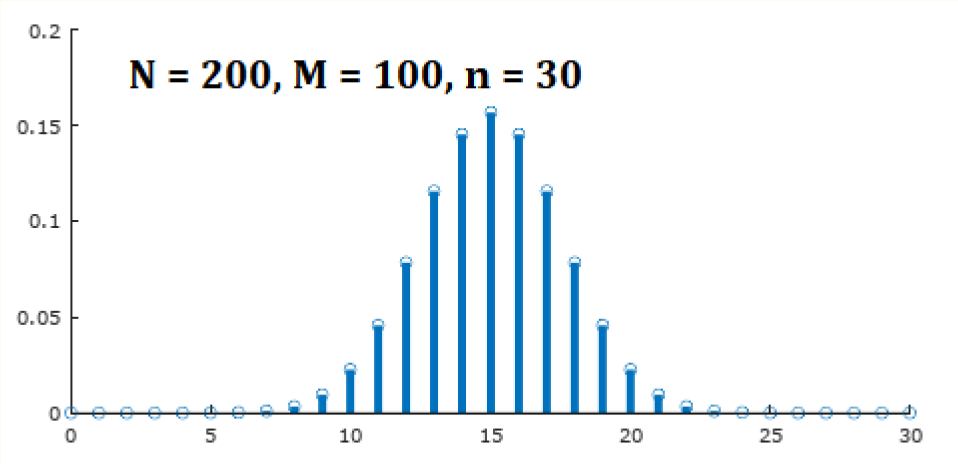
$$\mu = E[X] = n \times \frac{M}{N}$$

$$\sigma^2 = Var[X] = n \times \frac{M}{N} \times \frac{N-M}{N} \times \frac{N-n}{N-1}$$

Variable aleatoria hipergeométrica







Variable aleatoria hipergeométrica

$$X \quad Hipergeom\acute{e}trica\left(N, M, n\right)$$

$$\mu = E\left[X\right] = n \times \frac{M}{N}$$

$$\sigma^2 = Var\left[X\right] = n \times \frac{M}{N} \times \frac{N-M}{N} \times \frac{N-n}{N-1}$$

Interesante observar:

- es la proporción de elementos de la clase A, el valor esperado es (como en la binomial)
- El factor es menor que 1 la varianza en este caso es menor que la de una variable binomial para los mismos valores de y .

Variable aleatoria hipergeométrica

Relación con la binomial:

- Si la población y la cantidad de individuos de cada categoría es mucho más grande que el tamaño de la muestra () **no hace gran diferencia si la muestra se toma con o sin reposición** (¡hay un ejercicio en la guía!).

- Se puede **aproximar una variable aleatoria hipergeométrica por una binomial.**

De un mazo de naipes ordinarios (52 naipes) bien barajado se extraen cinco cartas, al azar (sin sustitución), para una mano de poker.

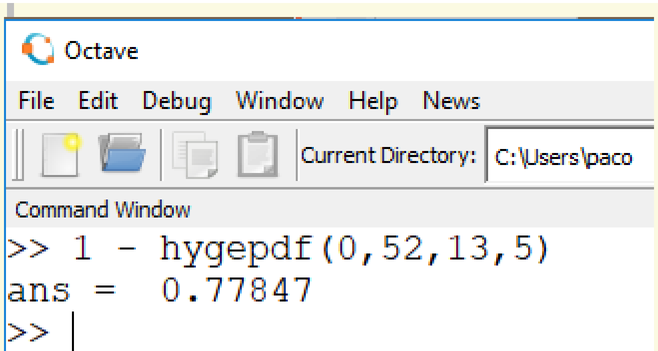
Calcule la probabilidad de sacar por lo menos un trebol.
Calcule la probabilidad de sacar por lo menos dos ases.

: número de tréboles en la mano de 5 cartas (el mazo de 52 cartas está partido en 4 clases, una de ellas es tréboles)

: número de ases en la mano de 5 cartas (en el mazo de 52 cartas hay 4 ases)

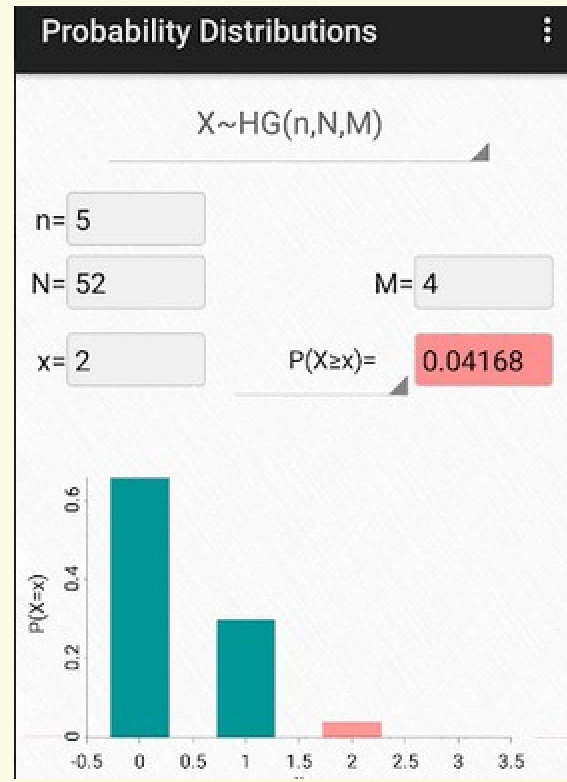
De un mazo de naipes ordinarios (52 naipes) bien barajado se extraen cinco cartas, al azar (sin sustitución), para una mano de poker.

Calcule la probabilidad de sacar por lo menos un trebol.



En Octave hygepdf calcula la función de probabilidad y puede verse el orden de los parámetros de invocación (k, N, M, n)

Calcule la probabilidad de sacar por lo menos dos ases



En la app mostrada se usa la misma notación introducida aquí. Se puede elegir calcular la probabilidad puntual o acumulada.

Variable aleatoria geométrica

Cuenta el número de repeticiones independientes de un experimento aleatorio hasta la primera ocurrencia de un evento de interés.

$$X \; Geom\acute{e}trica(p)$$

- Se realizan **experimentos idénticos e independientes**, con igual probabilidad de éxito en cada experimento.
- número de intentos hasta el primer éxito **(incluido)**.

Recorrido:

$$p_X(k) = (1-p)^{k-1} p \;\forall k \in R_X$$

Este es un ejemplo donde el recorrido de X no es finito, pero si numerable. La condición de cierre corresponde a la suma de una serie convergente.

Variable aleatoria geométrica

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \; si \, |r| < 1$$

Serie geométrica

$$\sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=0}^{\infty} (1-p)^k = ¿ \, \frac{p}{1-(1-p)} = 1 ¿$$

¡Tarea para la casa!:

$$\mu = E[X] = \frac{1}{p}$$

$$\sigma^2 = Var[X] = \frac{1-p}{p^2}$$

Variable aleatoria geométrica

Probabilidad de la cola

$$P(X>M)=1-P(X\le M)=1-\sum_{k=0}^{\infty}(1-p)^{k-1}p=\text{¿}(1-p)^{M}\text{¿}$$

Falta de memoria

$$P(X>M+N\vee X>M)=\frac{P(X>M+N)}{P(X>M)}=\frac{(1-p)^{M+N}}{(1-p)^{M}}=(1-p)^{N}$$

$$P(X>M+N\vee X>M)=P(X>N)$$

Esta propiedad se denomina de “falta de memoria”, porque la condición (un pasado) no modifica que excedió en a .

Variable aleatoria Poisson

Asociada a un especial proceso denominado de Poisson. Sirve para contar número de ocurrencias de un hecho en un tiempo dado.

- Número de llamadas telefónicas que llegan a un sistema.
- Número de accesos a un servidor.
- Número de fotones que emite un láser.
- Número de decaimientos (partículas emitidas) por una sustancia radioactiva.

Variable aleatoria Poisson

Asociada a un especial proceso denominado de Poisson. Sirve para contar número de ocurrencias de un hecho en un tiempo dado.

$$X \quad Poisson(\lambda)$$

Recorrido:

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} \; \forall k \in R_X$$

Variable aleatoria Poisson

$$p_X(k) = \frac{\lambda^k}{k!} e^{-\lambda} \; \forall k \in R_X$$

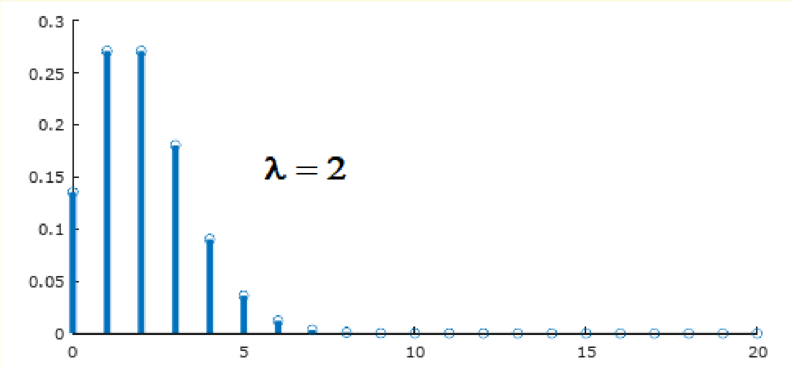
$$\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{+\lambda} = 1$$

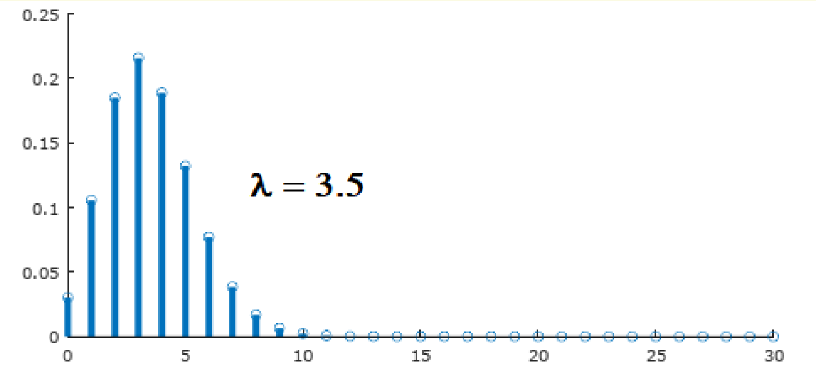
$$\mu = E[X] = \sum_{k=0}^{\infty} k \times \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \times e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$$

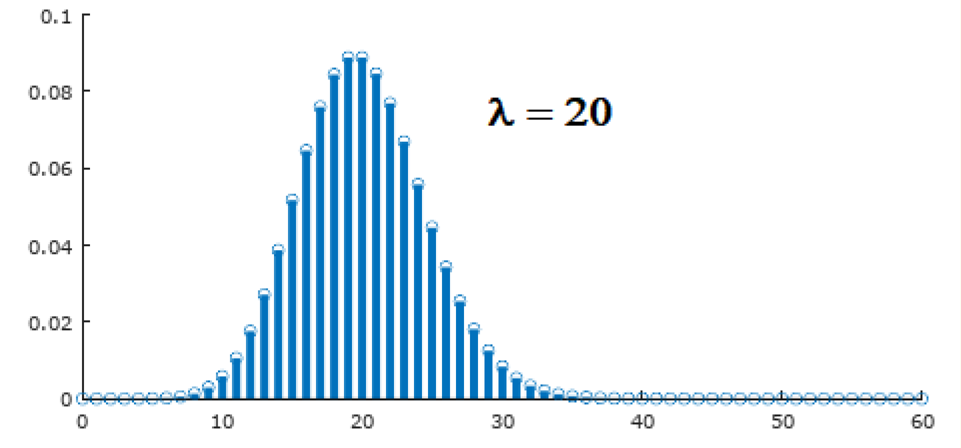
$$\sigma^2 = Var[X] = \lambda$$

¡Tarea para la casa!

Variable aleatoria Poisson







Suponga que hay 300 errores distribuidas al azar a lo largo de un libro de 600 páginas.El número de errores por página se puede asumir una variable aleatoria con distribución de Poisson de parámetro $\lambda = \frac{300}{600} = 0.5$. Calcule la probabilidad de que una página dada contenga;

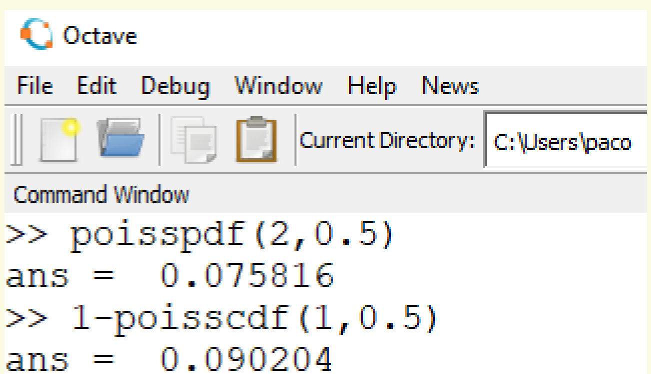
a) 2 errores exactamente

b) 2 ó más errores.

: número de errores en una página

$$P(X=k)=\frac{0.5^2}{2!}e^{-0.5}\approx 0.076$$

$$P(X\geq 2)=1-\frac{0.5^0}{0!}e^{-0.5}-\frac{0.5^1}{1!}e^{-0.5}\approx 0.0902$$



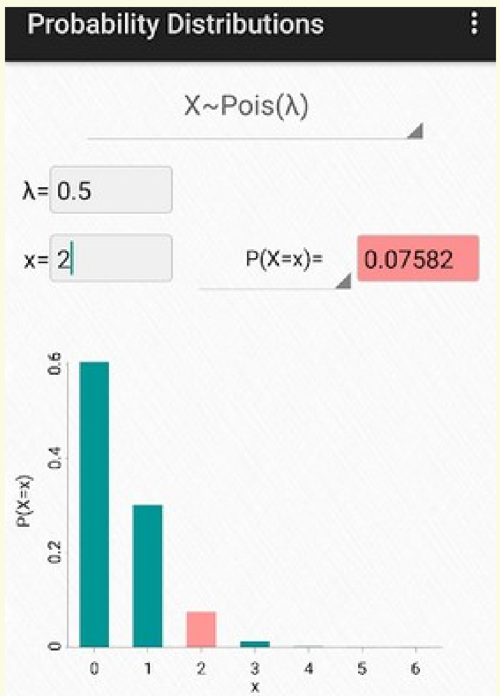
En Octave **poisspdf** calcula la función de probabilidad de la variable aleatoria de Poisson mientras que **poisscdf** calcula la función de distribución.

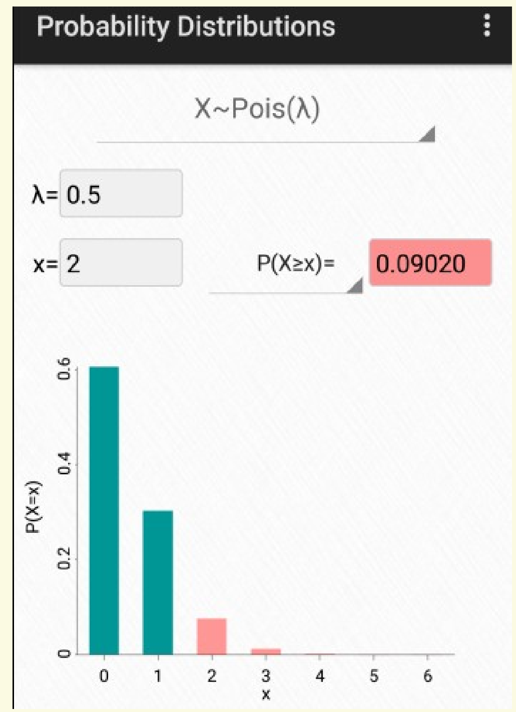
Suponga que hay 300 errores distribuidas al azar a lo largo de un libro de 600 páginas.El número de errores por página se puede asumir una variable aleatoria con distribución de Poisson de parámetro $\lambda = \frac{300}{600} = 0.5$. Calcule la probabilidad de que una página dada contenga;

a) 2 errores exactamente

b) 2 ó más errores.

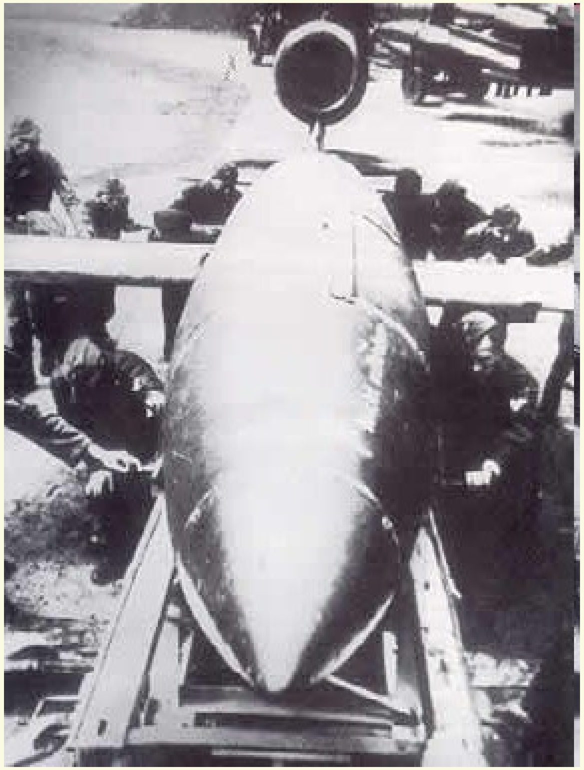
: número de errores en una página





Variable aleatoria Poisson

Bombas V1 y V2 sobre Londres en la Segunda Guerra Mundial



Durante la Segunda Guerra Mundial los alemanes lanzaron sobre Londres las temidas bombas voladoras V1 y V2. De alrededor de 5 m de longitud, cargando casi 1000 Kg de explosivos se lanzaban desde Calais en Francia (cerca de Londres) y al acabarse el combustible caían sobre Londres *aproximadamente al azar*.

Un oficial de la RAF relevó una región de 144 Km2 de Londres dividiéndola en 576 zonas de 0.25 Km2 de área. Durante el conflicto cayeron 537 bombas en esa región.

Variable aleatoria Poisson

Bombas V1 y V2 sobre Londres en la Segunda Guerra Mundial

Si la distribución de bombas se puede suponer al azar con ese promedio de bombas por zona entonces la fracción de zonas con 0, 1, 2, 3, … bombas se podría modelar con la distribución de Poisson.

$$P(X=k)=\frac{b^k}{k!}e^{-b}$$

donde es la cantidad de bombas y el promedio de bombas por zona y puede tomar los valores 0, 1, 2, …

La distribución de probabilidades de esta variable aleatoria depende de un parámetro. Como se dispone de datos entonces ese parámetro puede ***estimarse*** (dar una aproximación usando la media de los datos para dar esa estimación). Recordemos que, en este caso, el parámetro b es el valor esperado de la variable aleatoria.

Variable aleatoria Poisson

Bombas V1 y V2 sobre Londres en la Segunda Guerra Mundial

Bombas en una zona	Nro de zonas	Predicho con el modelo
0	229	226.7
1	211	211.4
2	93	98.5
3	35	30.6
4	7	7.1
5 o mas	1	1.8
	576	

Hubo 229 zonas que no fueron impactadas por bombas, 211 con exactamente 1 y así sucesivamente. Sólo 1 zona recibió 5 bombas. La suma de los productos entre las dos primeras columnas da el número total de bombas caídas. La última columna se calcula multiplicando el número total de zonas por la probabilidad calculada con el modelo de Poisson tomando como parámetro el promedio observado de bombas por zona .

Variable aleatoria Poisson

Relación con la binomial :

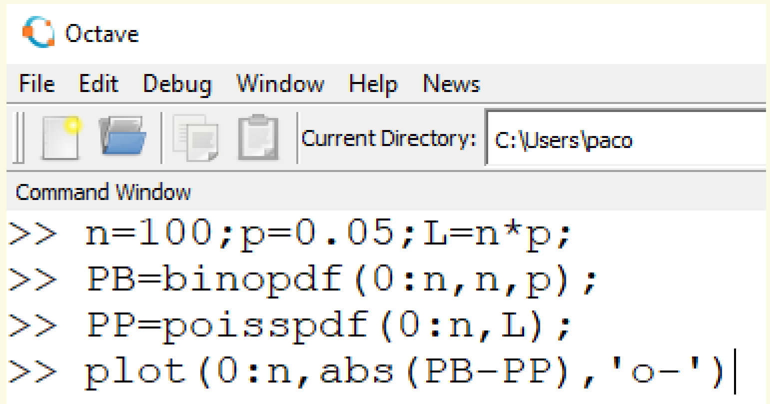
Si es grande () y es pequeño (), se puede aproximar la distribución de por la de .

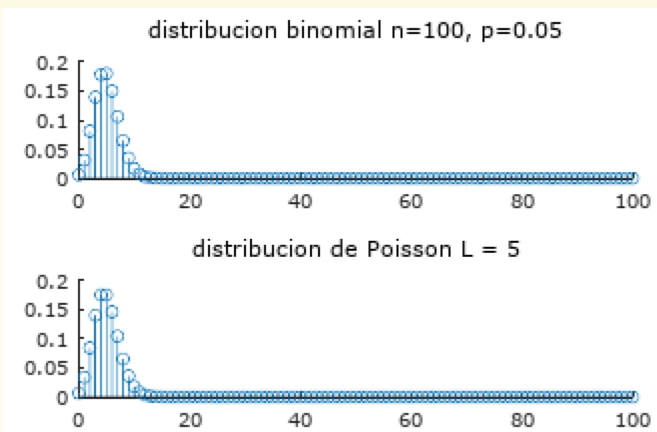
$$E[\boldsymbol{X}] = \boldsymbol{np} = E[\boldsymbol{Y}]$$

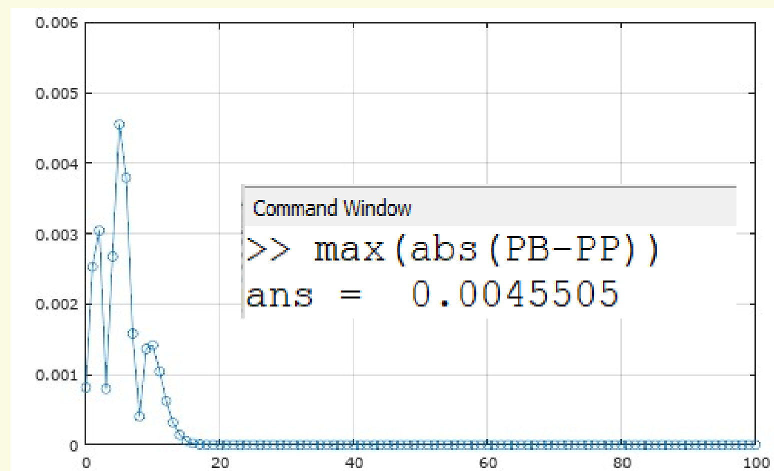
$$Var[X] = np(1-p) \approx np = Var[Y]$$

$$P(X \leq k) \approx P(Y \leq k)$$

Variable aleatoria Poisson







Una compañía de seguros garantiza pólizas de seguros individuales contra retrasos aéreos de más de doce horas. Una encuesta ha permitido estimar a lo largo de un año que cada persona tiene una probabilidad de cada de mil de ser víctima de un retraso aéreo que esté cubierto por este tipo de póliza y que la compañía aseguradora podrá vender una media de cuatro mil pólizas al año.

Se pide calcular la probabilidad de que el número de retrasos cubiertos por la póliza no pase de cuatro por año.

: número de retrasos cubiertos por la póliza en un año

$$E\left[X\right] = 4000 \times 0.001 = 4$$

Usando la aproximación de Poisson resulta:

$$P\left(X \leq 4\right) = \sum_{k=0}^{4} \frac{4^k}{k!} e^{-4} \approx 0.62884$$

Usando la distribución binomial con y resulta también 0.62884…