Clase 12: El proceso de Poisson

Matías Carrasco

26 de agosto de 2019

Resumen El proceso de Poisson es un modelo probabilístico similar al proceso de Bernoulli pero en tiempo continuo. De interés son la cantidad de éxitos hasta tiempo t (la distribución de Poisson) y el tiempo de espera entre dos éxitos consecutivos (la distribución exponencial).

Índice

El proceso de Poisson

En muchos fenómenos de interés, un éxito puede ocurrir en cualquier instante de tiempo. Por ejemplo, podríamos contar la cantidad de llamadas de teléfonos celulares que pasan a través de una torre de retransmisión entre las 9 y las 10 de la mañana, la cantidad de fallas en 100 metros de cable, la cantidad de clientes que llegan a la ventanilla de boletos entre las 12 del mediodía y las 2 de la tarde, o la cantidad de defectos en un rollo de pantalla de aluminio de 50 metros que tiene 20 cm de ancho. Para modelar este tipo de procesos precisamos una versión continua del proceso de Bernoulli. Este es el proceso de Poisson.

El proceso de Poisson

El número de éxitos en un intervalo continuo dado se ajusta a un proceso de Poisson de parámetro $\mu > 0$ si se cumplen las siguientes condiciones:

1. Las cantidades de éxitos en subintervalos disjuntos son independientes.
2. La probabilidad de un éxito en un subintervalo suficientemente chico de longitud h es aproximadamente μh.
3. La probabilidad de dos o más éxitos en un subintervalo suficientemente chico es esencialmente cero.

Supongamos que un experimento satisface las tres condiciones anteriores de un proceso de Poisson. Sea S el número de ocurrencias en un intervalo de longitud 1 (donde "longitud 1" representa una unidad de la cantidad bajo consideración). Nos gustaría encontrar una fórmula para $P(S = k)$, donde k es un entero no negativo.

Para esto, dividimos el intervalo en n subintervalos de igual longitud $1/n$. Si n es suficientemente grande (es decir, mucho más grande que k), aproximaremos la probabilidad de que haya k éxitos en el intervalo unidad por la probabilidad de que exactamente k de estos n subintervalos tengan un éxito. Esto está justificado porque la probabilidad de dos o más éxitos en cualquier subintervalo es esencialmente cero, según la condición 3. La probabilidad de un éxito en cualquier subintervalo de longitud $1/n$ es aproximadamente μ/n, según la condición 2.

Consideremos la ocurrencia o no de éxito en cada subintervalo como un ensayo de Bernoulli. Por la condición 1, tenemos una secuencia de n ensayos de Bernoulli con una probabilidad p aproximadamente igual a μ/n. Por lo tanto, una aproximación para $P(S = k)$ viene dada por la probabilidad binomial

$$\frac{n!}{k!(n-k)!}\left(\frac{\mu}{n}\right)^k\left(1-\frac{\mu}{n}\right)^{n-k}$$

Si n tiende a infinito, entonces

$$\begin{aligned}&\lim_{n\to\infty}\frac{n!}{k!(n-k)!}\left(\frac{\mu}{n}\right)^k\left(1-\frac{\mu}{n}\right)^{n-k}\\ &=\lim_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{n^k}\frac{\mu^k}{k!}\left(1-\frac{\mu}{n}\right)^n\left(1-\frac{\mu}{n}\right)^{-k}\end{aligned}$$

Ahora, para k fijo, tenemos

$$\begin{aligned}\lim_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{n^k}&=\lim_{n\to\infty}(1)\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)=1\\ \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n&=e^{-\mu}\\ \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^{-k}&=1\end{aligned}$$

Luego

$$\lim_{n\to\infty}\frac{n!}{k!(n-k)!}\left(\frac{\mu}{n}\right)^k\left(1-\frac{\mu}{n}\right)^{n-k}=\frac{\mu^k e^{-\mu}}{k!}=P(S=k)$$

La distribución de Poisson

Una variable discreta S tiene distribución de Poisson de parámetro μ si toma valores enteros mayores o iguales a cero con probabilidad

$$P(S=k)=\frac{\mu^k}{k!}e^{-\mu} \quad (k=0,1,2,\ldots)$$

Esto lo escribimos $S \sim \mathrm{Pois}(\mu)$.

La distribución de Poisson es sin duda la distribución discreta más importante. Como veremos en seguida, una gran variedad de observaciones se ajustan a ella. Juega un papel similar, para las variables discretas, al papel que juega la campana de Gauss para las mediciones continuas.

Es fácil ver que la distribución está bien definida, pues la suma de estas probabilidades es igual a 1. Basta recordar la serie de Taylor de la función exponencial:

$$e^{\mu}=\sum_{k=0}^{\infty}\frac{\mu^k}{k!},$$

de donde se deduce inmediatamente que

$$\sum_{k=0}^{\infty}P(S=k)=e^{-\mu}\sum_{k=0}^{\infty}\frac{\mu^k}{k!}=e^{-\mu}e^{\mu}=1.$$

Como veremos más adelante, el parámetro μ indica el número promedio de observaciones, o de forma equivalente $E(X)=\mu$.

Ejemplo 1 En una ciudad grande, las llamadas telefónicas al 911 llegan en un promedio de dos cada 3 minutos. Si se supone un proceso de Poisson para la llegada de las llamadas, ¿cuál es la probabilidad de que lleguen cinco o más llamadas en un período de 9 minutos?

Sea S el número de llamadas en un período de 9 minutos. Vemos que $E(S)=6$; es decir, en promedio, se recibirán seis llamadas durante un período de 9 minutos. Así,

$$P(S\geq 5)=1-P(S\leq 4)=1-\sum_{k=0}^{4}\frac{6^k e^{-6}}{k!}=1-0.285=0.715$$

■

La distribución de Poisson no solo es importante por derecho propio, sino que también se puede usar para aproximar las probabilidades de una distribución binomial. Anteriormente vimos que si S tiene una distribución de Poisson con el parámetro μ, entonces con n grande

$$P(S=k)\approx\binom{n}{k}\left(\frac{\mu}{n}\right)^k\left(1-\frac{\mu}{n}\right)^{n-k}$$

donde $p = \mu/n$, de modo que $\mu = np$ en la probabilidad binomial anterior. Es decir, si S tiene distribución binomial $\text{Bin}(n, p)$ con n grande y p pequeño, entonces

$$\frac{(np)^k e^{-np}}{k!} \approx \binom{n}{k} p^k (1-p)^{n-k}.$$

Esta aproximación es razonablemente buena si n es grande. Pero como μ es constante fijo en el argumento anterior, p debería ser pequeño, porque $np = \mu$. En particular, la aproximación es bastante precisa si $n \geq 20$ y $p \leq 0.05$ o si $n \geq 100$ y $p \leq 0.10$, pero no es mala en otras situaciones que violan estos límites, como $n = 50$ y $p = 0.12$.

Ejemplo 2 Un fabricante de luces de árboles de Navidad sabe que el 2 % de sus luces son defectuosas. Suponiendo independencia, el número de luces defectuosas en una caja de 100 luces tiene distribución binomial de parámetros $n = 100$ y $p = 0.02$. Para aproximar la probabilidad de que una caja de 100 de estas luces contenga como máximo tres luces defectuosas, utilizamos la distribución de Poisson con $\mu = 100(0.02) = 2$, lo que da

$$\sum_{k=0}^{3} \frac{2^k e^{-2}}{k!} = 0.857$$

Usando la distribución binomial, obtenemos, después de algunos cálculos tediosos,

$$\sum_{k=0}^{3} \binom{100}{k} (0.02)^k (0.98)^{100-k} = 0.859$$

Por lo tanto, en este caso, la aproximación de Poisson es extremadamente cercana al valor verdadero, pero mucho más fácil de calcular. ■

Varios ejemplos de variables Poisson

Veamos varios ejemplos de observaciones que se ajustan a la distribución de Poisson. En algunos de ellos, más precisamente en los ejemplos de la cerveza, las chips de chocolate y las lentejas, se trata de hecho de un proceso de Poisson espacial y no temporal, pero la deducción de la fórmula es exactamente la misma que para el caso temporal. La idea detrás sigue siendo la aproximación de Poisson a la binomial.

Poisson en la cerveza

En la producción de cerveza es importante conocer con precisión la cantidad de levadura utilizada para la fermentación. Demasiada levadura produce una cerveza amarga, y poca hace que la cerveza no fermente bien.

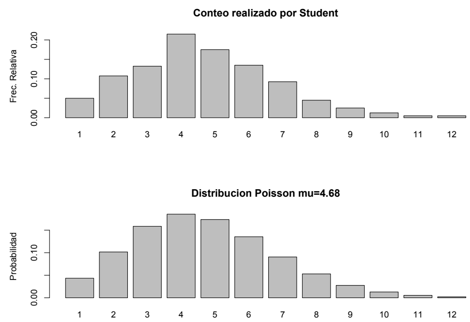


Figura 1: Comparación de la distribución de Poisson con $\mu = 4.68$ y el conteo experimental realizado por Student.

En 1904, la Guiness contrató al matemático Student[1] (William Gosset) para mejorar la calidad de la cerveza. Student ideó un modelo probabilístico para controlar la cantidad de levadura usada en la fermentación.

El modelo de Student era el siguiente: supongamos que un líquido (un cultivo diluido de células de levadura) es vertido y extendido sobre una placa, formando una capa delgada de 0.01 mm de espesor. Queremos estimar la cantidad promedio μ de células de levadura por unidad de área en la placa. La unidad de área es $1/400\ mm^2$ y la placa tiene en total N unidades de área. Debemos imaginarnos la placa dividida en N cuadrados de área $1/400\ mm^2$.

Asumiendo que el líquido ha sido bien mezclado, una célula dada tendrá la misma probabilidad de caer en cualquier unidad de área. En total hay μN células.

Student observó en el microscopio un área cuadrada de 1 mm^2 de la placa, que estaba dividida en 400 celdas cuadradas (esto se hace con un aparato llamado hemocitómetro), y contó el número de células en cada una de ellas. El resultado está en la Tabla 1.

Los valores de la tabla representan el número de células en cada una de las 400 celdas. Para comparar estos con la distribución de Poisson es mejor representar los valores de la tabla graficamente.

Lo que hacemos es observar cuántas celdas tienen, por ejemplo, 4 células. Al dividir este número por el total de celdas, es decir 400, tendremos la frecuencia relativa del valor 4, lo cual nos da una idea de la probabilidad de que caigan 4 células en una celda.

La Figura 1 muestra las frecuencias relativas del experimento realizado por Student. En la parte de abajo se compara con la función de probabilidad puntual de la distribución de Poisson de parámetro $\mu = 4.68$. Más adelante veremos métodos para cuantificar cuán buena es la aproximación de Poisson, respecto a la evidencia empírica aportada por los datos.

El parámetro μ debe interpretarse como la cantidad media de célu-

[1] William Gosset usaba el seudónimo Student para publicar sus investigaciones debido a que la compañía Guiness no autorizaba a sus empleados a revelar información importante sobre el proceso de producción de su cerveza.

2	2	4	4	4	5	2	4	7	7	4	7	5	2	8	6	7	4	3	4
3	3	2	4	2	5	4	2	8	6	3	6	6	10	8	3	5	6	4	4
7	9	5	2	7	4	4	2	4	4	4	3	5	6	5	4	1	4	2	6
4	1	4	7	3	2	3	5	8	2	9	5	3	9	5	5	2	4	3	4
4	1	5	9	3	4	4	6	6	5	4	6	5	5	4	3	5	9	6	4
4	4	5	10	4	4	3	8	3	2	1	4	1	5	6	4	2	3	3	3
3	7	4	5	1	8	5	7	9	5	8	9	5	6	6	4	3	7	4	4
7	5	6	3	6	7	4	5	8	6	3	3	4	3	7	4	4	4	5	3
8	10	6	3	3	6	5	2	5	3	11	3	7	4	7	3	5	5	3	4
1	3	7	2	5	5	5	3	3	4	6	5	6	1	6	4	4	4	6	4
4	2	5	4	8	6	3	4	6	5	2	6	6	1	2	2	2	5	2	2
5	9	3	5	6	4	6	5	7	1	3	6	5	4	2	8	9	5	4	3
2	2	11	4	6	6	4	6	2	5	3	5	7	2	6	5	5	1	2	7
5	12	5	8	2	4	2	1	6	4	5	1	2	9	1	3	4	7	3	6
5	6	5	4	4	5	2	7	6	2	7	3	5	4	4	5	4	7	5	4
8	4	6	6	5	3	3	5	7	4	5	5	5	6	10	2	3	8	3	5
6	6	4	2	6	6	7	5	4	5	8	6	7	6	4	2	6	1	1	4
7	2	5	7	4	6	4	5	1	5	10	8	7	5	4	6	4	4	7	5
4	3	1	6	2	5	3	3	3	7	4	3	7	8	4	7	3	1	4	4
7	6	7	2	4	5	1	3	12	4	2	2	8	7	6	7	6	3	5	4

Tabla 1: Distribución de células de levadura sobre 1 mm^2 dividido en 400 cuadrados.

las por unidad de área.

Poisson en los mundiales

La Tabla 2 muestra la frecuencia de partidos según la cantidad de goles en el mundial de fútbol de Francia 98. En total se jugaron 64 partidos y se hicieron 170 goles. Por ejemplo, hubo 5 partidos en los que no hubo goles, y hubo un solo partido en los que se convirtieron 7 goles. El número promedio de goles por partido es

$$\mu = \frac{170}{64} \approx 2.66.$$

Número total de goles por partido	0	1	2	3	4	5	6	7	Total
Frecuencia de partidos con esa cantidad de goles	5	11	12	18	11	6	0	1	64

Tabla 2: Goles en el mundial de Francia 98

En la Figura 2 se muestra la comparación con la distribución de Poisson de parámetro $\mu = 2.66$. Arriba de cada barra se muestra en rojo tanto la frecuencia relativa como la probabilidad teórica. La frecuencia relativa se obtiene dividiendo la cantidad de partidos en los que hubo k goles sobre 64 que es el total de partidos.

Notar que las diferencias entre las frecuencias observadas y las probabilidades teóricas están dentro de los rangos esperados para la cantidad de partidos estudiados. Dicho mal y pronto, 64 no es un número

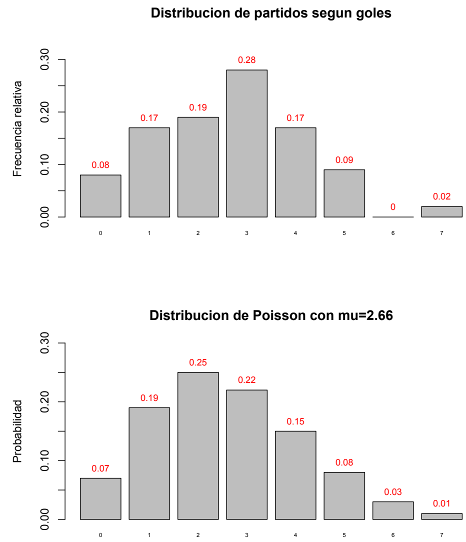


Figura 2: Arriba se muestra la distribución de partidos según goles en el mundial de Francia 98. Abajo se muestran las probabilidades teóricas de la distribución de Poisson de parámetro $\mu = 2.66$.

grande de datos para sacar conclusiones determinantes. Damos aquí simplemente una primera impresión, se podría estudiar si la distribución de partidos por goles se ajusta mejor en campeonatos más largos con más partidos. Es un trabajo que puede dar sus frutos, ¡nos puede ayudar a ganar algunos pesos en la próxima penca del mundial!

¡Poisson hasta en las galletitas!

En el curso de PyE del segundo semestre del 2016 se hizo el siguiente experimento: ¿Cuántas chips de chocolate hay en las galletitas de la marca *Pepitos*?

Para esto se repartieron varios paquetes de Pepitos entre los estudiantes. En lugar de comerlas, los estudiantes contaron cuántas chips de chocolate hay en cada galleta. Los datos obtenidos se muestran en la Tabla 3. Esta tabla se debe leer igual que la anterior, por ejemplo, hubo 6 galletas con 4 chips de chocolate.

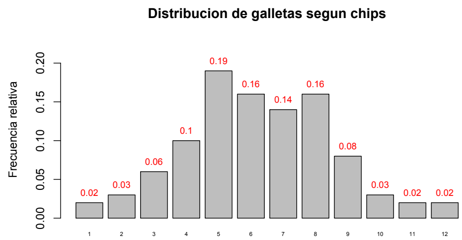
Número de chips de chocolate por galleta	1	2	3	4	5	6	7	8	9	10	11	12	Total
Frecuencia de galletas con esa cantidad de chips	1	2	4	6	12	10	9	10	5	2	1	1	63

Tabla 3: Chips de chocolate en las galletas Pepitos.

El número promedio de chips por galleta es

$$\mu = \frac{392}{63} \approx 6.2.$$

En la Figura 3 se muestra la comparación con la distribución de Poisson. De nuevo, la cantidad de datos no permite sacar conclusiones determinantes, pero se puede apreciar que la aproximación es bastante buena.



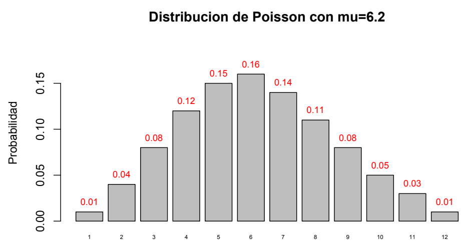


Figura 3: Arriba se muestra la distribución de galletas según la cantidad de chips de chocolate. Abajo se muestran las probabilidades teóricas de la distribución de Poisson de parámetro $\mu = 6.2$.

Poisson en las lentejas

En el primer semestre del 2017 fuimos por las lentejas. El experimento consistió en lanzar de forma aleatoria un puñado de lentejas sobre un papel cuadriculado. El papel contenía 100 cuadrados de 1 cm de lado.

La cantidad de lentejas lanzadas no coincide con las que al final quedan sobre el papel, así que el número promedio de lentejas por cuadrado depende de la realización del experimento.

En la Tabla 4 se muestra los datos obtenidos por un grupo de estudiantes.

Número de lentejas por cuadrado	0	1	2	3	4	5	Total
Frecuencia de cuadrados con esa cantidad de lentejas	52	32	13	2	0	1	100

Tabla 4: Lentejas distribuidas al azar sobre un papel cuadriculado.

El número medio de lentejas por cuadrado es entonces

$$\mu = \frac{69}{100} = 0.69.$$

En la Figura 4 se muestra la comparación con la distribución de Poisson de parámetro $\mu = 0.69$. Notar que la aproximación es muy buena a pesar de la poca cantidad de datos.

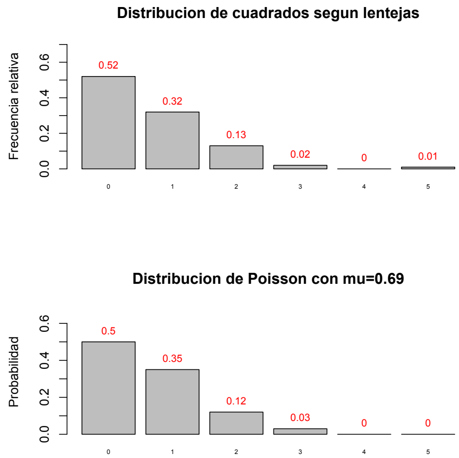


Figura 4: Arriba se muestra la distribución de cuadrados según la cantidad de lentejas. Abajo se muestran las probabilidades teóricas de la distribución de Poisson de parámetro $\mu = 0.69$.

Poisson en la bolsa de valores

Cuando una empresa cotiza en la bolsa de valores, el precio de sus acciones suben y bajan de forma poco predecible. Decimos que hay un *aumento* en el precio de las acciones de una empresa, si el precio de un determinado día es mayor al precio del día anterior.

Aunque parezca bastante sorprendente, detrás del aparente azar que esconden los cambios de precios encontramos alguna regularidades.

Por ejemplo, en la Tabla 5 se muestra el número de aumentos de las acciones de *Google* durante el período que va desde Enero del 2006 hasta Diciembre del 2010. Para ser más específicos, este período de tiempo se dividió en 60 meses. Para cada uno de estos meses, se contó cuántas veces hubo un aumento en el precio de las acciones en dicho més. La tabla muestra entonces la cantidad de meses en los que hubo k aumentos de precio.

Número de aumentos por mes	5	6	7	8	9	10	11	12	13	14	15	16	Total
Frecuencia de meses con esa cantidad de aumentos	1	6	2	9	9	10	5	8	5	4	0	1	60

Tabla 5: Distribución de meses según aumentos de precio en las acciones de Google.

El número medio de aumentos por mes es entonces

$$\mu = \frac{596}{60} \approx 9.93.$$

En la Figura 5 se muestra la comparación con la distribución de Poisson de parámetro $\mu = 9.93$. Notar que la aproximación es muy buena a pesar de la poca cantidad de datos.

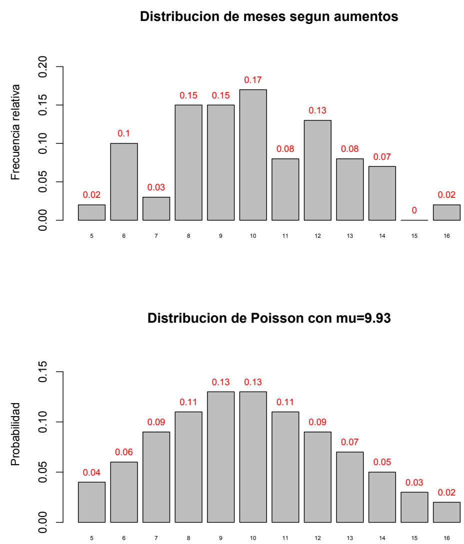


Figura 5: Arriba se muestra la distribución de meses según la cantidad de aumentos. Abajo se muestran las probabilidades teóricas de la distribución de Poisson de parámetro $\mu = 9.93$.

Como una primera aproximación no está mal, teniendo en cuenta que son solo 60 datos. ¿Seguirá siendo buena la aproximación si observamos la bolsa durante un período más largo de tiempo?

Esperanza y varianza de la distribución de Poisson

En la Figura 6 mostramos la función de probabilidad puntual de una variable S con distribución de Poisson para varios valores del parámetro μ. El valor de μ varía de 1 a 1 000. Notar cómo a medida que μ crece, la distribución de S se corre hacia la derecha y se concentra en valores cada vez mayores de k. Es interesante notar además la forma acampanada y simétrica de la distribución cuando μ es grande. Esto no es casualidad y lo estudiaremos más adelante.

Esperanza y varianza de la distribución de Poisson

Si $S \sim \text{Pois}(\mu)$, entonces

$$E(S) = \mu \quad \text{y} \quad Var(S) = \mu.$$

Demostración. De la definición de valor esperado:

$$E(S) = \sum_{k=0}^{\infty} k \frac{\mu^k e^{-\mu}}{k!} = \sum_{k=1}^{\infty} k \frac{\mu^k e^{-\mu}}{k!},$$

ya que el primer término ($k = 0$) de la serie se anula.

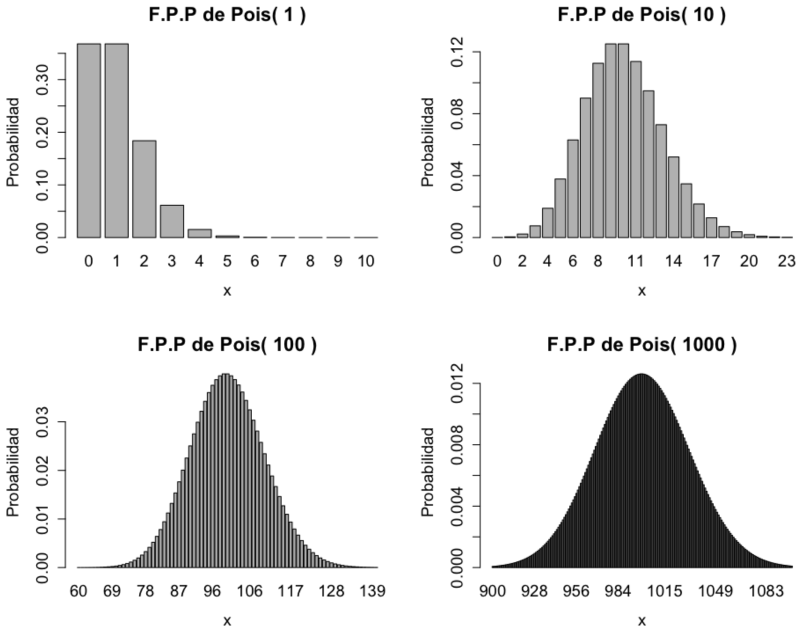


Figura 6: Distribución de Poisson para varios valores de μ. Para valores grandes de μ la forma del gráfico de la distribución se parece a la campana de Gauss.

Entonces

$$E(S) = \sum_{k=1}^{\infty} \frac{\mu^k e^{-\mu}}{(k-1)!} = \mu \sum_{k=1}^{\infty} \frac{\mu^{k-1} e^{-\mu}}{(k-1)!}.$$

Poniendo $j = k-1$ obtenemos

$$\mu \sum_{k=1}^{\infty} \frac{\mu^{k-1} e^{-\mu}}{(k-1)!} = \mu \sum_{j=0}^{\infty} \frac{\mu^{j} e^{-\mu}}{j!} = \mu.$$

De forma análoga podemos calcular la varianza. Para esto calculemos primero la esperanza de $S(S-1)$. De la definición tenemos que

$$\begin{aligned} E(S(S-1)) &= \sum_{k=0}^{\infty} k(k-1) \frac{\mu^k e^{-\mu}}{k!} \\ &= \mu^2 \sum_{k=2}^{\infty} \frac{\mu^{k-2} e^{-\mu}}{(k-2)!} = \mu^2. \end{aligned}$$

Pero por otro lado,

$$E(S(S-1)) = E\left(S^2\right) - E(S),$$

de donde $E(S^2) = \mu^2 + \mu$. De aquí deducimos que

$$Var(S) = E\left(S^2\right) - E(S)^2 = \mu^2 + \mu - \mu^2 = \mu.$$

Esto es lo que queríamos demostrar. $\square$

La distribución exponencial

La distribución exponencial es la versión continua de la distribución geométrica, y mide el tiempo entre éxitos consecutivos en el proceso de Poisson. En general se puede pensar a T como el tiempo que hay que esperar hasta que un éxito ocurra, siendo el tiempo en este caso continuo.

El mismo argumento que usamos para hallar la fórmula de la distribución de Poisson muestra que si en lugar de mirar un intervalo de longitud 1, miramos uno de longitud t, entonces la cantidad de éxitos S_t es una variable con distribución de Poisson de parámetro μt.

¿Cual es la densidad del tiempo de espera T hasta el primer éxito? Es más fácil calcular la fda de T primero y después derivarla. Por un lado

$$F(t) = P(T \leq t) = 1 - P(T > t).$$

Pero el evento $T > t$ quiere decir que no hubo éxitos en el intervalo $[0, t]$, o lo que es lo mismo que $S_t = 0$. Luego

$$F(t) = 1 - P(S_t = 0) = 1 - e^{-\mu t}.$$

Derivando obtenemos la densidad $p(t) = \mu e^{-\mu t}$ con $t > 0$. Por razones históricas se suele usar λ en lugar de μ para describir la distribución de una variable exponencial.

La distribución exponencial

Decimos que una variable aleatoria T tiene distribución exponencial de tasa (o parámetro) $\lambda > 0$, y lo notamos $T \sim \text{Exp}(\lambda)$, si T tiene densidad

$$p(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0, \\ 0 & \text{si } t < 0. \end{cases}$$

En la Figura 7 se muestra esta densidad para distintos valores de λ.

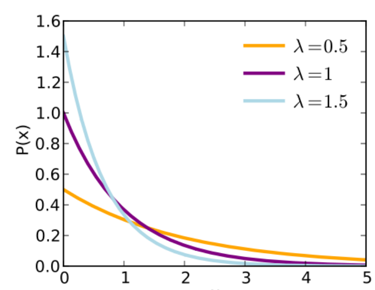


Figura 7: Densidad de una variable aleatoria con distribución Exponencial de parámetro λ.

De forma equivalente, para $0 \leq a \leq b$ tenemos

$$P(a < T \leq b) = \int_a^b \lambda e^{-\lambda t} dt = \left[e^{-\lambda t} \right]_a^b = e^{-\lambda a} - e^{-\lambda b}.$$

Para ver que la densidad exponencial está bien definida, basta tomar $a = 0$ y $b = +\infty$ en la ecuación anterior y observar que la integral es 1. Si tomamos $a = t > 0$ y hacemos $b = +\infty$, deducimos una expresión para la *función de supervivencia*

$$P(T > t) = e^{-\lambda t}.$$

Esta función decae exponencialmente a medida que t tiende a $+\infty$.

Una consecuencia importante de la expresión anterior es que la densidad exponencial presenta la propiedad de pérdida de memoria:

$$\begin{aligned} P(T > t + s | T > t) &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} = P(T > s). \end{aligned}$$

Si T representa el tiempo de vida útil de un objeto, la propiedad de pérdida de memoria dice que dado que ha sobrevivido hasta tiempo t, las chances de que sobreviva un tiempo adicional s son las mismas que las de sobrevivir un tiempo s contando desde el comienzo.

Algunos "objetos", como los átomos o los componentes eléctricos, tienen esta propiedad, y por lo tanto, su tiempo de vida se ajusta bien a una distribución exponencial. Pero la mayoría de las formas de vida biológicas no se ajustan a una distribución exponencial del tiempo de vida, porque experimentan un proceso de envejecimiento.

Para algo con una vida útil distribuida exponencialmente, λ es el valor constante de la tasa de mortalidad instantánea o de riesgo instantáneo. Es decir, λ mide la probabilidad de muerte por unidad de tiempo justo después del tiempo t, dada la supervivencia hasta el tiempo t. Para ver por qué, si consideramos un tiempo t y otro período de tiempo Δ, calculamos

$$\begin{aligned} P(T \leq t + \Delta | T > t) &= 1 - P(T > t + \Delta | T > t) \\ &= 1 - P(X > \Delta) \ \text{(pérdida de memoria)} \\ &= 1 - e^{-\lambda\Delta} \approx \lambda\Delta \end{aligned}$$

para valores pequeños de Δ.

Menos formalmente, para un incremento de tiempo infinitesimal dt, el resultado de este cálculo es que

$$P(T \leq t + dt | T > t) = \lambda dt,$$

o que

$$p(t) = \frac{P(t < T \leq t + dt)}{dt} = \lambda P(T > t).$$

Como el lado izquierdo es la densidad de T en el tiempo t, esto explica por qué la densidad exponencial en t es la tasa de riesgo λ multiplicada por la probabilidad $e^{-\lambda t}$ de supervivencia hasta el tiempo t: $p(t) = \lambda P(T > t)$. La característica de los tiempos de vida distribuidos exponencialmente es que la tasa de riesgo es constante, no dependiente de t. Otras distribuciones continuas en $[0, +\infty)$ corresponden a tasas de riesgo dependientes del tiempo $\lambda(t)$.

Esperanza y varianza de la distribución exponencial

Si $T \sim \mathrm{Exp}(\lambda)$, entonces

$$E(T) = \frac{1}{\lambda} \quad \text{y} \quad Var(T) = \frac{1}{\lambda^2}.$$

Demostración. La esperanza es

$$E(T) = \int_0^\infty (1 - F(t))dt = \int_0^\infty P(T > t)\, dt = \int_0^\infty e^{-\lambda t} dt = \frac{1}{\lambda}.$$

Notar la analogía con la distribución geométrica de parámetro p, para la cual la esperanza es $1/p$.

Para la varianza usaremos la fórmula $Var(T) = E\left(T^2\right) - E(T)^2$. El primer término lo obtenemos integrando por partes dos veces:

$$\begin{aligned} E\left(T^2\right) &= \int_0^\infty t^2 \lambda e^{-\lambda t} dt \\ &= \frac{1}{\lambda^2} \int_0^\infty u^2 e^{-u} du \\ &= \frac{1}{\lambda^2} \left[-e^{-u}(u^2 + 2u + 2) \right]_0^\infty = \frac{2}{\lambda^2}. \end{aligned}$$

Por tanto $Var(T) = 1/\lambda^2$. □

Ejemplo 3 Bajo condiciones de uso constantes, algunos tipos de componentes eléctricos, por ejemplo, fusibles y transistores, tienen una distribución de tiempo de vida que se ajusta bien por una distribución exponencial.

Tal componente no se desgasta gradualmente. Por el contrario, deja de funcionar de forma repentina e impredecible. No importa cuánto tiempo haya estado en uso el componente, la probabilidad de que sobreviva un intervalo de tiempo adicional de longitud Δ es siempre la misma. Esta probabilidad debe ser $e^{-\lambda\Delta}$ para alguna tasa λ, llamada tasa de falla en este contexto. La distribución de vida es entonces exponencial con la tasa λ. En general, mientras siga funcionando, tal componente es tan bueno como uno nuevo.

Supongamos que la tasa es $\lambda = 0.01$ por hora. Estimemos la probabilidad de que el transistor funcione por 50 horas. Basta calcular

$$P(T > 50) = e^{-\lambda 50} = e^{-0.5} \approx 0.606.$$

Dado que el transistor ha estado funcionando por 50 horas, ¿cuál es la probabilidad de que falle en el próximo minuto de uso?

De la interpretación de $\lambda = 0.01$ como la tasa de falla instantánea por hora, dado que se ha sobrevivido 50 horas, la probabilidad es aproximadamente $0.01 \times \frac{1}{60} \approx 0.00017$. ■

Ejemplo 4 Los átomos de isótopos radiactivos como carbono 14, uranio 235 o estroncio 90 permanecen intactos hasta un instante aleatorio cuando se descomponen repentinamente, lo que significa que se dividen o se convierten en otro tipo de átomo y emiten un pulso de radiación o partículas de algún tipo.

Este decaimiento radioactivo puede detectarse mediante un contador Geiger. Sea T el tiempo de vida, o el tiempo hasta la descomposición, de tal átomo, comenzando en algún momento arbitrario cuando el átomo está intacto. Es razonable suponer que la distribución de T debe tener la propiedad de pérdida de memoria. En consecuencia, hay una tasa λ, la tasa de desintegración para el isótopo en cuestión, tal que T tiene una distribución exponencial de parámetro λ.

Las probabilidades aquí tienen una interpretación clara debido a la gran cantidad de átomos típicamente involucrados (por ejemplo, unos pocos gramos de una sustancia contien del orden de 10^{24} átomos). Supongamos que un gran número N de tales átomos se descomponen independientemente el uno del otro. Entonces la proporción de estos N átomos que sobrevive hasta el tiempo t está cerca de $e^{-\lambda t}$, la probabilidad de supervivencia para cada átomo individual.

Este decaimiento exponencial de la masa de la sustancia radiactiva se ha verificado experimentalmente, lo que confirma la hipótesis de que las vidas de los átomos individuales se distribuyen exponencialmente. Las tasas de decaimiento λ para isótopos individuales se pueden medir con gran precisión, utilizando este decaimiento exponencial de la masa. Estas tasas no muestran una dependencia aparente de las condiciones físicas, como la temperatura y la presión.

Una forma común de indicar la tasa de decaimiento de un isótopo radiactivo es la vida media τ. Este es el tiempo que demora la desintegración de la mitad de la cantidad inicial del isótopo. Por lo que

$$e^{-\lambda\tau} = \frac{1}{2} \text{ o } \tau = \frac{\ln(2)}{\lambda}.$$

En otras palabras, la vida media τ es la mediana de la distribución del tiempo de vida del átomo

$$P(T \leq \tau) = P(T > \tau) = \frac{1}{2}.$$

El estroncio es un componente particularmente peligroso de los residuos de las explosiones nucleares. La sustancia es tóxica, se absorbe fácilmente en los huesos cuando se come, y tiene una larga vida media de aproximadamente 28 años. Suponiendo este valor para la vida media, calculemos:

1. La tasa de decaimiento λ: por lo anterior, esta es

$$\lambda = \frac{\ln(2)}{\tau} = 0.0248 \text{ por año}.$$

2. La probabilidad de que un átomo de estroncio 90 sobreviva al menos 50 años es

$$P(T > 50) = e^{-\lambda 50} = e^{-0.0248 \times 50} = 0.29.$$

3. La proporción de un gramo de estroncio 90 que queda después de 50 años es 0.29, pues esta es la probabilidad que calculamos antes.

4. El número de años que deben transcurrir luego de una explosión nuclear para que el 99 % del estroncio producido haya decaído es

$$e^{-0.0248t} = \frac{1}{100} \text{ de donde} t = \frac{\ln(100)}{0.0248} \approx 186 \text{ años.}$$

■

Ejemplo 5 Vamos a hacer un modelo de la atmósfera basado en la distribución exponencial. Supondremos que la atmósfera es un gas ideal y que se encuentra en equilibrio térmico a temperatura constante T.

Denotemos por $n(h)$ la densidad de moléculas de gas a la altura h:

$$n(h) = \frac{\text{\# de moléculas a la altura } h}{\text{unidad de volumen}}.$$

De la ecuación de estado de los gases ideales[2] sabemos que

$$P(h) = n(h)kT.$$

[2] Usualmente se la escribe $PV = nRT$ en donde n es el número de moles. Como estamos usando número de partículas hemos escrito la constante como k.

Además, para que el aire esté en equilibrio, debemos tener

$$P(h + dh) + mgn(h)dh = P(h),$$

en donde m es la masa de una molécula de aire y g la aceleración de la gravedad (que suponemos constante). Escribiendo $dP = P(h + dh) - P(h)$, podemos reescribir la relación anterior como

$$P'(h) = \frac{dP}{dh} = -mgn(h).$$

Derivando la ecuación de los gases $P'(h) = kTn'(h)$, de donde obtenemos la ecuación

$$n'(h) = -\frac{mg}{kT}n(h),$$

cuya solución es $n(h) = n(0)e^{-mgh/kT}$.

Observar que mgh es la energía potencial E_h de una molécula de aire a la altura h. El cálculo que hemos hecho nos dice que la probabilidad de encontrar una molécula de aire con energía (cercana a) E_h es proporcional a $e^{-E_h/kT}$.

Es decir, la distribución de energías de las moléculas de aire es exponencial de parámetro $1/kT$. En física esta distribución se llama *distribución de Boltzmann*. ■