3.5 – Variables aleatorias continuas

En la sección anterior se consideraron variables aleatorias discretas, o sea variables aleatorias cuyo rango es un conjunto finito o infinito numerable. Pero hay variables aleatorias cuyo rango son todos los números reales de un intervalo dado, (es decir es un conjunto infinito no numerable). Ejemplos de variables continuas podrían ser

X: “tiempo que tarda en llegar un colectivo a una parada”
Y: “tiempo de vida de un fusible”
Como ahora los valores de una v.a. continua no son contables no se puede hablar del *i*-ésimo valor de la v.a. *X* y por lo tanto $p(x_i) = P(X = x_i)$ pierde su significado. Lo que se hace es sustituir la función $p(x)$ definida sólo para $x_1, x_2, ...$, por una función $f(x)$ definida para todos los valores x del rango de *X*. Por lo tanto se da la siguiente definición de v.a. continua

Sea *X* una v.a.. Decimos que es ***continua*** si existe una función no negativa f, definida sobre todos los reales $x \in (-\infty, \infty)$, tal que para cualquier conjunto *B* de números reales

$$P(X \in B) = \int_B f(x)dx$$

O sea que la probabilidad de que *X* tome valores en *B* se obtiene al integrar la función *f* sobre el conjunto *B*.

A la función f la llamamos función densidad de probabilidad (f.d.p.).

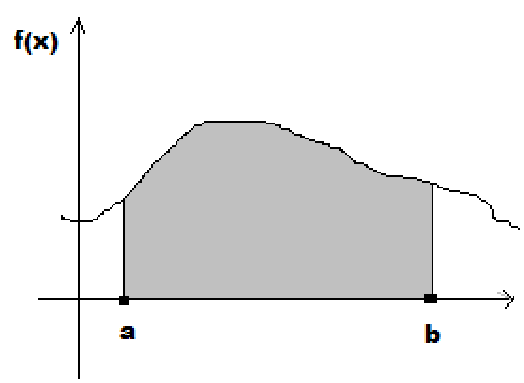
Observaciones:

1- Como *X* debe tomar algún valor real, entonces debe cumplirse que

$$1 = P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x)dx$$

2- Si *B* es el intervalo real $[a,b] = \{x \in R;\ a \le x \le b\}$ entonces

$$P(X \in B) = P(a \le X \le b) = \int_a^b f(x)dx$$



Notar que en este caso la probabilidad de que *X* tome valores en el intervalo $[a,b]$ es ***el área bajo f entre a y b***

3- Si en la observación anterior $a = b$ entonces

$$P(X = a) = \int_a^a f(x)dx = 0$$

Es decir la probabilidad que una v.a. continua tome algún valor fijado es cero. Por lo tanto, para una v.a. continua

$$P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b) = \int_a^b f(x)dx$$

Función de distribución acumulada

Sea X una v.a. continua. Se define la ***función de distribución acumulada de X*** (abreviamos F.d.a de *X*) como

$$F(x) = P(X \le x) \qquad -\infty < x < \infty$$

Si X tiene f.d.p. f(x) entonces

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt \qquad -\infty < x < \infty$$

Además

$$P(a \le X \le b) = \int_a^b f(x)dx = \int_{-\infty}^{b} f(x)dx - \int_{-\infty}^{a} f(x)dx = F(b) - F(a)$$

Observaciones:

1- Si X es una v.a. con *f.d.p. f(x)* y función de distribución acumulada $F(x)$ entonces

$$\frac{dF(x)}{dx} = \frac{d}{dx}\left(\int_{-\infty}^{x} f(t)dt\right) = f(x) \quad \text{donde} \quad F(x) \text{ sea derivable}$$

Es decir, se puede obtener la función de densidad de X a partir de su F.d.a.

2- Como en el caso discreto vale

Si $a < b$ entonces $F(a) \le F(b)$ (es decir *F(x)* es una función creciente)

Y además se cumple que

$$\lim_{x\to\infty} F(x) = \lim_{x\to\infty} \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{\infty} f(t)dt = 1$$

$$\lim_{x\to-\infty} F(x) = \lim_{x\to-\infty} \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{-\infty} f(t)dt = 0$$

Ejemplos:

1- Supongamos que *X* es una v.a. continua con *f.d.p.* dada por

$$f(x) = \begin{cases} C(4x - 2x^2) & \text{si } 0 \le x \le 2 \\ 0 & \text{caso contrario} \end{cases}$$

a) ¿Cuál es el valor de C?

b) Hallar $P(X > 1)$

c) Hallar la F.d.a. de *X*

Solución:

a) Por lo dicho en la observación 1, se debe cumplir que $1=\int_{-\infty}^{\infty} f(x)dx$, por lo tanto

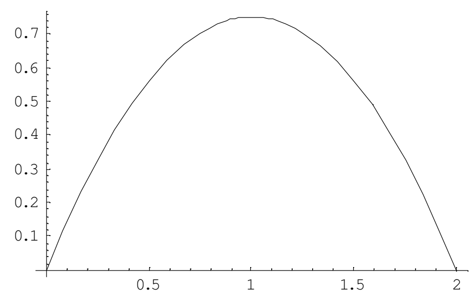
$$1=\int_{-\infty}^{\infty} f(x)dx=\int_{-\infty}^{0} 0dx+\int_{0}^{2} C\left(4x-2x^2\right)dx+\int_{2}^{\infty} 0dx=\int_{0}^{2} C\left(4x-2x^2\right)dx$$

Entonces

$$C\int_{0}^{2}\left(4x-2x^2\right)dx=C\left(4\frac{x^2}{2}-2\frac{x^3}{3}\right)\Bigg|_{0}^{2}=C\frac{8}{3}=1$$

$\therefore \quad C=\frac{3}{8}$

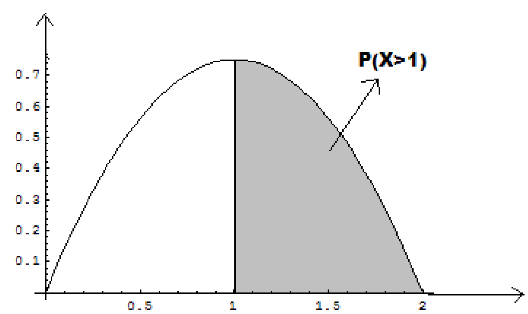
Es útil hacer un gráfico de la densidad



b) Para calcular la probabilidad que X sea mayor que 1, planteamos

$P(X>1)=$

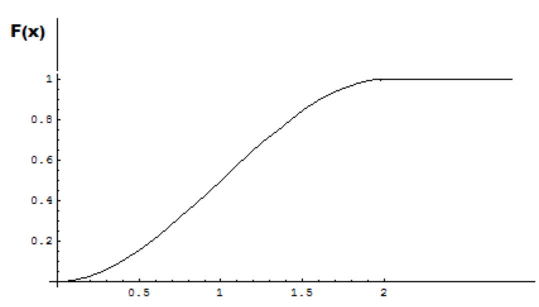
$$=\int_{1}^{2}\frac{3}{8}\left(4x-2x^2\right)dx=\frac{1}{2}$$



c) Para calcular la F.d.a. notar que tenemos tres casos: $x<0$, $0\le x\le 2$ y $x>2$, por lo tanto

$$F(x)=\begin{cases} 0 & \text{si} \quad x<0 \\ \int_{0}^{x}\frac{3}{8}\left(4t-2t^2\right)dt & \text{si} \quad 0\le x\le 2 \\ 1 & \text{si} \quad x>2 \end{cases} \qquad \text{es decir} \qquad F(x)=\begin{cases} 0 & \text{si} \quad x<0 \\ \frac{3}{4}x^2\left(1-\frac{x}{3}\right) & \text{si} \quad 0\le x\le 2 \\ 1 & \text{si} \quad x>2 \end{cases}$$

El gráfico de la F.d.a. es



Se podría haber calculado la $P(X>1)$ a partir de la F.d.a. de la siguiente forma

$$P(X > 1) = 1 - P(X \le 1) = 1 - F(1) = 1 - \frac{3}{4}1^2\left(1 - \frac{1}{3}\right) = 1 - \frac{3}{4} \times \frac{2}{3} = \frac{1}{2}$$

2- El tiempo de vida en horas que una computadora funciona antes de descomponerse es una v.a. continua con *f.d.p.* dada por

$$f(x) = \begin{cases} 0.01e^{-0.01x} & \text{si} \quad x \ge 0 \\ 0 & \text{si} \quad x < 0 \end{cases}$$

a) hallar la F.d.a. de X

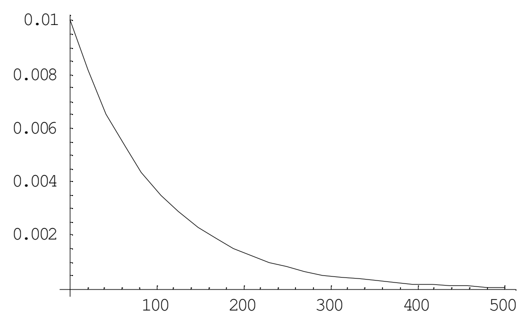
b) ¿Cuál es la probabilidad que la computadora funcione entre 50 y 150 horas antes de descomponerse?

c) ¿Cuál es la probabilidad que una computadora se descomponga antes de registrar 100 horas de uso?

d) ¿Cuál es la probabilidad que exactamente 2 de 5 computadoras se descompongan antes de registrar 100 horas de uso?. Asumir que las computadoras trabajan en forma independiente.

Solución:

a) Hacemos un gráfico de la densidad y entonces observamos claramente que hay dos casos a considerar para calcular la F.d.a.: $x \ge 0$ y $x < 0$



Si $x < 0$ entonces

$$F(x) = P(X \le x) = \int_{-\infty}^{0} 0dx = 0$$

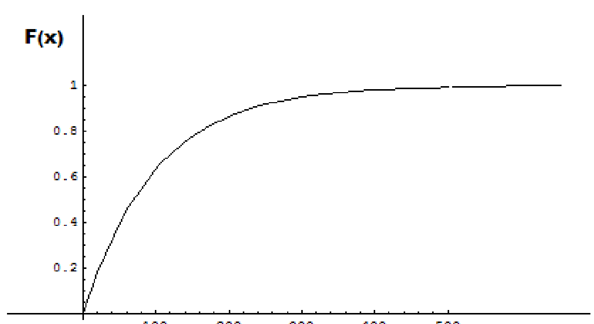
Y si $x \ge 0$ tenemos

$$F(x) = P(X \le x) = \int_{-\infty}^{0} 0dt + \int_{0}^{x} 0.01e^{-0.01t}dt =$$

$$= 0 + 0.01\left(\frac{e^{-0.01t}}{-0.01}\right)\Bigg|_0^x = 1 - e^{-0.01x}$$

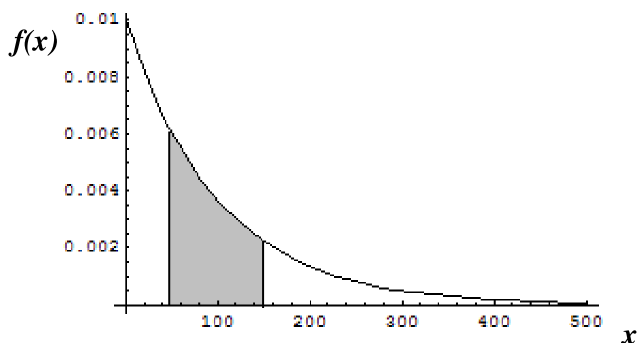
$$\therefore \quad F(x) = \begin{cases} 1 - e^{-0.01x} & \text{si } x \ge 0 \\ 0 & \text{caso contrario} \end{cases}$$

La gráfica de la F.d.a. es



Se pide calcular $P(50 \le X \le 150)$, lo hacemos con la F.d.a.:

$$P(50 \leq X \leq 150) = F(150) - F(50) = e^{-0.01\times 50} - e^{-0.01\times 150} = 0.3834$$



c) Hay que calcular $P(X < 100)$

$$P(X < 100) = F(100) = 1 - e^{-0.01\times 100} = 1 - e^{-1} \approx 0.63212$$

d) Podemos definir la v.a.

Y: "número de computadores entre 5 que se descomponen antes de las 100 horas de uso"
Entonces $Y \sim B(5,p)$ donde $p = P(X < 100)$
Por lo tanto hay que calcular

$$P(Y = 2) = \binom{5}{2} p^2 (1-p)^3 = \frac{5!}{2!3!}\left(1 - e^{-1}\right)^2 \left(e^{-1}\right)^3 \approx 0.198937$$

3.6 – Esperanza de una variable aleatoria continua

Para una v.a. discreta la $E(X)$ se definió como la suma de los $x_i\, p(x_i)$. Si X es una v.a. continua con *f.d.p.* *f(x)*, se define $E(X)$ sustituyendo la sumatoria por integración y $p(x_i)$ por *f(x)*.

La esperanza de una v.a. continua X con *f.d.p.* *f(x)* se define como

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$$

Ejemplo:
Para ciertas muestras de minerales la proporción de impurezas por muestra, es una v.a. X con f.d.p. dada por

$$f(x) = \begin{cases} \frac{3}{2}x^2 + x & \text{si } \ 0 < x < 1 \\ 0 & \text{caso contrario} \end{cases}$$

Hallar la esperanza de X

Solución:
Se plantea

$$E(X)=\int_0^1 x\left(\frac{3}{2}x^2+x\right)dx=\left.\left(\frac{3}{2}\frac{x^4}{4}+\frac{x^3}{3}\right)\right|_0^1=\left(\frac{3}{2}\times\frac{1}{4}+\frac{1}{3}\right)=\frac{17}{24}$$

A menudo se desea calcular la esperanza de una función de *X*, *Y* = *h(X)*, esto se puede hacer hallando previamente la densidad de *Y* y luego calcular *E(Y)* aplicando la definición anterior.
Otra forma de calcular *E(Y)* sin hallar la densidad de *Y* está dada por el siguiente

Teorema: Si *X* es una v.a. continua con *f.d.p. f(x)* y *h(X)* es cualquier función de *X*, entonces

$$E(h(X))=\int_{-\infty}^{\infty}h(x)f(x)dx$$

Dem.) sin demostración

Ejemplo: En el ejemplo anterior supongamos que el valor en dólares de cada muestra es $Y=h(X)=5-0.5X$. Encontrar la esperanza de *Y*

Podemos hallar la esperanza de *Y* encontrando previamente su *f.d.p.*
Para esto se encuentra la *F.d.a.* de *Y* , para luego hallar la densidad de *Y* derivando la *F.d.a*
Anotamos $G(y)$ y $g(y)$ a la *F.d.a* de *Y* y a la densidad de *Y* respectivamente

$$G(y)=P(Y\le y)=P(5-0.5X\le y)=P\left(X\ge\frac{5-y}{0.5}\right)=1-P\left(X\le\frac{5-y}{0.5}\right)=1-F\left(\frac{5-y}{0.5}\right)$$

Donde $F(x)=P(X\le x)$

Entonces

$$g(y)=\frac{d}{dy}\left(1-F\left(\frac{5-y}{0.5}\right)\right)=-f\left(\frac{5-y}{0.5}\right)\left(-\frac{1}{0.5}\right)=f\left(\frac{5-y}{0.5}\right)\left(\frac{1}{0.5}\right)=$$

$$=\begin{cases}\left(\frac{3}{2}\left(\frac{5-y}{0.5}\right)^2+\frac{5-y}{0.5}\right)\left(\frac{1}{0.5}\right) & \text{si} \quad 0<\frac{5-y}{0.5}<1\\ 0 & \text{caso contrario}\end{cases}$$

O sea

$$g(y)=\begin{cases}\left(\frac{3}{2}\left(\frac{5-y}{0.5}\right)^2+\frac{5-y}{0.5}\right)\left(\frac{1}{0.5}\right) & \text{si} \quad \frac{9}{2}<y<5\\ 0 & \text{caso contrario}\end{cases}$$

Ahora calculamos la $E(Y)$

$$E(Y)=\int_{9/2}^{5}y\left(\frac{3}{2}\left(\frac{5-y}{0.5}\right)^2+\frac{5-y}{0.5}\right)\frac{1}{0.5}dy=\frac{223}{46}$$

Aplicando el teorema anterior los cálculos se reducen:

$$E(Y)=E(h(X))=\int_0^1 \underbrace{(5-0.5x)}_{h(x)}\underbrace{\left(\frac{3}{2}x^2+x\right)}_{f(x)}dx=\frac{223}{46}$$

Notar que de la misma forma que en el caso discreto, si $h(x)=ax+b$, es decir si h es una función lineal, aplicando las propiedades de linealidad de la integral tenemos

$$E(aX+b)=aE(X)+b$$

En el ejemplo anterior se podía encontrar la esperanza de Y haciendo

$$E(Y)=E(h(X))=E(5-0.5X)=5-0.5E(X)=5-0.5\times\frac{17}{24}=\frac{223}{46}$$

Varianza de una variable aleatoria continua

Sea X una v.a. continua con *f.d.p. f(x)* y sea $E(X)=\mu$, entonces la varianza de X es

$$V(X)=\sigma_X^2=E\left[(X-\mu)^2\right]=\int_{-\infty}^{\infty}(x-\mu)^2 f(x)dx$$

La interpretación de la varianza de una v.a. continua es la misma que para el caso discreto. Además sigue valiendo la igualdad

$$V(X)=E\left(X^2\right)-\mu^2$$

Pues en la demostración hecha para el caso discreto si sustituyen las sumatorias por integrales. Por la misma razón, también vale que

$$V(aX+b)=a^2V(X) \qquad \text{y} \qquad \sigma_{aX+b}=|a|\sigma_X$$

Ejemplo:
Calculamos la varianza de $Y=5-0.5X$

$$V(Y)=0.5^2V(X)$$

$$V(X)=E(X^2)-\mu^2=E(X^2)-\left(\frac{17}{24}\right)^2$$

$$E(X^2)=\int_0^1 x^2\left(\frac{3}{2}x^2+x\right)dx=\left.\left(\frac{3}{2}\frac{x^5}{5}+\frac{x^4}{4}\right)\right|_0^1=\left(\frac{3}{2}\times\frac{1}{5}+\frac{1}{4}\right)=\frac{11}{20}$$

$$\therefore \quad V(X)=\frac{11}{20}-\left(\frac{17}{24}\right)^2 \quad\Rightarrow\quad V(Y)=0.5^2\left(\frac{11}{20}-\left(\frac{17}{24}\right)^2\right)=\frac{139}{11520}$$