



Curso Propedéutico de Estadística-Probabilidad clase 3





Temario

- Otras Distribuciones discretas
 - -Distribución Binomial Negativa
 - -Distribución de Poisson
- Funciones de probabilidad continua
 - -Distribución uniforme
- Función de distribución acumulada





La distribución binomial negativa se basan en un experimento que satisface las siguientes condiciones:

- 1. El experimento consiste en una secuencia de ensayos independientes.
- 2. Cada ensayo puede dar por resultado un éxito (S) o una falla (F).
- 3. La probabilidad de éxito es constante de un ensayo a otro, por tanto,

$$P(S \text{ en el ensayo } i) = p \quad \text{con } i = 1, 2, 3, ...$$





¿Cuál es la variable aleatoria de interés en la distribución binomial negativa ?

La variable aleatoria de interés es X

→ El número de fallas que preceden al éxito r-ésimo

Se llama variable aleatoria binomial negativa porque, en contraste con la variable aleatoria binomial, el número de éxitos es fijo y el número de ensayos es aleatorio.

¿Distribución binomial?

Binomial -> el número de ensayos es fijo y el número de éxitos es aleatorio





¿Cuál es son los posibles valores que toma la variable discreta binomial negativa?

Naturales (numerables) \rightarrow 0,1,2,....

La función de masa de probabilidad de la variable aleatoria binomial negativa X con los parámetros r = número de éxitos (S) y p = P(S)

$$nb(x; r, p) = {x + r - 1 \choose r - 1} p^{r} (1 - p)^{x} \quad x = 0, 1, 2, \dots$$

nb(7; r=3, *p*=0.5)



Si la probabilidad de éxito es 0.5, ¿cuál es la probabilidad de obtener 7 fracasos antes del 3er éxito?

nb(2; r=10, p=0.2)



Si la probabilidad de éxito es 0.2, ¿cuál es la probabilidad de obtener 2 fracasos antes del 10° éxito?





¿Por qué se obtiene esa función de masa de probabilidad?

$$nb(x; r, p) = {x + r - 1 \choose r - 1} p^{r} (1 - p)^{x} \quad x = 0, 1, 2, \dots$$

nb(7; r=3, *p*=0.5)



Si la probabilidad de éxito es 0.5, ¿cuál es la probabilidad de obtener 7 fracasos antes del 3er éxito?

*El último ensayo debe ser un éxito.

*y antes del último éxito debe haber exactamente 2 éxitos en los primeros 9 ensayos;

$$nb(7;3,p) = \left\{ \binom{9}{2} \cdot p^2 (1-p)^7 \right\} \cdot p = \binom{9}{2} \cdot p^3 (1-p)^7$$





$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

$$\binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{4!}{2! \cdot 2!} = \frac{4 \cdot 3 \cdot \cancel{2} \cdot \cancel{1}}{\cancel{2} \cdot \cancel{1} \cdot 2 \cdot 1} = 3 \cdot 2 = 6$$





Ejercicio 1

Un pediatra desea reclutar **5 parejas**, cada una de las cuales espera a su **primer hijo**, para participar en un nuevo **programa de parto natural**.

Sea S ={está de acuerdo en participar} $\rightarrow p = P$ (una pareja seleccionada al azar está de acuerdo en participar). Si p = 0.2 (con)

¿Cuál es la probabilidad de que **15 parejas deban ser entrevistadas** antes de encontrar **5 que estén de acuerdo** en participar?



Es decir, ¿cuál es la probabilidad de que ocurran 10 fallas antes del 5to éxito?

Al sustituir r = 5, p = 0.2 y x = 10 en nb(x; r, p)

$$nb(10; 5, 0.2) = {14 \choose 4} (0.2)^5 (0.8)^{10} = 0.034$$





Observaciones:

- 1) En algunas fuentes la variable aleatoria binomial negativa es: el **número de ensayos** *X* + *r* en lugar del número de fallas.
- 2) En el caso especial r=1, la función de masa de probabilidad es

$$nb(x; 1, p) = (1 - p)^{x}p$$
 $x = 0, 1, 2,...$ Distribución geométrica

3) La media y varianza están dadas por:

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





Ejercicio 2:

En la serie de campeonato de la NBA (National Basketball Association), el equipo que gane 4 de 7 juegos será el ganador. Suponga que los equipos *A* y *B* se enfrentan en los juegos de campeonato y que el equipo *A* tiene una probabilidad de 0.55 de ganarle al equipo *B*.

¿Cuál es la probabilidad de que el equipo A gane la serie en 6 juegos?

$$nb(x; r, p) = {x + r - 1 \choose r - 1} p^{r} (1 - p)^{x} \quad x = 0, 1, 2, \dots$$

$$\binom{5}{3}0.55^4(1-0.55)^{6-4} = 0.1853$$





Se dice que una variable aleatoria discreta X tiene una distribución de Poisson con parámetro μ (μ > 0) si la función de masa de probabilidad de X es:

$$p(x; \mu) = \frac{e^{-\mu} \cdot \mu^x}{x!}$$
 $x = 0, 1, 2, 3, ...$ Media dist.

Se extiende a todos los números enteros no negativos (numerable)

Las distribuciones binomial y binomial negativa se dedujeron partiendo de un experimento compuesto de ensayos -> No hay un experimento simple para la Poisson.

La distribución Poisson expresa la probabilidad de que ocurra un **determinado número de eventos.**

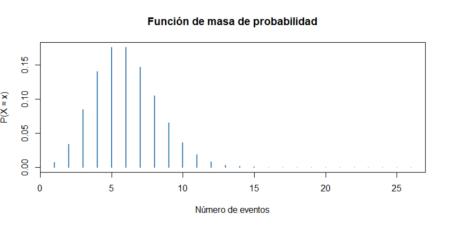
Algunas veces se estandarizan los eventos durante cierto intervalo de tiempo (o intervalo espacial).



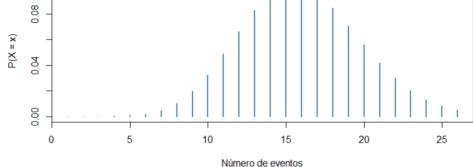


$$p(x; \mu) = \frac{e^{-\mu} \cdot \mu^x}{x!}$$
 $x = 0, 1, 2, 3, ...$

Gráficas de su masa de probabilidad $\mu=5$ (izquierda) y $\mu=15$ (derecha) :



Función de masa de probabilidad







$$p(x; \mu) = \frac{e^{-\mu} \cdot \mu^x}{x!}$$
 $x = 0, 1, 2, 3, ...$

No es evidente por inspección que $p(x; \mu)$ especifica una función de masa de **probabilidad legítima**

¿A qué nos referimos con que sea legítima?

- 1) En primer lugar, $p(x; \mu) > 0$ para cada valor x posible, debido a la exigencia de que $\mu > 0$
- 2) El hecho de que $\Sigma p(x; \mu)$ =1 es una consecuencia de la **expansión en** series de Maclaurin de e^{μ}

$$e^{\mu} = 1 + \mu + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \dots = \sum_{x=0}^{\infty} \frac{\mu^x}{x!}$$

$$1 = \sum_{x=0}^{\infty} \frac{e^{-\mu} \cdot \mu^x}{x!}$$





Ejercicio 1:

Sea X el número de trampas (cierto tipo de defectos) en un tipo particular de transistor metalóxido semiconductor, y suponga que tiene una distribución de Poisson con μ = 2.

¿Cuál es la probabilidad de que haya exactamente tres trampas?

$$P(X = 3) = p(3; 2) = \frac{e^{-2}2^3}{3!} = 0.180,$$

¿Cuál es la probabilidad de que haya a lo más tres trampas?

$$P(X \le 3) = \sum_{x=0}^{3} \frac{e^{-2}2^x}{x!} = 0.135 + 0.271 + 0.271 + 0.180 = 0.857$$





La distribución Poisson se puede ver cómo límite de otra distribución discreta.

¿Cuál distribución discreta?

Suponga que en la función de masa de probabilidad binomial $b(x; n, p), n \to \infty$ y $p \to 0$ de tal modo que np tienda a un valor $\mu > 0$. Entonces $b(x; n, p) \to p(x; \mu)$.

De acuerdo con esta proposición, en cualquier experimento binomial en el cual n es grande y p pequeña, $b(x; n, p) \approx p(x; \mu)$, donde $\mu = np$

Como regla empírica esta aproximación puede ser aplicada con seguridad si *n > 50 y np < 5.





La distribución Poisson se puede ver cómo límite de la distribución binomial

x	n=30, p=0.1	n = 100, p = 0.3	n = 300, p = 0.1	Poisson, $\mu = 3$
0	0.042391	0.047553	0.049041	0.049787
1	0.141304	0.147070	0.148609	0.149361
2	0.227656	0.225153	0.224414	0.224042
3	0.236088	0.227474	0.225170	0.224042
4	0.177066	0.170606	0.168877	0.168031
5	0.102305	0.101308	0.100985	0.100819
6	0.047363	0.049610	0.050153	0.050409
7	0.018043	0.020604	0.021277	0.021604
8	0.005764	0.007408	0.007871	0.008102
9	0.001565	0.002342	0.002580	0.002701
10	0.000365	0.000659	0.000758	0.000810





La distribución Poisson es una distribución que posee media y varianza iguales.

Si X tiene una distribución de Poisson con parámetro μ , entonces $E(X) = V(X) = \mu$.

Puesto que $b(x; n, p) \to p(x; m)$ a medida que $n \to \infty$, $p\to 0$, $np\to \mu$, la media y la varianza de una **variable binomial** deberán **aproximarse** a las de una **variable de Poisson**.

¿Cuál es la media y varianza de la binomial?





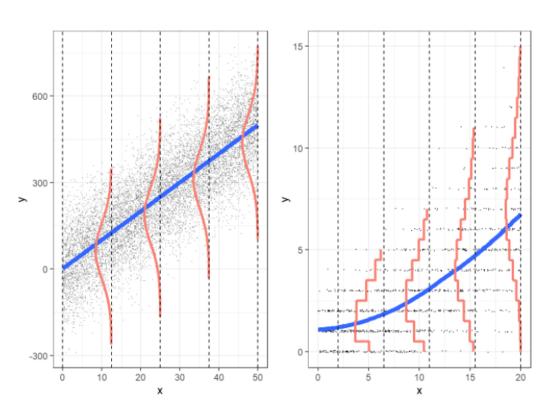
Estos límites son $np \rightarrow \mu$ y $np(1 - p) \rightarrow \mu$.

Aplicaciones de las distribuciones Poisson y Binomial negativa

Dato cultural:

Regresión
Lineal

Distribución
normal





Distribución Poisson

 \bigcirc

Binomial
Negativa
(cuando la
media no es
igual a la
varianza).





Una variable aleatoria es **continua** si las siguientes dos condiciones se cumplen:

- 1a. Su conjunto de valores posibles se compone de todos los números que hay en un solo intervalo sobre la línea de numeración (posiblemente de extensión infinita)→
- **1b.** todos los números en una unión disjunta de intervalos (por ejemplo, [0,10] ∪ [20,30])
- **2.** Ningún valor posible de la variable tiene probabilidad positiva, esto es, P(X = c)=0 con cualquier valor posible de c.

El estudio de variables continuas requiere matemáticas continuas de **cálculo**: **integrales** y **derivadas**.





Ejemplo:

En el estudio de la ecología de un lago se mide la **profundidad** en lugares seleccionados, entonces *X* la **profundidad** en ese lugar es una **variable aleatoria continua**. Si *B* es la profundidad máxima. → ¿qué valores toma la v.a. X?

Observación

Se podría argumentar que aunque en principio las variables como altura, peso y temperatura son continuas, en la práctica las **limitaciones de los instrumentos de medición** nos restringen a un mundo discreto.

Sin embargo, los modelos continuos a menudo representan muy bien de forma aproximada situaciones del mundo real y con frecuencia es más fácil trabajar con matemáticas continuas que con matemáticas discretas.

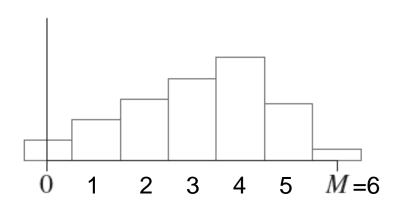




Volviendo al ejemplo del lago, sea M la profundidad máxima (en metros), así que cualquier número en el intervalo [0, M] es un valor posible de X.

Si se "discretiza" X midiendo la profundidad al metro más cercano, entonces los valores posibles son enteros no negativos menores o iguales a M

Se toma una muestra y se hace un **histograma** para representar la **distribución discretizada**:



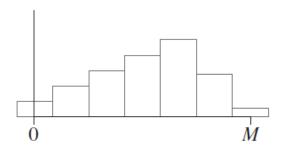
¿Qué es un histograma?

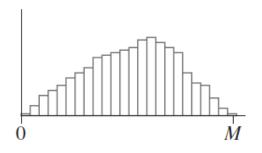
Es una representación gráfica de una variable en forma de barras, donde la superficie de cada barra es proporcional a la frecuencia de los valores representados.

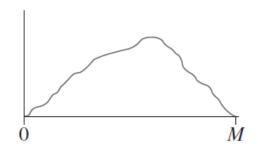




¿Qué pasa si se obtienen mediciones más finas de la profundidad?







La curva suave de la derecha representa la distribución de probabilidad de una variable continua.

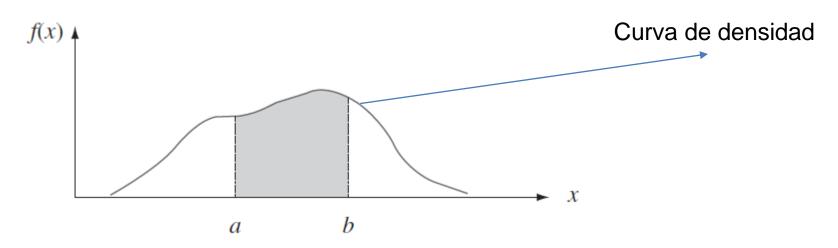
La **probabilidad** de que la profundidad en un punto seleccionado al azar se encuentre entre a y b es simplemente el **área bajo la curva** regular entre a y b





Sea X una variable aleatoria continua. Entonces, una **distribución de probabilidad** o **función de densidad de probabilidad** (pdf) de X es una función f(x) de modo tal que para dos números cualesquiera a y b con $a \le b$,

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







Para que f(x) sea una función de densidad de probabilidad legítima debe satisfacer las dos siguientes condiciones:

1.
$$f(x) \ge 0$$
 con todas las x

2.
$$\int_{-\infty}^{\infty} f(x) dx = \text{ área bajo toda la gráfica de } f(x) = 1$$





Propiedad 2 (v.a. continuas)

2. Ningún valor posible de la variable tiene probabilidad positiva, esto es, P(X = c)=0 con cualquier valor posible de c.

¿Por qué?

$$P(X = c) = \int_{c}^{c} f(x) dx = \lim_{\varepsilon \to 0} \int_{c-\varepsilon}^{c+\varepsilon} f(x) dx = 0$$

El hecho de que P(X = c) = 0 cuando X es continua tiene una importante **consecuencia práctica**: la probabilidad de que X quede en algún intervalo entre a y b **no depende** de si el **límite inferior** a o **el límite superior** b están incluido en el cálculo de probabilidad

$$P(a \le X \le b) = P(a < X < b) = P(a < X \le b) = P(a \le X < b)$$





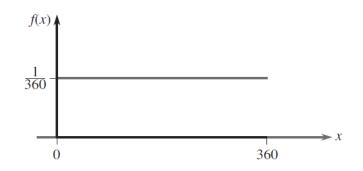
Ejemplo:

Considere la línea de referencia que conecta la válvula de un neumático con el punto central, y sea *X* el ángulo medido en el sentido de las manecillas del reloj.



¿Qué valores puede tomar X?

$$f(x) = \begin{cases} \frac{1}{360} & 0 \le x < 360\\ 0 & \text{de lo contrario} \end{cases}$$



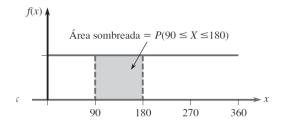
¿Es una función de densidad legítima?





¿Cuál es la **probabilidad** de que el ángulo sea de entre 90° y 180°?

$$P(90 \le X \le 180) = \int_{90}^{180} \frac{1}{360} \, dx = \frac{x}{360} \Big|_{x=90}^{x=180} = \frac{1}{4} = 0.25$$



¿Cuál es la **probabilidad** de que el ángulo de ocurrencia esté dentro de 90° respecto a la línea de referencia?



$$P(0 \le X \le 90) + P(270 \le X < 360) = 0.25 + 0.25 = 0.50$$





Del ejemplo anterior, debido a que siempre que $0 \le a \le b \le 360$ en el ejemplo $P(a \le X \le b)$ depende sólo del **anch**o **b** - **a** del **intervalo**, se dice que X tiene una **distribución uniforme**

Se dice que una variable aleatoria continua X tiene una **distribución uniforme** en el intervalo [A, B] si la función de densidad de probabilidad de X es

$$f(x; A, B) = \begin{cases} \frac{1}{B - A} & A \le x \le B\\ 0 & \text{de lo contrario} \end{cases}$$





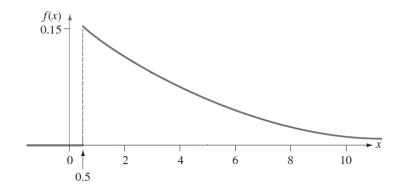
Ejemplo 2:

"Intervalo de tiempo" en el flujo de tránsito es el lapso transcurrido entre el tiempo en que un automóvil termina de pasar por un punto fijo y el instante en que el siguiente vehículo comienza a pasar por ese punto.

Sea *X* el intervalo de tiempo para dos automóviles consecutivos seleccionados al azar en una autopista durante un periodo de tráfico intenso.

La siguiente función de densidad representa esta variable aleatoria.

$$f(x) = \begin{cases} 0.15e^{-15(x-0.5)}. & x \ge 0.5\\ 0 & \text{de lo contrario} \end{cases}$$







¿Es una función de densidad válida?

$$\int_{-\infty}^{\infty} f(x) dx = \int_{0.5}^{\infty} 0.15e^{-0.15(x-0.5)} dx = 0.15e^{0.075} \int_{0.5}^{\infty} e^{-0.15x} dx$$
$$= 0.15e^{0.075} \cdot \frac{1}{0.15} e^{-(0.15)(0.5)} = 1 \qquad \int_{a}^{\infty} e^{-kx} dx = (1/k)e^{-k \cdot a}$$

¿Cuál es la **probabilidad** de que el intervalo de tiempo sea cuando mucho de 5 segundos?

$$P(X \le 5) = \int_{-\infty}^{5} f(x) dx = \int_{0.5}^{5} 0.15e^{-0.15(x - 0.5)} dx$$





¿Cuál es la **probabilidad** de que el intervalo de tiempo sea cuando mucho de 5 segundos?

$$P(X \le 5) = \int_{-\infty}^{5} f(x) dx = \int_{0.5}^{5} 0.15e^{-0.15(x - 0.5)} dx$$

$$= 0.15e^{0.075} \int_{0.5}^{5} e^{-0.15x} dx = 0.15e^{0.075} \cdot \left(-\frac{1}{0.15} e^{-0.15x} \Big|_{x=0.5}^{x=5} \right)$$

$$= e^{0.075}(-e^{-0.75} + e^{-0.075}) = 1.078(-0.472 + 0.928) = 0.491$$

$$= P(\text{menos de 5 s}) = P(X < 5)$$





La función de distribución acumulada F(x) de una variable aleatoria discreta es la probabilidad $P(X \le x)$ (un número real x)

Para una variable aleatoria discreta X, se obtiene al sumar la función de masa de probabilidad f(u) a lo largo de todos los valores posibles u que satisfagan $u \le x$

$$F(x) = \mathrm{P}[X \leq x] = \sum_{u \leq x} f(u)$$





La función de distribución acumulada F(x) de una variable aleatoria es la probabilidad $P(X \le x)$ (un número real x)

La función de distribución acumulada de una variable aleatoria continua se obtiene integrando la función de densidad de probabilidad f(y) entre los límites $-\infty$ y x

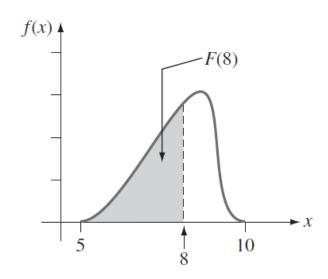
La **función de distribución acumulada** F(x) de una variable aleatoria continua X se define para todo número x como

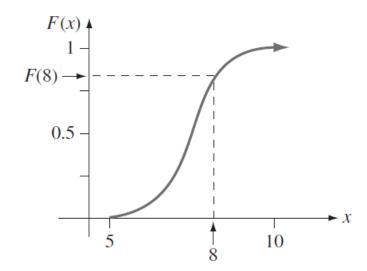
$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y) \, dy$$





La función de distribución acumulada F(x) de una variable aleatoria la probabilidad $P(X \le x)$ (un número real x)







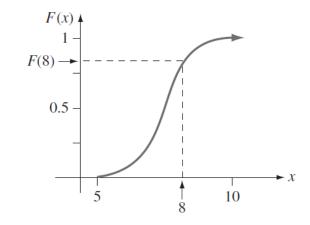


La función de distribución acumulada F(x) satisface las siguientes propiedades:

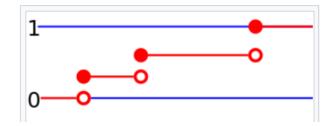
1.
$$0 \le F(x) \le 1$$
.

2.
$$\lim_{x o \infty} F(x) = 1$$
.

3.
$$\lim_{x \to -\infty} F(x) = 0$$
.



- 4. Es monótona no decreciente, es decir, si $x_1 \leq x_2$ entonces $F(x_1) \leq F(x_2)$.
 - 5. Es continua por la derecha, es decir, $\lim_{x o a^+} F(x) = F(a^+)$.

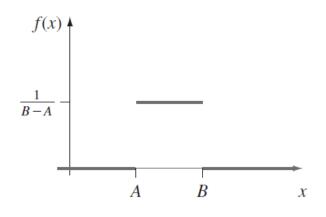






Ejercicio 1:

Sea X el espesor de una cierta lámina de metal con distribución uniforme en [A, B]. La función de densidad se muestra en la figura siguiente:



¿Cuál es la función de distribución acumulada F(x)?

Para
$$x < A \Rightarrow F(x) = 0$$



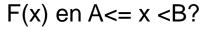
$$F(x) = 0$$

Para
$$x > B$$
 \Rightarrow $F(x) = 1$

$$F(x) = 1$$

$$F(x) = \begin{cases} 0 & x < A \\ & A \le x < B \\ 1 & x \ge B \end{cases}$$

¿Cuál es la función de distribución acumulada







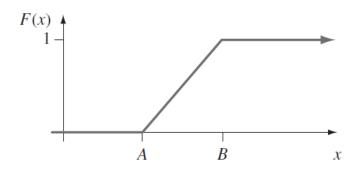
Ejercicio 1:

Sea **X el espesor** de una cierta lámina de metal con **distribución uniforme** en [A, B]. La función de densidad se muestra en la figura siguiente:

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{A}^{x} \frac{1}{B - A} dy = \frac{1}{B - A} \cdot y \Big|_{y = A}^{y = x} = \frac{x - A}{B - A}$$

$$F(x) = \begin{cases} 0 & x < A \\ \frac{x - A}{B - A} & A \le x < B \\ 1 & x \ge B \end{cases}$$

¿Cómo se ve la gráfica de F(x)?







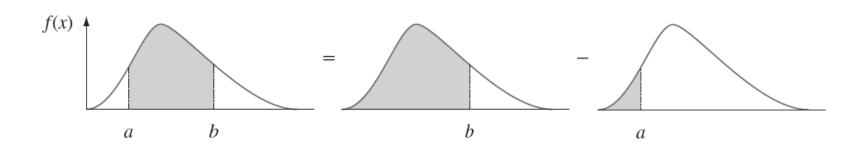
Un uso importante que se le puede dar a la función de distribución acumulada es que las probabilidades de intervalos pueden ser calculadas empleando F(x).

Sea X una variable aleatoria continua con función de densidad de probabilidad f(x) y función de distribución acumulada F(x). Entonces para cualquier número a,

$$P(X > a) = 1 - F(a)$$

y para dos números cualesquiera a y b con a < b,

$$P(a \le X \le b) = F(b) - F(a)$$







Ejemplo:

Suponga que la función de densidad de probabilidad de la magnitud X de una carga dinámica sobre un puente (en newtons) está dada por

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \le x \le 2\\ 0 & \text{de lo contrario} \end{cases}$$

¿Cuál es la función de distribución acumulada F(x)?

Para cualquier número x entre 0 y 2

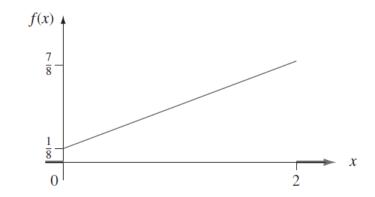
$$F(x) = \int_{-\infty}^{x} f(y) \, dy = \int_{0}^{x} \left(\frac{1}{8} + \frac{3}{8}y\right) dy = \frac{x}{8} + \frac{3}{16}x^2 \quad \Rightarrow \quad F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{8} + \frac{3}{16}x^2 & 0 \le x \le 2 \\ 1 & 2 < x \end{cases}$$

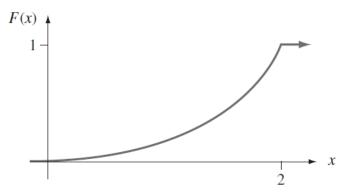




$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \le x \le 2\\ 0 & \text{de lo contrario} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{8} + \frac{3}{16}x^2 & 0 \le x \le 2 \\ 1 & 2 < x \end{cases}$$





¿Cuál es la probabilidad de que la carga esté entre 1 y 1.5?

$$P(1 \le X \le 1.5) =$$

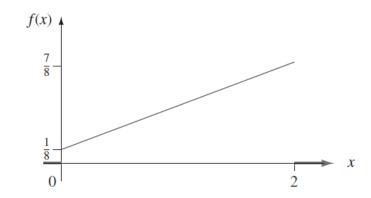
$$= F(1.5) - F(1) = \left[\frac{1}{8} (1.5) + \frac{3}{16} (1.5)^2 \right] - \left[\frac{1}{8} (1) + \frac{3}{16} (1)^2 \right] = \frac{19}{64} = 0.297$$

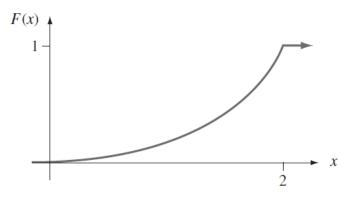




$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x & 0 \le x \le 2\\ 0 & \text{de lo contrario} \end{cases}$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{8} + \frac{3}{16}x^2 & 0 \le x \le 2 \\ 1 & 2 < x \end{cases}$$





¿Cuál es la probabilidad de que la carga sea mayor a 1?

$$= 1 - P(X \le 1) = 1 - F(1) = 1 - \left[\frac{1}{8} (1) + \frac{3}{16} (1)^2 \right] = \frac{11}{16} = 0.688$$





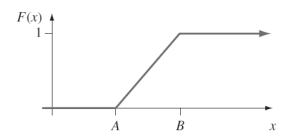
Se puede obtener la función de densidad f(x) a partir de la función de distribución acumulada es que las probabilidades de intervalos pueden ser calculadas empleando F(x).

¿Cómo se obtiene?

Si X es una variable aleatoria continua con función de densidad de probabilidad f(x) y función de distribución acumulada F(x), entonces en cada x que hace posible que la derivada F'(x) exista, F'(x) = f(x).

Ejemplo: Distribución continua uniforme en intervalo [A,B]

F(x) es derivable excepto en x = A y x = B, donde la gráfica de F(x) tiene esquinas. Como F(x) = 0 para x < A y F(x) = 1 para x > B, F'(x) = 0 = f(x) $(-\infty, A)$ y (B, ∞)



$$F'(x) = \frac{d}{dx} \left(\frac{x - A}{B - A} \right) = \frac{1}{B - A} = f(x)$$







Centro de Investigación en Matemáticas, A.C.

Unidad Monterrey

CIDICS, Campus de la Salud, UANL Ave. Carlos Canseco s/n Col Mitras Centro, Monterrey N.L., México, C.P.66460 México Tel +52 (81) 8329 4000 Ext 1778 cimat@cimat.mx / www.cimat.mx





