DISTRIBUCIONES DISCRETAS DE PROBABILIDAD RESUMEN

Distribución Discreta Uniforme

Distribución de Bernoulli

Distribución Binomial

Distribución de Poisson

Distribución Geométrica

Distribución Binomial Negativa

Distribución Hipergeométrica

Modelos teóricos de distribuciones de probabilidad.
Funciones de masa teóricas que pueden resultar adecuadas para determinadas variables aleatorias

En el caso de una variable discreta el comportamiento de la variable aleatoria queda descrito por la probabilidad de que la variable X tome un determinado **valor especifico x (P(X=x)).**

Describen la probabilidad de ocurrencia de resultados específicos en un conjunto discreto de valores posibles. Estas distribuciones están definidas en un conjunto finito o infinito numerable de valores, lo que significa que solo pueden tomar valores discretos, como enteros o contables.

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución Discreta Uniforme

Tipo de distribución de probabilidad en la que cada valor posible dentro de un rango finito tiene la misma probabilidad de ocurrir. Es decir, todos los valores tienen una probabilidad igual de ser seleccionados.

- Con *n* valores posibles, la probabilidad de que ocurra cada valor es **1/*n***
- Comúnmente utilizada en situaciones donde no hay razón para esperar que un valor sea más probable que otro dentro del rango

Por ejemplo, considera un dado de seis caras. Cada número del 1 al 6 tiene la misma probabilidad de salir cuando se lanza el dado, lo que significa que sigue una distribución discreta uniforme. En este caso, ***n*=6** (los seis posibles valores del dado), y la probabilidad de que salga cada número es **1/6**

$P(X = k) = \frac{1}{n}$

Donde:

- $P(X = k)$ es la probabilidad de que la variable aleatoria tome el valor k.
- n es el número total de valores posibles que puede tomar la variable aleatoria.
- k es el valor específico que puede tomar la variable aleatoria.

- Notación: $X \sim U(n)$
- Campo de variación: $X = \{1, 2, \ldots, n\}$
- Función de cuantía: $\boxed{P(X = x_i) = 1/n, \ \forall n = 1, \ldots n}$
- Función de distribución: $\boxed{F(x_i) = P(X \leq x_i) = \sum_{i=1}^{k} 1/n = k/n}$

- Esperanza:

$$\boxed{\mathbb{E}(X) = \sum_{i=1}^{n} x_i P(X = x_i) = \sum_{i=1}^{n} x_i \cdot \frac{1}{n} = \frac{1+n}{2}}$$

- Varianza:

$$\boxed{\mathbb{V}\text{ar}(X) = E[X^2] - (E[X])^2 = \sum_{i=1}^{n} x_i^2 \frac{1}{n} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}}$$

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución de Bernoulli

BERNOULLI se utiliza para procesos aleatorios que tienen dos resultados. **(DICOTOMICO)**

- Modeliza situaciones en las que tenemos un experimento aleatorio con dos posibles resultados que denominaremos: **éxito o fracaso**.
- Las variables pueden tener 2 valores númericos: 0 y 1. Donde 1 corresponde a un evento y 0 a un no evento.
- La distribución está completamente definida con un único parámetro **p** que se refiere a la probabilidad de obtener éxito.
- La probabilidad no se ve afectada por el conocimiento de los resultados anteriores, es decir, los diferentes ensayos son independientes.
- La probabilidad **q** de fracaso viene dada por **q = 1 – p.**

- Notación: $X \sim Bernoulli(p)$ o $X \sim B(p)$
- Campo de variación: $X = \{0, 1\}$
- $P(X = 1) = p$ y $P(X = 0) = q = 1 - p$
- Función de cuantía: $\boxed{P(X = x) = p^x q^{1-x} \qquad x = 0, 1}$
- Función de distribución: $$\boxed{F(x) = \begin{cases} 0 & x < 0 \\ q & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}}$$
- Esperanza: $$\boxed{\mathbb{E}(X) = \sum_{i=0,1} x_i p_i = p}$$
- Varianza: $$\boxed{\mathbb{V}\mathrm{ar}(X) = \sum_{i=0,1} x_i^2 p_i - p^2 = pq}$$

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

- Distribución de Bernoulli
  - Distribución Binomial
  - Distribución Geométrica
  - Distribución Binomial Negativa

Distribución Binomial B(n,p)

- Modeliza situaciones en las que el mismo experimento de Bernoulli se repite **n veces**, dando lugar cada una de ellas a resultados completamente **independientes** de los anteriores.
- Queda definida con dos parámetros:
  - **n:** el número de veces que se repite el experimento dicotómico.
  - **p:** la probabilidad de obtener éxito.
- La variable aleatoria cuenta el número de éxitos obtenidos en los **n** ensayos independientes.
- Ejemplos:
  - Un elemento pasa o no pasa una inspección
  - Un partido político gana o pierde

- Notación: $X \sim Binomial(n, p)$ o $X \sim B(n, p)$
- Campo de variación: $X = \{0, 1, 2, \ldots, n\}$
- Función de cuantía: $\boxed{P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}}$
- Esperanza:

$$\boxed{\mathbb{E}(X) = np}$$

- Varianza:

$$\boxed{\mathbb{V}\text{ar}(X) = npq}$$

- Propiedad: La suma de v.a. binomiales independientes es otra v.a. binomial. Es decir, si $X_i \sim B(n_i; p)$ $(i = 1, 2, \ldots, K)$ entonces

$$Y = \sum_{i=1}^{K} X_i \sim B\left(\sum_{i=1}^{K} n_i, p\right)$$

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución de Poisson

Poisson es una distribución de probabilidad que describe el **número de eventos raros** que ocurren **en un intervalo de tiempo o espacio fijo**, dado que estos eventos ocurren de manera independiente y a una tasa constante.

- **La distribución de Poisson es la ley de los sucesos raros.**
- **Como en la binomial, contamos el número de veces que se presenta el suceso.**
- **Útil para modelar eventos raros pero bien definidos, adecuada cuando la tasa de ocurrencia de eventos es baja y el intervalo de tiempo o espacio es grande.**

Por ejemplo, el número de llamadas recibidas en un centro de atención telefónica en una hora, el número de errores tipográficos en una página de texto o el número de partículas radiactivas emitidas en un período de tiempo, pueden seguir una distribución de Poisson.

Características

- **Eventos independientes:** Los eventos deben ocurrir de manera independiente unos de otros. El hecho de que un evento ocurra no afecta la probabilidad de que ocurra otro evento.
- **Tasa constante:** La tasa promedio de ocurrencia de eventos debe ser constante en todo el intervalo de tiempo o espacio considerado.

Función de cuantía:

$$\boxed{P(X=k)=\frac{e^{-\lambda}.\lambda^k}{k!}}$$

Donde:

- $P(X=k)$ es la probabilidad de que ocurran exactamente k eventos en el intervalo de interés.
- λ es la tasa promedio de ocurrencia de eventos en el intervalo.
- e es la base del logaritmo natural (aproximadamente 2.71828).
- $k!$ es el factorial de k.

- Esperanza: $\boxed{\mathbb{E}\left(X\right)=\lambda}$
- Varianza: $\boxed{\mathbb{V}\text{ar}\left(X\right)=\lambda}$

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

- **Distribución de Bernoulli**
  - **Distribución Binomial**
  - **Distribución Geométrica**
  - **Distribución Binomial Negativa**

Distribución Geométrica G(p)

- Modeliza situaciones basadas en experimentos de Bernoulli, en las que nos fijamos en el número de fracasos que se producen antes de obtener el primer éxito.
- Queda definida a partir de su parámetro **p** que, al igual que en la distribución de Bernoulli, se refiere a la probabilidad de éxito.
- Implica la existencia de una **dicotomía** de posibles resultados y la independencia de las pruebas entre sí.
- Una v.a. X posee una distribución geométrica, si esta es la suma del numero de fracasos obtenidos hasta la aparición del primer éxito en la sucesión.

- Notación: $X \sim Geometrica(p)$ o $X \sim G(p)$
- Campo de variación: $X = \{0, 1, 2, \ldots\}$
- Función de cuantía: $\boxed{P(X = x) = q^x \cdot p}$ donde q = 1 - p
- Esperanza: $\boxed{\mathbb{E}(X) = \frac{q}{p}}$
- Varianza: $\boxed{\mathbb{V}ar(X) = \frac{q}{p^2}}$

- Para la distribución Geométrica no se cumple la propiedad aditiva.
- **Propiedad:** falta de memoria

$$P(X \geq a + b / X \geq a) = P(X \geq b)$$

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

- **Distribución de Bernoulli**
  - **Distribución Binomial**
  - **Distribución Geométrica**
  - **Distribución Binomial Negativa**

Distribución Binomial Negativa (BN(r,p)

- La distribución binomial negativa modeliza situaciones basadas en experimentos de Bernoulli, en las que contabilizamos los fracasos antes del r-ésimo éxito.
- Queda definida a partir de los parámetros **r** y **p**, que se refieren al número de éxitos que se pretende alcanzar y a la probabilidad de éxito respectivamente.
- La **distribución geométrica** es un **caso particular** de la d**istribución binomial negativa** para la que r = 1, es decir: **BN(1; p) = G(p).**
- Proceso:
  - Número indefinido de ensayos separados (independientes)
  - Termina cuando se alcanza un número especifico de resultados favorables
  - Experimento con extracción se llevara a cabo con devolución
- Ejemplo: sacar una bola de una urna y luego devolverla, significa que después de c/ensayo, la bola se devuelve a la urna antes del siguiente ensayo.
  - Número de caras que se obtienen al lanzar la moneda antes de obtener cruz por tercera vez.
  - Número de llamadas que realiza un comercial antes de hacer su décima venta.

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución Binomial Negativa (BN(r,p))

- Notación: $X \sim BN(r,p)$
- Campo de variación: $X = \{0, 1, 2, \ldots\}$
- Función de cuantía: $\boxed{P(X = x) = \binom{x+r-1}{x} q^x \cdot p^r}$
- Esperanza: $\boxed{\mathbb{E}(X) = r \cdot \frac{q}{p}}$
- Varianza: $\boxed{\mathbb{V}ar(X) = r \cdot \frac{q}{p^2}}$

$$P(X = k) = \binom{k+r-1}{k} \times p^r \times (1-p)^k$$

Donde:

- $P(X = k)$ es la probabilidad de observar k éxitos antes de obtener r fracasos.
- k es el número de éxitos observados.
- r es el número fijo de fracasos.
- p es la probabilidad de éxito en un solo ensayo de Bernoulli.

La notación $\binom{k+r-1}{k}$ representa el coeficiente binomial, que calcula el número de formas en que se pueden elegir k éxitos de una secuencia de $k + r - 1$ ensayos.

- Para la distribución Binomial Negativa sí se cumple la propiedad aditiva:

$$X = \sum_i^n X_i, \;\; iid \; X_i \sim G(p) \rightarrow X \sim BN(n; p)$$
$$X = \sum_i^n X_i, \;\; iid \; X_i \sim BN(r_i, p) \rightarrow X \sim BN\left(\sum_{i=1}^n r_i; p\right)$$

1. Si sumamos n variables aleatorias geométricas iid X_i que siguen una distribución geométrica con parámetro p, la suma $X = \sum_{i=1}^n X_i$ sigue una distribución binomial negativa con parámetros n y p.
2. Si sumamos n variables aleatorias binomiales negativas iid X_i que siguen una distribución binomial negativa con parámetro r_i y p, la suma $X = \sum_{i=1}^n X_i$ sigue una distribución binomial negativa con parámetro $\sum_{i=1}^n r_i$ y p.

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución Binomial Negativa (BN(r,p))

Ejemplo

En una competición de ajedrez, el vencedor de cada eliminatoria es el que logre primero la tercera victoria. Si un jugador tiene un porcentaje de partidas ganadas del 70%, y asumiendo que ningún otro participante logra clasificarse antes,
calcule:

a) ¿Cuál es la probabilidad de que el jugador tenga que jugar 5 partidas para clasificarse?
b) ¿Y la de que juegue solo 3 partidas?
c) ¿Cuál es el número esperado de partidas que jugará para clasificarse?

- Notación: $X \sim BN(r, p)$
- Campo de variación: $X = \{0, 1, 2, \dots\}$
- Función de cuantía: $\boxed{P(X = x) = \binom{x+r-1}{x} q^x \cdot p^r}$
- Esperanza: $\boxed{\mathbb{E}(X) = r \cdot \frac{q}{p}}$
- Varianza: $\boxed{\mathbb{V}\text{ar}(X) = r \cdot \frac{q}{p^2}}$

Solución

X=Número de partidas perdidas antes de ganar la tercera.

$X \sim BN(3; 0,7)$

a. *¿Cuál es la probabilidad de que el jugador tenga que jugar 5 partidas para clasificarse?*
$P(X = 2) = \binom{2+3-1}{2} \cdot 0,3^2 \cdot 0,7^3 = 0,1852$

b. *¿Y la de que juegue solo 3 partidas?*
$P(X = 0) = \binom{0+3-1}{0} \cdot 0,3^0 \cdot 0,7^3 = 0,343$

c. *¿Cuál es el número esperado de partidas que jugará para clasificarse?*
El número esperado de partidas que perderá para clasificarse es:
$E(X) = 3 \cdot \frac{0,3}{0,7} = 1,2857$
Por lo que el número esperado de partidas que jugará para clasificarse es 4,2857.

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución Hipergeométrica – Hipergeométrica(N,n,p)

- Modela la probabilidad de obtener un número especifico de éxitos en una muestra aleatoria sin reemplazo, de un conjunto finito de elementos que contiene tanto éxitos como fracasos.

Consideremos una población con **N elementos de dos clases A y A* excluyentes**, de los cuales n^A son de la clase A y n^{A^*} de la clase A*, con $n^A+n^{A^*}=N$.
Al tomar un elemento de esta población, la probabilidad que proceda de una de las clases será n^A / N
Sea ahora el experimento consistente **en tomar n elementos consecutivos** de esta población **sin remplazamiento**.
Nuestro interés radica en **calcular la probabilidad de obtener exactamente X=k elementos de la clase A en las n extracciones sin remplazo**.

Ejemplo:
Una urna con N bolas de 2 colores, N1 de color blanco y N2 de color negro, de forma que N = N1 + N2. De dicha urna se extraen n bolas sin reemplazamiento y estamos interesados en el número de bolas blancas extraídas.

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución Hipergeométrica – Hipergeométrica(N,n,p)

- El ejemplo es equivalente a tener una población con N individuos, de los que N1 cumplen una determinada característica y el resto no.
- De esta población se extrae mediante muestreo aleatorio simple sin reemplazamiento una muestra de tamaño n y nos interesa saber el número de individuos que tienen la citada característica dentro de la muestra.

- La distribución hipergeométrica queda definida mediante sus parámetros:

  N: Tamaño de la población.

  n: Tamaño de la muestra.

  p: Probabilidad de cumplir la característica que nos interesa $p = \frac{N_1}{N}$

- Notación: $X \sim Hipergeometrica(N, n, p)$

  $X \sim Hipergeometrica(N, N_1, n)$
- Campo de variación: $X = \{0, 1, 2, \ldots n\}$
- Función de cuantía: $\boxed{P(X = x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}}$

- Esperanza: $\boxed{\mathbb{E}(X) = n \cdot p}$
- Varianza: $\boxed{\mathbb{V}\mathrm{ar}(X) = n \cdot p \cdot q \cdot \frac{N-n}{N-1}}$

$$E(X)=np=\frac{nn_A}{N};\ V(X)=npq\frac{N-n}{N-1}=\frac{nn_An_{A^*}(N-n)}{N^2(N-1)};$$

$$\sigma=\sqrt{npq\frac{N-n}{N-1}}=\sqrt{\frac{nn_An_{A^*}(N-n)}{N^2(N-1)}}$$

DISTRIBUCIONES DISCRETAS DE PROBABILIDAD

Distribución Hipergeométrica – Hipergeométrica(N,n,p)

Se dice que una v.a. X sigue una ley *hipergeométrica* de parámetros N, n y p cuando su función de masa sea:

$$P(X=k)=\frac{\binom{n_A}{k}\binom{n_{A^*}}{n-k}}{\binom{N}{n}}=\frac{\binom{N_p}{k}\binom{N_q}{n-k}}{\binom{N}{n}}$$

N elementos de dos clases A y A* excluyentes, de los cuales n^A son de la clase A y n^{A^*} de la clase A*, con $n^A+n^{A^*}=N$.

Al tomar un elemento de esta población, la probabilidad que proceda de una de las clases será n^A / N

Sea ahora el experimento consistente **en tomar n elementos consecutivos** de esta población **sin remplazamiento**.

interés radica en **calcular la probabilidad de obtener exactamente X=k elementos de la clase A en las n extracciones sin remplazamiento**

•Su diferencia con la distribución binomial es que, en aquélla, las probabilidades permanecían constantes a lo largo de todas las pruebas mientras que en la distribución hipergeométrica, **las probabilidades varían de una prueba a otra (extracciones sin remplazamiento).**

•Sin embargo, **si N es grande respecto a n**, las probabilidades varían muy poco de una prueba a la siguiente, por lo que en estos casos se puede decir que la variable hipergeométrica sigue aproximadamente una distribución binomial.

n/N <0,1

EJEMPLO HIPERGEOMÉTRICA

Una empresa de 100 trabajadores sabe que el 30% de su personal habla inglés y el resto no. Se selecciona aleatoriamente un grupo de 10 trabajadores para realizar un viaje a Londres.

Se pide:

a) Probabilidad de que 4 de los 10 seleccionados hablen inglés.

b) Probabilidad de que al menos 2 de los 10 seleccionados hablen inglés.

c) Número esperado de individuos que hablarán inglés en el grupo de viaje.

Solución

X=Número de trabajadores seleccionados para el viaje que habla inglés.

$X \sim Hipergeometrica(100; 10; 0,3)$

- a) *Probabilidad de que 4 de los 10 seleccionados hablen inglés.*

$$N_1 = p \cdot N = 0,3 \cdot 100 = 30$$
$$N_2 = N - N_1 = 100 - 30 = 70$$
$$P(X = 4) = \frac{\binom{30}{4}\binom{70}{10-4}}{\binom{100}{10}} = 0,2076$$

- b) *Probabilidad de que al menos 2 de los 10 seleccionados hablen inglés.*

$$P(X \geq 2) = 1 - P(X < 2) = 1 - P(X = 0) - P(X = 1) =$$
$$1 - \frac{\binom{30}{0}\binom{70}{10-0}}{\binom{100}{10}} - \frac{\binom{30}{1}\binom{70}{10-1}}{\binom{100}{10}} = 1 - 0,0229 - 0,1127 = 0,8644$$

- Notación: $X \sim Hipergeometrica(N, n, p)$
  $X \sim Hipergeometrica(N, N_1, n)$
- Campo de variación: $X = \{0, 1, 2, \ldots n\}$
- Función de cuantía: $\boxed{P(X = x) = \frac{\binom{N_1}{x}\binom{N_2}{n-x}}{\binom{N}{n}}}$
- Esperanza: $\boxed{\mathbb{E}(X) = n \cdot p}$
- Varianza: $\boxed{\mathbb{V}ar(X) = n \cdot p \cdot q \cdot \frac{N-n}{N-1}}$

- c) *Número esperado de individuos que hablarán inglés en el grupo de viaje.*

$$E(X) = 10 \cdot 0,3 = 3$$
$$V(X) = 10 \cdot 0,3 \cdot 0,7 \cdot \frac{100-10}{100-1} = 1,9$$

EJEMPLO HIPERGEOMÉTRICA

Un fabricante de chips los empaqueta en lotes de 25. El comprador los inspecciona tomando una muestra de 3 chips y acepta el lote si menos de 2 son defectuosos.

- Calcular la probabilidad de que el comprador acepte un lote con 6 chips defectuosos
- ¿Cuál es el número esperado y la varianza en los 3 chips inspeccionados?

(a) Se tiene un conjunto de $N = 25$ chips, en los que hay $N_1 = 6$ defectuosos, y $N - N_1 = 19$ no defectuosos. Extraemos $n = 3$ sin reemplazamiento. Consideramos la variable aleatoria $X =$ "número de chips defectuosos en los 3 seleccionados",

$$X \sim \mathcal{H}(N = 25, N_1 = 6, n = 3) \; .$$

La probabilidad de que el comprador acepte el lote es:

$$P[X < 2] = P[X = 0] + P[X = 1] = \frac{\binom{6}{0}\binom{19}{3}}{\binom{25}{3}} + \frac{\binom{6}{1}\binom{19}{2}}{\binom{25}{3}} = 0.8674,$$

$$\begin{aligned} E[X] &= 3 \times \frac{6}{25} = 0.72 \; , \\ Var[X] &= 3 \times \left(\frac{6}{25}\right) \times \left(\frac{19}{25}\right) \times \left(\frac{22}{24}\right) = 0.5016 \; . \end{aligned}$$

EJEMPLO HIPERGEOMÉTRICA

Supongamos que un distribuidor recibe de un fabricante un producto que viene en cajas de 1000 unidades. El fabricante afirma que su proceso de fabricación garantiza que cada lote lleva a lo sumo un 5% de piezas defectuosas. Para comprobar esta afirmación el recepcionista abre una caja y examina aleatoriamente tres unidades. ¿Cuál es la probabilidad de que al menos una de las tres piezas sea defectuosa, supuesto que la afirmación del fabricante es cierta?

$$P(X\geq 1)=1-P(X=0)=1-\frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$$

$$=1-\frac{\binom{50}{0}\binom{950}{3}}{\binom{1000}{3}}=0.14276$$

Podríamos aproximar esta probabilidad por medio de una binomial, ya que n/N=3/1000=0.003

$$P(X\geq 1)=1-P(X=0)=1-\binom{3}{0}(0.05)^0(0.95)^3=1-0.857375=0.142625$$

Tabla resumen de las distribuciones discretas

Distribución	Valores	Parámetros	Definición de la variable	Observaciones
Uniforme discreta	a, a+1, a+2, …,b	a: mínimo b: máximo	Variable que puede tomar n valores distintos con la misma probabilidad cada uno de ellos	
Binomial	0, 1, 2, …, n	n: número de pruebas p: probabilidad de éxito	Número de éxitos en n pruebas independientes de un experimento con probabilidad de éxito constante	Esta distribución se aplica a poblaciones finitas cuando los elementos se toman al azar y con reemplazo, y a poblaciones conceptualmente infinitas cuando el proceso es estable y sin
Multinomial	X_i: 0, 1, 2, … (i= 1, …, m)	n: número de pruebas m: nº de resultados posibles p_i: probabilidad del suceso i	Número de veces que ocurren m sucesos disjuntos en n pruebas independientes	Se aplica cuando se tiene un proceso estable y sin memoria
Hipergeométrica	de max{0,n-(N-R)} a min{R,n}	N: tamaño de la población R: número de éxitos n: número de pruebas	Número de éxitos en una muestra de tamaño n, extraída sin reemplazo de una población de tamaño N que contiene R éxitos	Es equivalente a la distribución binomial cuando el muestreo se hace sin reemplazo. Si el tamaño de la población es grande ambas distribuciones se pueden considerar prácticamente iguales
Geométrica	0, 1, 2, …	p: probabilidad de éxito	Número de fracasos antes de obtener un éxito por primera vez	Se utiliza en la distribución de tiempos de espera y tiene la propiedad de "falta de memoria"
Binomial negativa	0, 1, 2, …	r: número de éxitos p: probabilidad de éxito	Número de fracasos antes de obtener el r-ésimo éxito	Cuando r=1 se obtiene la distribución geométrica
Pascal	r, r+1, r+2, …	r: número de éxitos p: probabilidad de éxito	Número de pruebas necesarias para obtener r éxitos	Se relaciona con la binomial negativa de la siguiente manera: Pascal(r,p)=BN(r,p)+r
Poisson	0, 1, 2, …	λ: tasa de ocurrencia	Número de ocurrencias de un evento "raro" o poco frecuente en un intervalo o espacio continuo de tiempo	El proceso que genera una distribución de Poisson es estable y no tiene memoria. La distribución binomial se aproxima por la Poisson si n es grande y p pequeña, siendo λ=np

VARIABLES ALEATORIAS CONTINUAS

ARGUMENTO

- **Las v.a. continuas** son aquellas que **pueden tomar un número infinito de valores** dentro de un intervalo específico o en todo el conjunto de números reales. A diferencia de **las v.a. discretas**, que solo **pueden tomar valores específicos y aislados**, las v.a. continuas pueden adoptar cualquier valor dentro de un rango dado.

- Por ejemplo, la altura de una persona, el tiempo que tarda en completarse una tarea o la temperatura ambiente son ejemplos de variables aleatorias continuas. Estas variables pueden tomar valores en un rango continuo y no se limitan a valores discretos específicos.

- Para describir completamente una variable aleatoria continua, se utiliza una **función de densidad de probabilidad (PDF)**, que asigna una densidad de probabilidad a cada valor posible de la variable en el rango dado. La probabilidad de que la variable aleatoria tome un valor dentro de un intervalo específico se calcula **integrando la función de densidad de probabilidad sobre ese intervalo**.

VARIABLES ALEATORIAS CONTINUAS

Características

- Una variable aleatoria continua X, es una función real definida en el espacio de probabilidad asociado a un experimento aleatorio. Medida numérica que se obtiene como resultado de un proceso aleatorio:

$$X : \Omega \rightarrow \mathbf{R}$$

Para cada elemento muestral en el espacio muestral Ω (omega), la variable aleatoria X asigna un valor real.

- Se dice que una variable aleatoria es **continua** si su función de distribución, **F(x)**, es continua y su primera derivada **existe y es continua**.

- El conjunto de puntos donde se verifican estas condiciones constituye el **campo de variación** de la variable aleatoria.

- En la práctica utilizaremos modelos de probabilidad continuos para modelizar variables que presenten campos de variación con gran variabilidad.

- Ejemplos.-
  - Los ingresos de las diferentes empresas de un sector.
  - Los tiempos de computación de un proceso de carga de datos.
  - En videojuegos: salud de un personaje, velocidad de movimiento, poder de ataque, nivel de experiencia…

VARIABLES ALEATORIAS CONTINUAS

Función de distribución

- En el contexto de una variable aleatoria continua *X*, **la función de distribución acumulativa *F*(*x*) proporciona la probabilidad de que la variable aleatoria *X* sea menor o igual a un valor dado *x***. En otras palabras, es la probabilidad acumulada hasta ese valor *x*. La función de distribución *F(x)* puede expresarse como:

$$\boxed{P(X \leq x) = F(x) = \int_{-\infty}^{x} f(x)\mathrm{d}x}$$

Donde la integral definida de la **función de densidad de probabilidad *f*(*x*) desde menos infinito hasta *x* representa el área bajo la curva de la función de densidad de probabilidad hasta el punto *x***.

- **Probabilidad acumulada:** La probabilidad de que la variable aleatoria X sea menor o igual a x se puede expresar como la integral definida de la función de densidad de probabilidad f(x) desde menos infinito hasta x.

- **Propiedades**
  - Como es una probabilidad, se cumple que **$0 \leq F(x) \leq 1$**.
  - Al ser **acumulativa** nunca es decreciente, es decir, sí consideramos dos valores específicos de la **variable *x*, *a*** y ***b*** entonces **$F(a) \leq F(b)$** es porque **$a < b$**.
  - ***F* (−∞) = 0 y *F* (+∞) = 1**

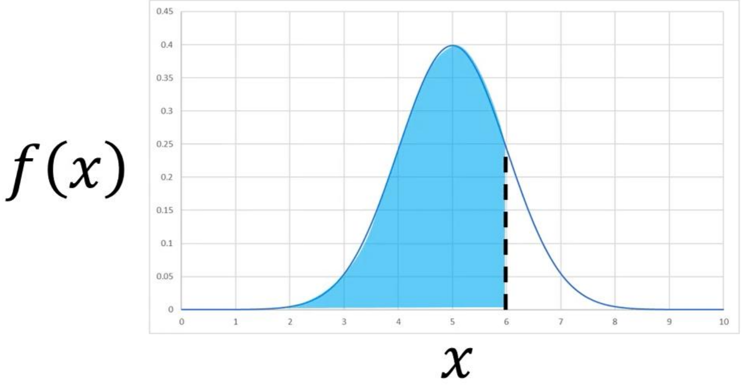
$$P(a < X \leq b) = F(b) - F(a) = \int_{-\infty}^{b} f(x)\mathrm{d}x - \int_{-\infty}^{a} f(x)\mathrm{d}x = \int_{a}^{b} f(x)\mathrm{d}x$$

VARIABLES ALEATORIAS CONTINUAS

Función de distribución

Definición formal:
Sea X una variable aleatoria continua con función de densidad de probabilidad f(x). La función de distribución acumulativa de X es la función:

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



De forma gráfica, *F(x)* es el área bajo la curva de densidad a la izquierda de *x*. **Recordemos que cuando trabajamos con la función de densidad, área es probabilidad.**

Además, tenemos algunas **fórmulas interesantes** que nos servirán para resolver los problemas.

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

$$P(a \leq X \leq b) = F(b) - F(a) \quad ; \quad a < b$$

La siguiente fórmula nos permite pasar de la función de distribución acumulativa a la función de densidad de probabilidad:

$$f(x) = \frac{dF(x)}{dx}$$

Al ser una función acumulativa, esta no puede ser decreciente.

VARIABLES ALEATORIAS CONTINUAS

Función de densidad

- A la derivada de la función de distribución F (x), F ′(x), la denominamos función de densidad y se representa por f (x).

- En términos de probabilidad:

$$P\left(a < X < b\right) = \int_{a}^{b} f\left(x\right)\mathrm{d}x$$

representa la probabilidad de que la variable aleatoria X tome un valor entre a y b.

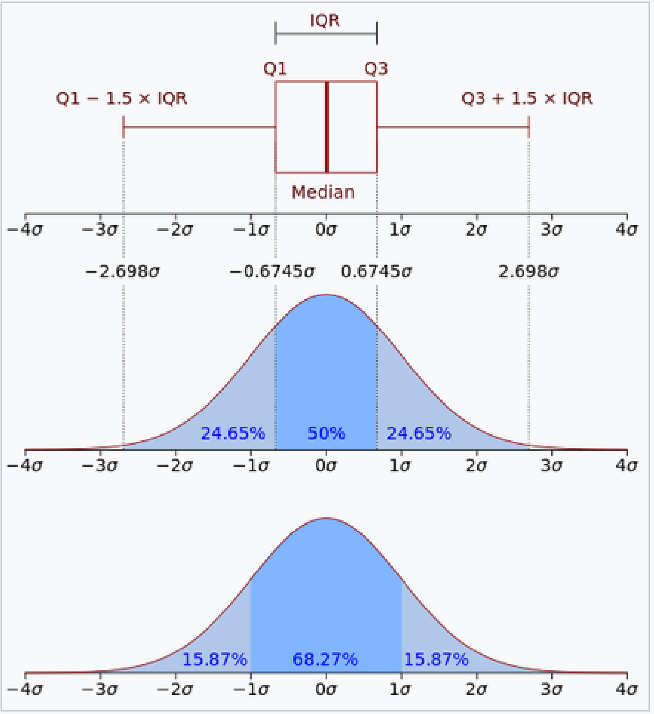


Diagrama de Caja y función de densidad de probabilidad de una distribución normal $N(0, \sigma^2)$.

- Hay que tener cuenta que f (x), al contrario que F (x), no es una probabilidad, por lo cual puede ser mayor que uno, e incluso infinito, pero nunca negativa.

VARIABLES ALEATORIAS CONTINUAS

Ejemplo

Dada la función de distribución: $F(x) = 2x - x^2$ en [0, 1].

(a) Calcular la función de densidad.
(b) Comprobar que el área bajo la función de densidad es 1.
(c) Hallar la probabilidad del suceso $\{X \leq 0.2\}$ mediante la función de densidad y la función de distribución.

Solución

(a)

$$f(x) = F'(x) = 2 - 2x \qquad en \ [0,1]$$

(b)

$$\int_0^1 (2-2x)\,dx = 2x - x^2\Big|_0^1 = (2\cdot 1 - 1^2) - (2\cdot 0 - 0^2) = 1$$

$$f(x) = \frac{dF(x)}{dx}$$

$$\boxed{P(a < X < b) = \int_a^b f(x)dx}$$

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

(c)

$$P(X \leq 0.2) = \int_0^{0.2} (2-2x)\,dx = 2x - x^2\Big|_0^{0.2} = $$
$$= (2\cdot 0.2 - 0.2^2) - (2\cdot 0 - 0^2) = 0.36$$

$$P(X \leq 0.2) = F(0.2) = 2\cdot 0.2 - 0.2^2 = 0.36$$

VARIABLES ALEATORIAS CONTINUAS

Propriedades

- En el cálculo se utiliza la **función de densidad** más habitualmente que la de distribución, pues proporciona sobre la variable aleatoria una información **más directa y manejable**.

- Comprobamos lo dicho con los siguientes **ejemplos:**

  - Si la función de densidad es **paralela al eje horizontal**, sea cual sea la posición de un intervalo arbitrario, siempre la probabilidad es la **misma**.

  - Si es una **recta decreciente**, la variable aleatoria tiene una probabilidad más alta de presentar valores pequeños que grandes.

  - Si es una **recta creciente**, la variable aleatoria tiene una probabilidad más alta de presentar valores grandes que pequeños.

  - En una función de **forma de U** es más fácil encontrar valores pequeños o grandes que intermedios.

  - Si la función de densidad tiene la **forma de ∩** la situación es la inversa.

- No frecuente, pero es **posible** encontrar variables aleatorias con distribuciones mixtas, es decir, con campo de variación donde en una parte de él la variable es continua y en otra discreta.

VARIABLES ALEATORIAS CONTINUAS

Ejercicio propuesto

Considere la variable aleatoria con función de densidad:

$$f(x) = k \cdot x^3 e^{-x^2/2}, \quad para\ x \in [0, \infty)$$

Se pide:

- Calcule el valor de k para que sea realmente una función de densidad.
- Calcule la función de distribución.
- Calcule las siguientes probabilidades: $P(1 < X \leq 3)$, $P(X < 2)$ y $P(-5 < X < 5)$.

VARIABLES ALEATORIAS CONTINUAS

Esperanza

La **esperanza** matemática en distribuciones continuas se calcula del siguiente modo:

$$\boxed{\mathbb{E}(X) = \int_{-\infty}^{\infty} x\, f(x)\, \mathrm{d}x}$$

Ejemplo

Calcule la esperanza de la variable continua definida por:

$$f(x) = 4x^3 \qquad \text{para } 0 \leq x \leq 1$$

Solución

$$\mathbb{E}(X) = \int_0^1 x \cdot 4x^3 \mathrm{d}x = \left.\frac{4x^5}{5}\right|_0^1 = 4\left[\frac{1^5}{5} - \frac{0^5}{5}\right] = \frac{4}{5}$$

VARIABLES ALEATORIAS CONTINUAS

Varianza

La **varianza** se calcula del siguiente modo:

$$\boxed{\mathbb{V}\mathrm{ar}\left(X\right) = \sigma^2 = \mu_2 = \mathbb{E}\left[\left(X-\mu\right)^2\right] = \mathbb{E}\left(X^2\right) - \mu^2}$$

donde

$$\mathbb{E}\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f\left(x\right) \mathrm{d}x$$

Ejemplo

Calcule la varianza de la variable continua definida por:

$$f\left(x\right) = 4x^3 \qquad \textit{para } 0 \leq x \leq 1$$

Solución

$$\begin{aligned} \mathbb{E}\left(X^2\right) &= \int_0^1 x^2 \cdot 4x^3 \mathrm{d}x = \left.\frac{4x^6}{6}\right|_0^1 = 4\left[\frac{1^6}{6} - \frac{0^6}{6}\right] = \frac{2}{3} \\ \mathbb{V}\mathrm{ar}\left(X\right) &= \frac{2}{3} - \left(\frac{4}{5}\right)^2 = 0.0267 \end{aligned}$$

VARIABLES ALEATORIAS CONTINUAS

Ejercicio

Considere la variable aleatoria con función de densidad:

$$f(x) = kx^2 - 3x + 1, \quad para\ x \in [0, 1]$$

Se pide:

- Calcule el valor de k para que sea realmente una función de densidad.
- Calcule la función de distribución.
- Calcule las siguientes probabilidades: $P(0,2 < X \leq 0,7)$, $P(X < -0,5)$ y $P(0,7 < X < 1,5)$.
- Calcule la esperanza y la varianza.

DISTRIBUCIONES DE PROBABILIDAD CONTINUAS

DESCRIPCIÓN

1. Las distribuciones de probabilidad continuas son **modelos matemáticos que describen la probabilidad de ocurrencia de eventos dentro de un rango continuo de valores reales**. Estas distribuciones son esenciales en la teoría de la probabilidad y se utilizan ampliamente en diversas disciplinas.

2. Es cierto que, a diferencia de las distribuciones de probabilidad discretas, **las distribuciones de probabilidad continuas tienen una gama infinita de posibles resultados**, lo que **permite que las variables** aleatorias asociadas con estas distribuciones **tomen cualquier valor dentro de un rango continuo de números reales**.

3. **La función de densidad de probabilidad (PDF)** es una característica fundamental de las distribuciones de probabilidad continuas. **Describe la probabilidad relativa de que una variable aleatoria tome un valor en un intervalo dado**. **La integral de la función de densidad de probabilidad sobre un intervalo específico proporciona la probabilidad de que la variable aleatoria caiga dentro de ese intervalo.**

4. Las distribuciones de probabilidad continuas que vamos a estudiar en este tema serán:

   - **Distribución Uniforme**
   - **Distribución Exponencial**
   - **Distribución Normal**

   Cada una de ellas se utiliza para modelar diferentes tipos de fenómenos y tienen propiedades distintivas que las hacen útiles en diferentes contextos.