Variables Aleatorias y Funciones de Densidad

Dr. Delfino Vargas Chanes

Facultad de Economía
Universidad Nacional Autónoma de México

13 de marzo de 2024

Índice

Funciones

Definición 1.1

Sean X, Y conjuntos cualesquiera. Una función f de X a Y es un objeto que, $\forall\, x \in X$ se le asigna un **único** elemento $y \in Y$. Denotamos estos objetos como $f : X \to Y$. Llamamos a X el dominio de f, mientras que Y es el codominio de f.

- Decimos que una función $f : X \to Y$ está **bien definida** si cumple con lo siguiente
  - Totalidad: $\forall\, x \in X\ \exists\, y$ tal que $f(x) = y$
  - Existencia: $\forall\, x \in X \Rightarrow f(x) \in Y$
  - Unicidad: $\forall\, x \in X$ solo hay una $y \in Y$ tal que $f(x) = y$
- Ejemplos de funciones
  - $f : \mathbb{Z} \to \mathbb{Z}$ definida por $f(x) = 2x$
  - $F : \mathbb{R}^3 \to \mathbb{R}$ definida como $F(K, N) = AK^{\alpha}L^{1-\alpha}$
  - $\phi : \mathbb{R} \to \mathbb{R}$ definida como $\phi(x) = x^2 + x + 3$

Propiedades

Definición 1.2

Sean X, Y conjuntos cualesquiera y $f : X \to Y$ una función. Decimos que es inyectiva si $\forall\, x, y \in X$ si $f(x) = f(y) \Rightarrow x = y$

Definición 1.3

Sean X, Y conjuntos cualesquiera y $f : X \to Y$ una función. Decimos que es suprayectiva si $\forall\, y \in Y\ \exists\, x \in X$ tal que $f(x) = y$

Definición 1.4

Sean X, Y conjuntos cualesquiera y $f : X \to Y$ una función. Decimos que es biyectiva si es suprayectiva e inyectiva.

Variable Aleatoria

Definición 2.1

Dado un experimento en un espacio muestral Ω, una *variable aleatoria* X es una función que va de Ω a $\mathbb{R}$, es decir

$$X : \Omega \to \mathbb{R}$$

- Estaremos refiriéndonos a las variables aleatorias como VA.
- Una variable aleatoria X es una función que asocia un valor numérico $X(s)$ a cada resultado posible de un experimento aleatorio $s \in \Omega$.
- En un espacio muestral $\Omega : \{AA, AS, SA, SS\}$, las siguientes son VA:
  - El número de Águilas es una VA con valores $0, 1, 2$.
  - Y es una VA que es 1 si el primer lanzamiento es Águila y 0 si no.
  - En ese caso se le asigna a los eventos AA y AS el valor 1.

Variable Aleatoria

- A continuación, más ejemplos de VA.
- Escojamos una persona al azar del salón, y sea el color de sus ojos X.
- El número de embarazos después del uso correcto de anticonceptivos.
- La cantidad de bacteria en un estudio de crecimiento bacterial.
- El tiempo en el que Tadej Pogačar completará el Tour de France 2024.
- El valor del IPC México (MXX) al cierre del mercado mañana.
- La fuente de la aleatoriedad de la variable es el experimento mismo, en el que de un evento $s \in \Omega$ se escoge mediante una función de probabilidad P. Definiremos estas mas adelante.
- Antes del experimento, no sabemos cual es la VA, solo podemos calcular su probabilidad. Después del experimento, si la VA es X entonces su valor número para un experimento s es $X(s)$.

VA Discreta

Definición 2.2

Decimos que una variable aleatoria X es discreta si hay una lista finita de valores $a_1, a_2, ...a_n$ o una lista infinita numerable $a_1, a_2, ...$ tales que

$$\sum_i P(X = a_i \text{ para alguna } i = 1, 2, ...) = 1$$

Es decir, los valores constituyen un conjunto numerable, sea finito o infinito.

Función de Masa de Probabilidad

- Hemos afirmado que una variable aleatoria discreta puede tomar in número infinito de valores.
- En la práctica podemos usar los conceptos de probabilidad para poder establecer las reglas de asignación de los valores que toman las VA.
- Para ello adoptamos el concepto de función de distribución de probabilidad.

Función de Masa de Probabilidad

Definición 2.3

La función de masa de probabilidad (PMF) de una VA discreta es la función p_X dada por $p_X(X) = P(X = x)$. Formalmente:

$$p(X = x) = P\left(\{s \in \Omega \mid X(s) = x\}\right) = P(X^{-1}(x))$$

- Al escribir $P(X = x)$ estamos utilizando $X = x$ para denotar un evento, que consiste en todos los resultados $s \in \Omega$ para los cuales la VA X asigna el valor $x \in \mathbb{R}$.

Teorema 2.1

Sea X una VA discreta. Su PMF p_X cumple las siguientes propiedades:

1. $\forall\, j \in 1, \ldots \Rightarrow p_X(x_j) > 0$. Las VA X con probabilidad 0 no se enlistan.
2. $\sum\limits_{i=1}^{\infty} p_X(x_i) = 1$

Demostración.

La primera es trivial por la no negatividad de la probabilidad. Para la segunda, se tiene que:

$$\sum_{i=1}^{\infty} P(X = x_i) = P\left(\bigcup_{i=1}^{\infty}\{X = x_i\}\right) = P(X = x_1 \text{ o } X = x_2 \text{ o } \ldots) = 1$$

$\square$

Ejemplos

- La implicación de estas propiedades con un ejemplo de tres lanzamientos de una moneda, con X la VA observar si cae águila.
- Estos valores se pueden reordenar en una tabla más sencilla, ya que de ocho lanzamientos uno no cae águila.
- Por ejemplo se tiene $P(X=0)=\frac{1}{8}$ o $P(X=2)=\frac{3}{8}$

Valor de $X=x$	$p_X(x)$
0	$\frac{1}{8}$
1	$\frac{1}{8}$
2	$\frac{3}{8}$
3	$\frac{1}{8}$
Total	1

Introducción

- En esta parte se revisan las distribuciones básicas donde la variable de respuesta es binaria (dos categorías).
- Un modelo de probabilidad es una forma que modela la distribución de probabilidad y el comportamiento de una variable aleatoria. Las probabilidades se expresan en términos de cantidades relevantes llamadas parámetros.
- En esta parte también introducimos el concepto de pruebas de hipótesis. Este concepto será crucial para el resto del curso.

Éxito y Fracaso

- A continuación, ejemplos de donde los elementos de la población presentan una dicotomía (respuesta binaria):
- Inspección de varios artículos que provienen de una línea de producción y se cuenta el número de artículos defectuosos.
- Levantar una encuesta de votantes y observar cuántos de ellos votan a favor de una propuesta.
- Analizar los especímenes de sangre tomados de ratas y contar cuántas de ellas son portadoras de cierta enfermedad.
- Examinar historias de casos de un número de nacimientos y contar cuantos fueron inducidos por cesária

Distribución Bernoulli

Definición 3.1

Se dice que una VA X tiene distribución Bernoulli con parámetro p, si $P(X = 1) = p$ y $P(X = 0) = 1 - p$, donde $p \in (0, 1)$. Se escribe como $X \sim \text{Bern}(p)$. El símbolo $\sim$ se lee "se distribuye como"

- Cualquier VA con valores posibles 0 y 1, tiene una distribución $\text{Bern}(p)$, con un parámetro p asociado.
- Este número p determina que distribución en particular se tiene.
- No solo hay una distribución Bernoulli, pero toda una familia.
- Cualquier evento en un espacio muestral cualquiera tiene una VA asociada a el, igual a 1 si sucede y a 0 si no sucede.

Variable Indicatriz

Definición 3.2

La VA indicatriz de un evento A es la VA que es igual a 1 si ocurre, y a 0 si no ocurre. Denotamos la VA indicatriz de A como I_A. Notemos que, $I_A \sim \text{Bern}(p)$ con $p = P(A)$.

Ensayos Bernoulli

Definición 3.3

Un experimento que resulta o en un *éxito* o un *fracaso*, pero no ambas, se llama un ensayo Bernoulli. Una VA Bernoulli es en realidad un el indicador del éxito en un ensayo Bernoulli, es 1 si hay éxito, y 0 si hay fracaso. Denotamos entonces al parámetro p como la probabilidad de éxito de una distribución Bern(p).

Distribución Binomial

Definición 3.4

Suponga que se realizan n ensayos Bernoulli independientes, cada uno con la misma probabilidad p asociada. Sea X una VA que denota el número de éxitos. La distribución de X se llama distribución binomial, de parámetros n y p. Se denota como $X \sim \mathrm{Bin}(n, p)$, donde $n \in \mathbb{Z}^{+}$ y $p \in (0, 1)$.

Distribución Binomial

Teorema 3.1

Si $X \sim \mathrm{Bin}(n, p) \Rightarrow$ la PDF de X es

$$p_X = P(X = x) = \binom{n}{k} p^k (1-p)^{n-k}$$

para $k = 0, 1, ..., n$, con $P(X = k) = 0$ de lo contrario.

Demostración de la PDF

Demostración.

Cuando uno ha obtenido k éxitos en n ensayos, también se obtienen $(n-k)$ fracasos. La probabilidad de k éxitos y $(n-k)$ fracasos cuando el órden importa es

$$p^k(1-p)^{n-k}$$

Cuando el órden no importa, también podemos combinar k éxitos y $(n-k)$ fracasos. Este número es el número de posibilidad de encontrar k objetos de n objetos, o sea, el coeficiente binomial

$$\binom{n}{k}$$

Debemos multiplicar estas dos expresiones para obtener la probabilidad de k éxitos y $(n-k)$ fracasos. Pero esto es la PDF. $\square$

PDF Binomial

- Esta PDF es válida porque cumple con las propiedades mencionadas.
- Es evidente que es no negativa.
- En cuanto a que su suma sea igual a 1, veamos que

$$1 = 1^n = (p + 1 - p)^n = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

- Esto por el teorema del binomio, que postula

$$(x + y)^n = \sum_{k=0}^{n} \binom{n}{k} y^k \cdot x^{n-k}$$

Teorema de Binomial

Teorema 3.2

Sea $X \sim \mathrm{Bin}(n, p)$ y definamos a $q = 1 - p \Rightarrow n - X \sim \mathrm{Bin}(n, q)$

Demostración.

Veamos que $n - X$ tiene la PDF de la binomial. Sea $Y = n - x$

$$P(Y = k) = P(X = n - k) = \binom{n}{n-k} p^{n-k}(q)^k = \binom{n}{k} q^k (p)^{n-k}$$

$\square$

Ejemplos Binomial I

- De acuerdo a las leyes de Mendell si tenemos que si hay solo plantas rojas y blancas la heredabilidad es del 25 % para las flores rojas.
- Suponga que se cruzan 5 pares de flores rojas y blancas. Entonces ¿Cuál es la probabilidad de lo siguiente?
  - ¿No haya plantas rojas?
  - ¿Haya 4 o más plantas rojas?
- Dado que los ensayos ocurren con diferentes plantas (orígenes diferentes), entonces se cumple la independencia.
- Sea X la variable aleatoria que denota el evento de las plantas rojas de entre las 5 posibles.
- Si el evento E se define como obtener plantas rojas. Entonces la teoría Mendeliana especifica que $P(E) = .25$ y tiene una distribución $X \sim \mathrm{Bin}(n = 5, p = .25)$.

Ejemplos Binomial I

- Entonces se requiere calcular

$$P(X = 0) = \binom{5}{0} p^0 (1-p)^{5-0} = 0.75^5 = 0.237$$

$$P(X \geqslant 4) = \binom{5}{4} p^4 (1-p)^{5-4} + \binom{5}{5} p^5 (1-p)^0$$

- Realizando los cálcuos se tiene que

$$= 0.015 + 0.001 = 0.016$$

Ejemplos Binomial II

- Tiremos un dado. Sea X el número de veces que aparece el número 6 en los primeros 10 lanzamientos, y Y el número de lanzamientos necesarios para obtener un 3.
- En efecto, $X \sim \mathrm{Bin}(10, \frac{1}{6})$. Escribamos la función de masa de probabilidad de X, para algún $i = 1, ..., 10$

$$P(X = k) = \binom{10}{k} \binom{1}{6}^{k} 1 - \binom{1}{6}^{10-k}$$

- Veamos este ejemplo de Wikipedia
- Suponga una moneda con maña. La tranza es que la probabilidad que salga sol es de 0.3. La probabilidad de obtener 4 soles en 6 lances es

$$P(X = 4) = \binom{6}{4} 0.3^4 (1 - 0.3)^{6-4} = 0.059535$$

Función de distribución acumulada

Definición 3.5

Decimos que la función de distribución acumulada (CDF) de una VA llamada X cualquiera es la función F_X dada por $F_X(x) = P(X \leqslant x)$

Teorema 3.3

Si $X \sim \mathrm{Bin}(n, p) \Rightarrow$ la CDF de X es

$$F_X = P(X \leqslant x) = \sum_{i=0}^{\lfloor k \rfloor} \binom{n}{i} p^i (1-p)^{n-i}$$

CDF

- Las CDF deben de cumplir con las siguientes propiedades.
- Si $x \leqslant y \Rightarrow F_X(x) \leqslant F_X(y)$
- $\forall\ a \Rightarrow F_X(a) = \lim_{x \to a^+} F_X(x)$, la función es continua por la derecha.
- $\lim_{x \to -\infty} F_X(x) = 0$ y $\lim_{x \to \infty} F_X(x) = 1$

Binomial en R

- Existen tres funciones asociadas a la Binomial en RStudio.
- `dbinom` es la PDF Binomial. Toma en cuenta el valor de x, y los parámetros n, llamado size, y p, llamado prob.
- Por ejemplo, recordemos nuestro ejemplo de la moneda con maña

$$P(X=4)=\binom{6}{4}0.3^4(1-0.3)^{6-4}=0.059535$$

En este caso `dbinom(x=4, size=6, prob=0.3)`.

- `pbinom` es la CDF de la binomial.
- `rbinom` genera números que siguen una distribución binomial para los parámetros elegidos. En las tres son un valor, size y prob. Sin embargo, para esta, se tiene que es cuantas VA queremos generar.

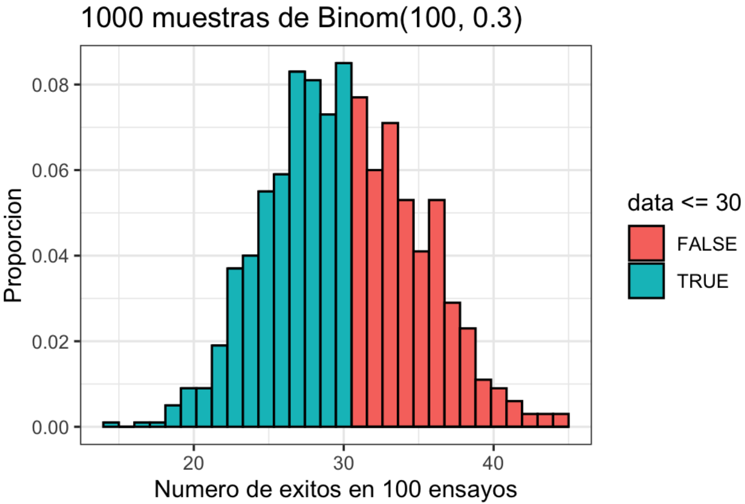
Cálculo de dbinom, pbinom y rbinom en R

```
dbinom(x=4, size=6, prob=0.3)
[1] 0.059535

pbinom(30, 100, 0.3)
[1] 0.5491236

#Codigo para la grafica empieza aqua

set.seed(10)
rbinom(1, 100, 0.3)
binomial_data <- rbinom(1000, 100, 0.3)
binomial_data <- as.data.frame(binomial_data)
names(binomial_data) <- c
```

Cálculo de dbinom, pbinom y rbinom en R

```
pbinom(30, 100, 0.3)

binomial_data %>% ggplot() +
geom_histogram(aes(x = data,
                   y = stat(count / sum(count)),
                   fill = data <= 30),
               color = 'black') +
theme_bw() +
labs(x = 'Numero de exitos en 100 ensayos',
   y = 'Proporcion',
   title = '1000 muestras de Binom(100, 0.3)')
```

Gráfica de pbinom



Funciones de VA

- Note que las funciones de VA también son VA aleatorias. e.g., para X una VA aleatoria cualquiera, X^2, $\sin(x)$, $\log(X)$ también son VA.

Definición 3.6

Para un experimento cualquiera s en un espacio muestral Ω una VA X y una función $g : \mathbb{R} \rightarrow \mathbb{R} \Rightarrow g(X)$ es la VA que mapea o transforma s a

$$g(X(s))$$

$\forall\, s \in \Omega$

Independencia

Definición 3.7

Decimos que dos VA X y Y son independientes si

$$P(X \leqslant x, Y \leqslant y) = P(X \leqslant x)P(Y \leqslant y)$$

$\forall\, x, y \in \mathbb{R}$

Definición 3.8

Decimos que n VA $X_1, ..., X_n$ son independientes si

$$P(X_1 \leqslant x_1, ..., X_n \leqslant x_n) = P(X_1 \leqslant x_1)...P(X_n \leqslant x_n)$$

$\forall\, x_1, ..., x_n \in \mathbb{R}$

Valor Esperado

Definición 4.1

El valor esperado de una VA discreta X cuyos valores posibles son $x_1, x_2, \ldots$ se define como

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} x_i P(X = x_i)$$

- La definición de valor esperado es similar a la de la media ponderada.
- En este caso, los *pesos* son las probabilidades.
- Es decir, el valor esperado de una VA X es una media ponderada de los valores que puede tomar X, ponderado por sus probabilidades.

Ejemplos

- Sea X una VA con $P(X=1)=0.2$, $P(X=2)=0.3$, y $P(X=3)=0.5$. ¿Cuál es $\mathbb{E}[X]$?
- Utilicemos la definición anterior:

$$\sum_{i=1}^{3} x_i P(X=x_i) = 1\cdot 0.2 + 2\cdot 0.3 + 3\cdot 0.5 = 2.3$$

- Sea X el resultado de lanzar un dado sin maña. En este caso las x_i toman el valor de $1,2,3,4,5,6$, las caras del dado.
- Todas las x_i tienen la misma $P(X=x_i)=\frac{1}{6}$. ¿Cuál es $\mathbb{E}[X]$?

$$\sum_{i=1}^{6} x_i P(X=x_i) = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6}$$

$$= \frac{1}{6}(1+2+3+4+5+6) = 3.5$$

Linealidad de $\mathbb{E}$

Teorema 4.1

Sea X una VA cualquiera y $\lambda \in \mathbb{R} \Rightarrow$

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

y además

$$\mathbb{E}[\lambda \cdot X] = \lambda \cdot \mathbb{E}[X]$$

Valor Esperado de la Binomial

Teorema 4.2

Sea $X \sim \mathrm{Bin}(n, p)$ una VA. Entonces, $\mathbb{E}[X] = np$

Demostración.

$$\mathbb{E}[X] = \sum_{k=1}^{n} k \binom{n}{k} p^k q^{n-k}$$

$$= \sum_{k=1}^{n} n \binom{n-1}{k-1} p^k q^{n-k}$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} q^{(n-1)-(k-1)}$$

$$= np \cdot (p+q)^n = 1^n \cdot np = np$$

□

Varianza y Desviación Estándar

Definición 4.2

Sea $\mu = \mathbb{E}[X]$Definimos a la varianza de una VA X como

$$\mathrm{Var}[X] = \mathbb{E}[X - \mu]^2$$

La raíz cuadrada de la varianza se llama desviación estandar

$$\mathrm{SD}[X] = \sqrt{\mathrm{Var}[X]}$$

Varianza

Teorema 4.3

Para cualquier VA X

$$\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mu^2$$

Demostración.

$$\begin{aligned} \mathrm{Var}[X] &= \mathbb{E}[X-\mu]^2 = \mathbb{E}\left[X^2 - 2\mu X + \mu^2\right] \\ &= \mathbb{E}\left[X^2\right] - 2\mu\mathbb{E}[X] + \mu^2 = \mathbb{E}\left[X^2\right] - 2\mu^2 + \mu^2 \\ &= \mathbb{E}\left[X^2\right] - \mu^2 \end{aligned}$$

$\square$

Varianza

- La varianza cumple las siguientes propiedades:
- $\mathrm{Var}[X+\lambda] = \mathrm{Var}[X]$
- $\mathrm{Var}[\lambda \cdot X] = \lambda^2 \cdot \mathrm{Var}[X]$
- Si X y Y son independientes, $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$
- Las únicas VA con varianza cero son las constantes.

Variables Aleatorias Continuas

- Hasta ahora hemos estado trabajando con VA discretas, es decir, que son numerables y podemos enlistar.
- En esta parte, estaremos trabajando con variables aleatorias continuas.
- Estas VA, pueden tomar cualquier valor en $\mathbb{R}$ en un intervalo.
- Antes de definirlas, necesitamos la intuición de lo que es una integral.

Integral

- Al no ser un curso de cálculo, no daremos una definición formal.
- Sea $f : [a, b] \to \mathbb{R}$. Detrás de la integral, está la motivación de aproximar el área bajo la curva de la función $y = f(x)$, donde $a \leqslant x \leqslant b$.
- Tomamos N particiones del eje x en un intervalo de a a b dividiendolo entre ese número. $\Delta x = \frac{x_m}{N}$
- Estas particiones son la *base* del rectangulo. Para la altura, consideramos el valor de la función f, lo que se proyecta en el eje de las y.

$$\text{Área} \approx \sum_{i=1}^{N} f_i \Delta x$$

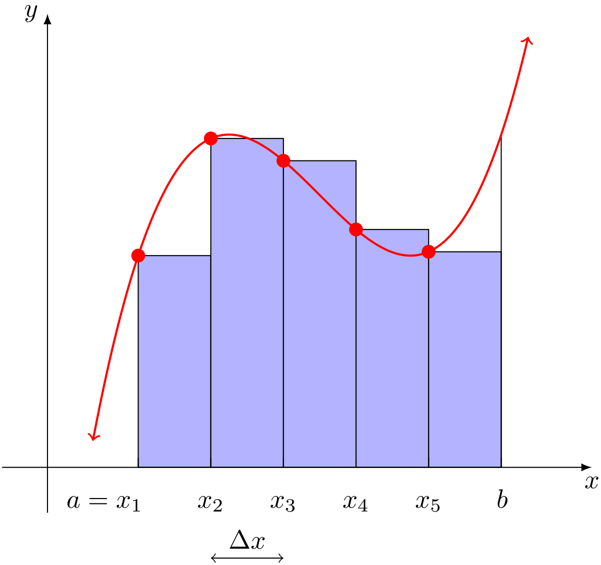
- Mientras más grande sea N, más rectángulos tenemos, y más se aproxima al valor del área bajo la curva.

Integral

- Notemos que cuando $N \to \infty$, la fracción $\Delta = \frac{x_m}{N} \to 0$
- La integral es cuando tomamos el límite cuando la base de los rectángulos, es decir $\Delta x \to 0$

$$\text{Área} = \int_a^b f(x)dx = \lim_{\Delta x \to 0} \sum_{i=1}^{N} f_i \Delta x$$

- Se *calcula* el área entre a y b, donde dx es el diferencial de x.
- Presentamos una animación para facilitar la intuición.


y
x
a = x1
x2
x3
x4
x5
b
Δx

VA continua

Definición 5.1

Decimos que la función de distribución acumulada (CDF) de una VA llamada X cualquiera es la función F_X dada por $F_X(x) = P(X \leqslant x)$

Definición 5.2

Decimos que una VA X es continua si su CDF es diferenciable.

VA Continua

Definición 5.3

Sea $F_X(x)$ la CDF de una VA continua X. Entonces, denotamos a $f(x)$ como

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

y le llamamos a $f(x)$ la PDF de la VA X.

- Notemos que podemos escribir a la CDF como

$$F_X(x) = \int_{-\infty}^{x} f(t)dt$$

VA Continua

Definición 5.4

Si una VA continua X tiene una PDF $f(x)$, y se tiene que $a < b$, entonces, la probabilidad de que X caiga en el intervalo $[a, b]$ es

$$P(a \leqslant x \leqslant b) = \int_a^b f(x)dx$$

Propiedades de la PDF

Teorema 5.1

Para que una PDF sea válida debe de cumplir lo siguiente:

$$f(x) \geqslant 0$$

No negativa, y se integra a 1.

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

Notemos que son las mismas que para una VA discreta, nada más que cambiamos la suma por la integral.

Valor Esperado de VA Continuas

Definición 5.5

El valor esperado de una variable aleatoria continua con PDF f es

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx$$

Como nota, la varianza en VA continuas se computa de la misma forma que en la parte anterior.

Distribución Normal

Definición 5.6

Decimos que una VA X tiene distribución normal con media μ y varianza σ^2 (es decir, de parámetros μ y σ)

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

si y solo si su PDF es la siguiente

$$f(x) = \mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$

Distribución Normal

Demostración.

Veamos que integra a 1 (Pendiente) $\square$

Distribución Normal

Teorema 5.2

Si X es una VA continua con distribución normal, entonces

$$\mathbb{E}[X] = \mu \qquad \text{y} \qquad \mathrm{Var}[X] = \sigma^2$$

La demostración de este teorema se realizará más adelante.

Distribución Normal Estandar

Definición 5.7

Decimos que una VA X tiene distribución normal estandarizada si sigue una distribución normal con media 0 y desviación estándar 1.

$$X \sim \mathcal{N}(0,1)$$

es decir, su PDF es

$$\varphi(x) = \mathcal{N}(x;0,1) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{x^2}{2}\right]$$

La CDF se denota como

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)dt$$

La variable estandarizada

- Si X es una variable aleatoria tal que $X \sim \mathcal{N}(\mu, \sigma^2)$, entonces la variable aleatoria Z se llama variable aleatoria estandarizada

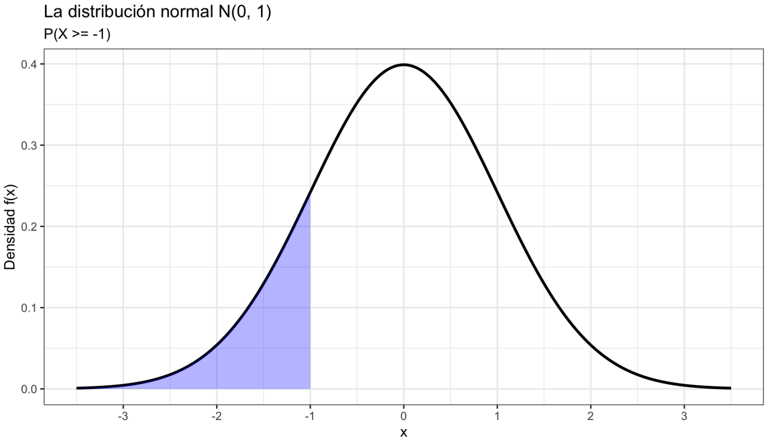
$$X = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

- La probabilidad en el intervalo (a, b) es el área bajo la curva.
- $P[a < X < b]$ se obtiene evaluando la diferencia entre las dos áreas. Área a la izquierda de b menos el área a la izquierda de a.
- $P[b < X]$ es el área a la izquierda de b. Veamos ejemplos.

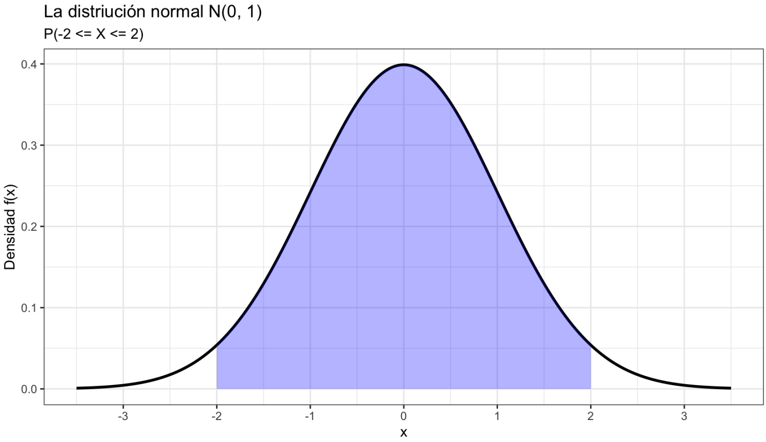
Comandos de R

- `dnorm (x ,mean = 0, sd = 1)` calcula $P(X = x)$, para una normal estandarizada.
- `pnorm(q, mean = 0, sd = 1, lower.tail = TRUE)` calcula $P(X \leqslant x)$, mientras que `lowe.tail FALSE` calcula $P(X > x)$.
- Supongamos una VA $X \sim \mathcal{N}(0, 1)$, calculemos la probabilidadde que X, se utiliza para leer
- Utilizamos el comando `pnorm(-1, 0, 1)` $= 0.1586553$
- Podemos hacerlo para un intervalo, calculemos $P(-2 \leqslant X \leqslant 1)$, que se realizó con el código `pnorm(1, 0, 1) - pnorm(-2, 0, 1)` $= 0.8185946$
- Ahora para $P(-2 \leqslant X \leqslant 2)$, se tiene que `pnorm(2, 0, 1)-pnorm(-2, 0, 1)`$=0.9544997$.

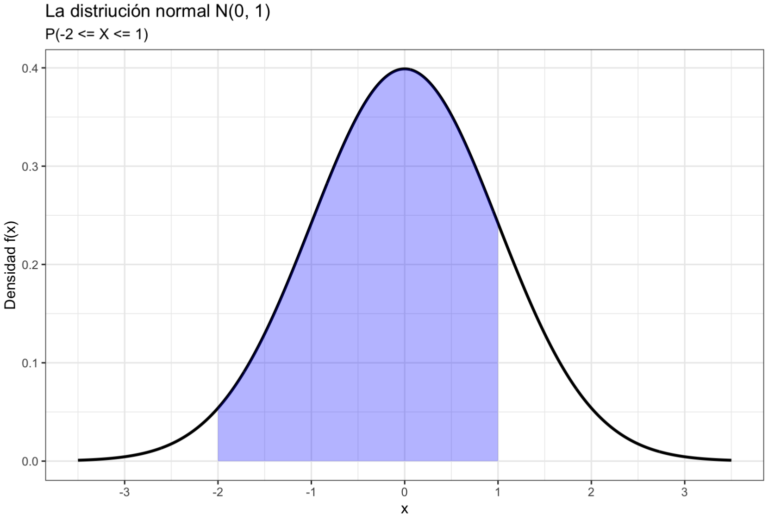
Gráfica en R


La distribución normal N(0, 1)
P(X >= -1)
Densidad f(x)
0.4
0.3
0.2
0.1
0.0
-3
-2
-1
0
1
2
3
x

Gráfica en R


La distriución normal N(0, 1)
P(-2 <= X <= 2)
Densidad f(x)
0.4
0.3
0.2
0.1
0.0
-3
-2
-1
0
1
2
3
x

Gráfica en R


La distriución normal N(0, 1)
P(-2 <= X <= 1)
0.4
0.3
0.2
0.1
0.0
Densidad f(x)
-3
-2
-1
0
1
2
3
x

Código en R

```
data.frame(x = c(-3.5, 3.5)) %>% ggplot(aes(x)) +
stat_function(fun = dnorm,
              n = 1000,
              args = list(mean = 0,
                          sd = 1),
              size = 1) +
geom_area(stat = 'function',
          fun = dnorm,
          fill = 'blue',
          xlim = c(-1, -3.5),
          alpha = 0.3) +
theme_bw() +
labs(y = 'Densidad f(x)',
   title = 'La distribucion normal N(0, 1)',
   subtitle = 'P(X >= -1)') +
scale_x_continuous(breaks = seq(-3, 3))
```

Código en R

```
data.frame(x = c(-3.5, 3.5)) %>% ggplot(aes(x)) +
stat_function(fun = dnorm,
              n = 1000,
              args = list(mean = 0,
                          sd = 1),
              size = 1) +
geom_area(stat = 'function',
          fun = dnorm,
          fill = 'blue',
          xlim = c(-2, 2),
          alpha = 0.3) +
theme_bw() +
labs(y = 'Densidad f(x)',
   title = 'La distriucion normal N(0, 1)',
   subtitle = 'P(-2 <= X <= 2)') +
scale_x_continuous(breaks = seq(-3, 3))
```

Código en R

```
data.frame(x = c(-3.5, 3.5)) %>% ggplot(aes(x)) +
stat_function(fun = dnorm,
              n = 1000,
              args = list(mean = 0,
                          sd = 1),
              size = 1) +
geom_area(stat = 'function',
          fun = dnorm,
          fill = 'blue',
          xlim = c(-2, 1),
          alpha = 0.3) +
theme_bw() +
labs(y = 'Densidad f(x)',
   title = 'La distriucion normal N(0, 1)',
   subtitle = 'P(-2 <= X <= 1)') +
scale_x_continuous(breaks = seq(-3, 3))
```

Ejemplos

- Suponga que el IQ de unos monos tiene promedio 9.5 y desviación estándar de 1.7. Encontremos la proporción de monos con IQ menor a 5.

$$\texttt{pnorm (5, 9.5, 1.7)} = 0.004059761$$

- ¿Qué tal la proporción de monos con IQ menor a 9?

$$\texttt{pnorm (9, 9.5, 1.7)} = 0.384334$$

- Los frascos de jabón de AMLO suelen tener un promedio de volumen 90 ml, y desviación estandar, calcule la proporción del jabón entre 84.5 y 90.1

$$\texttt{pnorm(90.5,93,5.3) - pnorm(84.5,93,5.3)} = 0.2641892.$$

Bibliografía

Bernoulli, J. (1713). *Ars coniectandi*. Impensis Thurnisiorum, fratrum.

Bhattacharyya, G. K., y Johnson, R. A. (1977). *Statistical concepts and methods*. Wiley.

Bickel, P. J., y Doksum, K. A. (2015). *Mathematical statistics: basic ideas and selected topics, volumes i-ii package*. CRC Press.

Fair, R. C. (1978). A theory of extramarital affairs. *Journal of political economy*, *86*(1), 45–61.

Groeneveld, R. A. (1998). A class of quantile measures for kurtosis. *The American Statistician*, *52*(4), 325–329.

Wackerly, D., Mendenhall, W., y Scheaffer, R. L. (2014). *Mathematical statistics with applications*. Cengage Learning.

¡Gracias por su atención!