Clase 4 Introducción Variables Aleatorias

Curso Introducción al Análisis de datos con R para la acuicultura.

Dr. José A. Gallardo y Dra. María Angélica Rueda.
jose.gallardo@pucv.cl | Pontificia Universidad Católica de Valparaíso

15 July 2021

PLAN DE LA CLASE

1.- Introducción

- Clasificación de variables aleatorias.
- Observar una variable cuantitativa continua.
- Predecir una variable cuantitativa continua.
- Identificar variables aleatorias discretas.
- Observar distribuciones de variables aleatorias discretas.

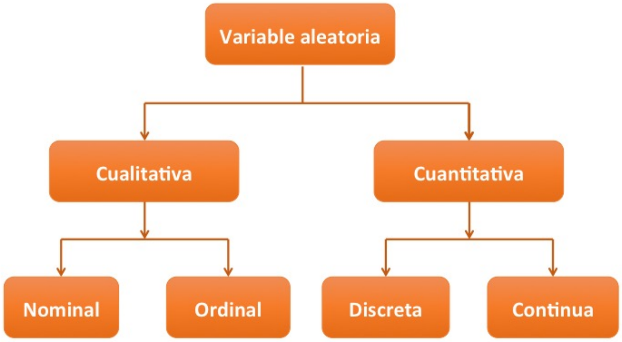
2.- Práctica con R y Rstudio cloud

- Observa y predice una variable aleatoria continua con distribución Normal.

- Observa y predice variables aleatorias discretas con distribución Bernoulli o Binomial.

- Elabora un reporte dinámico en formato pdf.

TIPOS DE VARIABLES ALEATORIAS



CASOS ESPECIALES

1.- Variable aleatoria binaria: Posee dos resultados posibles; por ejemplo, éxito o fracaso, macho o hembra, sano o enfermo, (0,1).

2.- Variable aleatoria dependiente del tiempo:

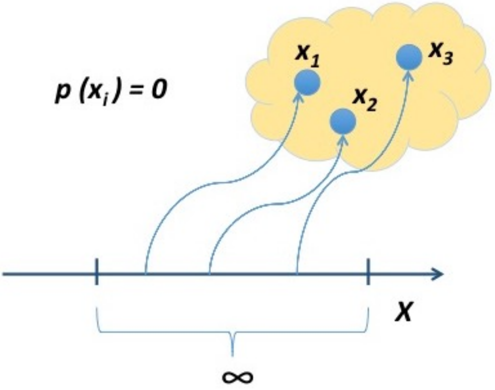
a) Discreta: Días a la muerte de un organismo o fallo de un componente en un sistema en un tiempo t.

b) Continua: Señales de sensores ambientales o señales biométricos.

Algunas de estas variables se conocen como **series de tiempo** y en términos estrictos son más bien una *sucesión de variables aleatorias* a través del tiempo.

VARIABLE ALEATORIA CUANTITATIVA CONTINUA

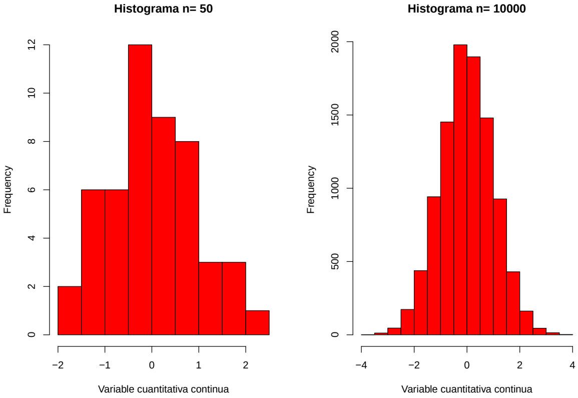
Definición: Puede tomar cualquier valor dentro de un intervalo (a,b), (a,Inf), (-Inf,b),(-Inf,Inf) y la probabilidad que toma cualquier punto es 0, debido a que existe un número infinito de posibilidades.



OBSERVAR UNA VARIABLE CUANTITATIVA CONTINUA

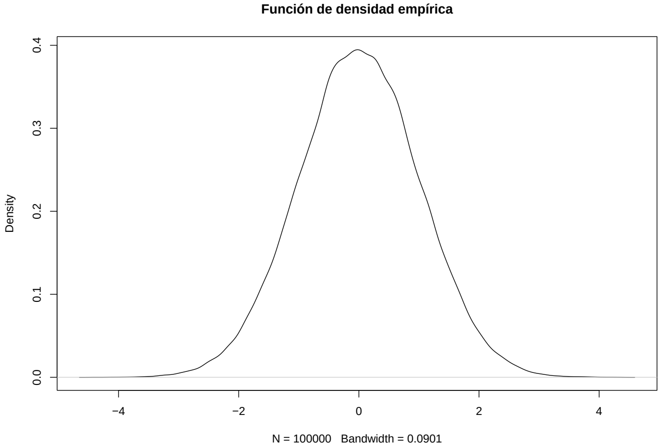
Al observar con un histograma **hist()** notamos que:

1. La frecuencia o probabilidad en un intervalo es distinta de cero.
2. Cuando aumenta el **n** muestral se perfila una distribución llamada **normal**.



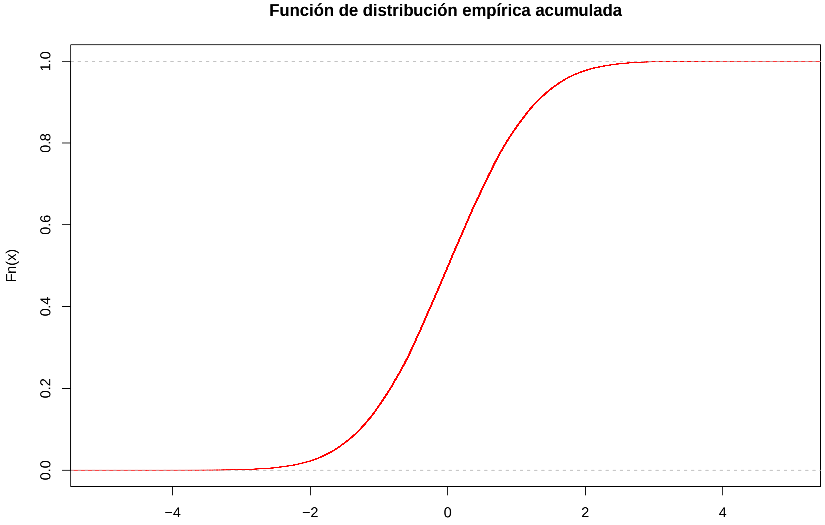
PREDECIR UNA VARIABLE CUANTITATIVA CONTINUA

Podemos predecir la probabilidad de que la variable aleatoria tome un determinado valor usando la función de densidad empírica **density()**.



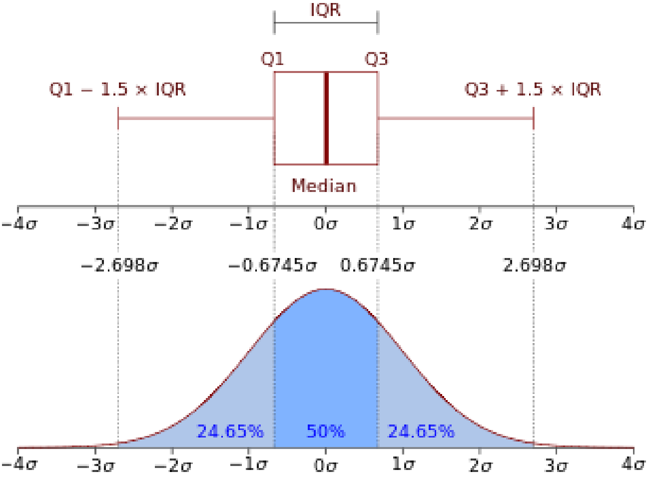
PREDECIR VARIABLES CONTINUAS 2

Podemos predecir la probabilidad de que la variable aleatoria tome un valor menor o igual a un determinado valor, usando la función de distribución empírica acumulada **ecdf()**.



OBSERVAR CON BOXPLOT

Las gráficas de cajas y bigotes (**boxplot()**) son muy adecuadas para observar variables aleatorias continuas.



VARIABLES ALEATORIAS DISCRETAS

Las variables aleatorias discretas son aquellas que presentan un número contable de valores; por ejemplo:

- **Número de salmones que están en un estanque** (10, 50, 70, etc.).
- **Número de parásitos presentes en tilapias** (1, 3, 5, 6, etc.).
- **Número de días a la muerte durante el desafío** (1, 2, 3,. . . , 40).

IMPORTANCIA DE IDENTIFICAR Y ANALIZAR VARIABLES ALEATORIAS DISCRETAS

- En gran parte, la ***distribución de variables aleatorias discretas*** puede ser **asimétrica a derecha o a izquierda**.
- Cuando las variables en estudio son conteos, proporciones o variables binarias (0 y 1); deben ser tratadas como **variables aleatorias discretas**.
- Según sea la variable aleatoria discreta, tendrá una función de distribución de probabilidad asociada **(Bernoulli, Binomial, Binomial Negativa, Poisson, entre otras)**.
- Es importante identificar la naturaleza que tiene nuestra variable en estudio, y así evitar errores en los análisis estadísticos que llevemos a cabo.

EXPERIMENTO BERNOULLI

Se saca un camarón al azar de una piscina, la probabilidad de que tenga sindrome de la mancha blanca es de 0.35. Sea X=1 si el camarón tiene sindrome de la mancha blanca y X=0 en el caso de que no tenga sindrome de la mancha blanca. ¿Cuál es la distribución de X?

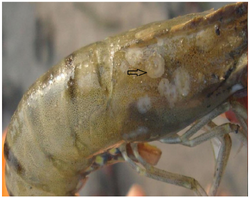
$$f(x) = P(X = x) = \begin{cases} 1 - p & ;si \quad x = 0 \\ p & ;si \quad x = 1 \end{cases}$$

$$f(x) = P(X = x) = p^x (1 - p)^{1-x}; x = 0; 1$$

$$X \sim Be(p)$$

EJEMPLO BERNOULLI

X: nº de veces que sale un camarón con sindrome de la mancha blanca al sacarlo una sola vez de la piscina.



	Fracaso	Éxito
x	0	1
f(x)=P(X=x)	1-p	p
	0.65	0.35

EXPERIMENTO BINOMIAL

Es un experimento que debe cumplir las siguientes condiciones:

1. El experimento consta de una secuencia de **n** ensayos idénticos.

2. En cada ensayo hay dos resultados posibles. A uno de ellos se le llama **éxito** y al otro, **fracaso**.

3. La probabilidad de éxito es constante de un ensayo a otro, nunca cambia y se denota por **p**. Por ello, la probabilidad de fracaso será **1-p**.

4. Los ensayos son **independientes**, de modo que el resultado de cualquiera de ellos ***no*** influye en el resultado de cualquier otro ensayo.

EJEMPLO BINOMIAL

Se tomó una muestra al azar de **10** salmones de diferentes estanques y se registró el evento “con parásitos” o “sin parásitos” durante 5 días. La probabilidad de tener parásitos es (para todos los salmones evaluados) igual a 0.25.

	Fracaso	Éxito
f(x)=P(X=x)	1-p	p
	0.75	0.25

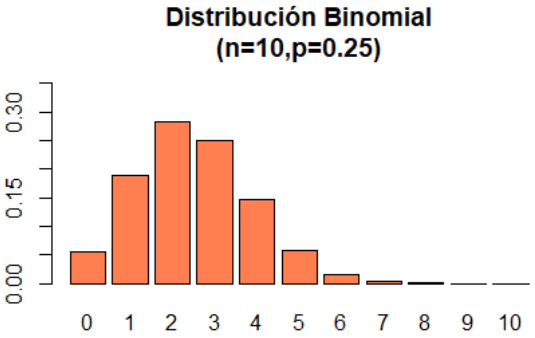
EJEMPLO DE EXPERIMENTO BINOMIAL

¿Cuál es la probabilidad de que 7 salmones tengan parásitos de los 10 salmones?

$$f(x) = P(X = x) = p^x(1-p)^{n-x}; x = 0, 1, 2, ..., 10$$

$$P(X = 7) = \binom{10}{7} 0.25^7(0.75)^{10-7} = \frac{10!}{7!(10-7)!} 0.25^7(0.75)^3 = 0.0031$$

EJEMPLO DE EXPERIMENTO BINOMIAL (DISTRIBUCIÓN)



PRÁCTICA ANÁLISIS DE DATOS

1.- Guía de trabajo Rmarkdown disponible en drive.
Clase 04-Introducción variables aleatorias

2.- La tarea se realiza en Rstudio.cloud, proyecto (**Clase 04-Introducción variables aleatorias**).

RESUMEN DE LA CLASE

- Identificamos y clasificamos variables aleatorias.
- Observamos una variable cuantitativa continua usando histogramas y boxplot.
- Predecimos el comportamiento de una variable cuantitativa continua usando funciones de densidad y de distribución acumulada.
- Estudiamos sobre variables aleatorias discretas y algunas distribuciones de probabilidad asociadas (Bernoulli o Binomial).