

Variables aleatorias continuas en R

Sesión 8

Natalie Julian - www.nataliejulian.com

Estadística UC y Data Scientist en Zippedi Inc.

Variable aleatoria continua

Una característica importante de una variable de tipo continua X es que dentro de cualquier intervalo de posibles valores de X existen infinitos posibles valores de X , es decir, es *densa*, esto no sucede en las variables discretas.

Ejemplos de variables aleatorias continuas

Algunas variables aleatorias continuas que son de particular interés modelar y estudiar en la actualidad:

- Vulnerabilidad de las edificaciones ante sismos u otros fenómenos
- Percepción de satisfacción de vida de la población de un país
- Tasas de transmisión de enfermedades respiratorias
- Precio promedio del dólar por semana
- Rating de series de anime

Muestras aleatorias de variables continuas

Distribución	Muestra aleatoria	Función de densidad	Función de distribución
Unif(a,b)	<code>runif(n= ,min=a,max=b)</code>	<code>dunif(x= ,min=a,max=b)</code>	<code>punif(q= ,min=a,max=b)</code>
Exp(λ)	<code>rexp(n= , rate=lambda)</code>	<code>dexp(x= ,rate=lambda)</code>	<code>pexp(q= ,rate=lambda)</code>
Gamma(ν,λ)	<code>rgamma(n= ,shape=,rate=)</code>	<code>dgamma(x= ,shape=,rate=)</code>	<code>pgamma(q= ,shape=,rate=)</code>
Normal(μ,σ)	<code>rnorm(n= ,mean=mu, sd=sigma)</code>	<code>dnorm(x= ,mean=mu,sd=sigma)</code>	<code>pnorm(q= ,mean=mu,sd=sigma)</code>
Normal(0,1)	<code>rnorm(n=)</code>	<code>dnorm(x=)</code>	<code>pnorm(q=)</code>

Ejemplo: Rating de animes distribuye normal

Asuma que la variable rating de anime distribuye normal con parámetros $\mu = 6.5$ y $\sigma = 1$.

a) ¿Cuál es la probabilidad de que el rating de un anime sea superior a 7?

Ejemplo: Rating de animes distribuye normal

Asuma que la variable rating de anime distribuye normal con parámetros $\mu = 6.5$ y $\sigma = 1$.

a) ¿Cuál es la probabilidad de que el rating de un anime sea superior a 7?

Necesitamos determinar $P(X > 7)$, note que:

$$P(X > 7) = 1 - P(X \leq 7) = 1 - F_X(7)$$

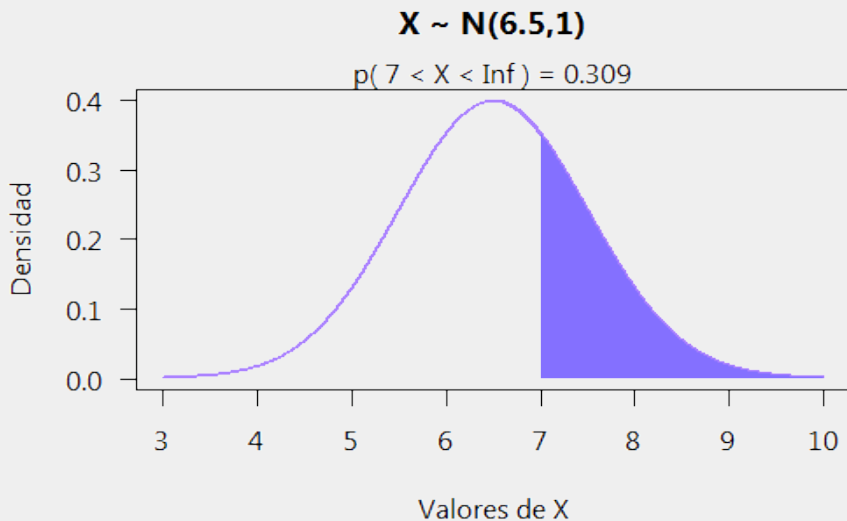
$$1 - \text{pnorm}(7, \text{mean}=6.5, \text{sd}=1) = 1 - 0.6914625$$

$$P(X > 7) = 0.3085375$$

Ejemplo: Rating con distribución normal

b) Muestre en un gráfico qué valor representa $P(X > 7)$. Comente.

Función de distribución como área



Ejemplo: Rating con distribución normal

c) Plantee la probabilidad como área y cáculelela en R.

Ejemplo: Rating con distribución normal

- c) Plantee la probabilidad como área y cálculela en R.

Se define la función densidad y ésta se integra a partir del valor 7:

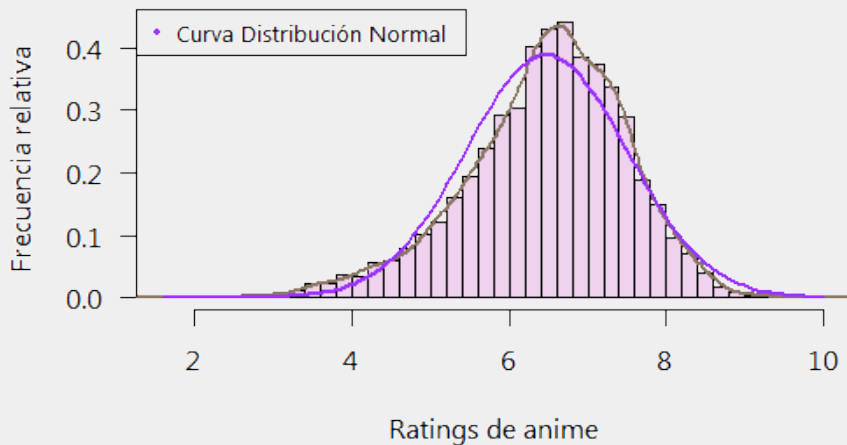
```
densidad<-function(x){  
  1/sqrt(2*pi*sd2) * exp(-(x - mu)2/(2 * sd2))  
}
```

```
integrate(densidad,lower=a,upper=Inf)  
0.3085375 with absolute error < 7.3e-07
```

Ejemplo: Rating con distribución normal

- d) Cargue los datos `anime` y muestre en un gráfico la distribución del rating. Sobreponga una curva de distribución normal. Comente.

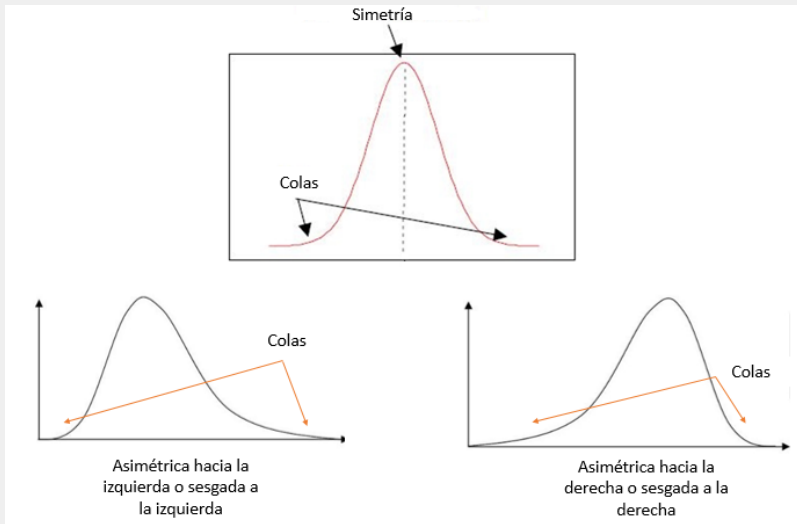
Histograma de los ratings de anime



*Revisar **siempre** que el histograma de frecuencias relativas de entre 0 y 1, a veces, puede ocurrir que de mayor a uno, en dicho caso, se debe reponderar la variable.*

Comportamiento de curvas de densidad

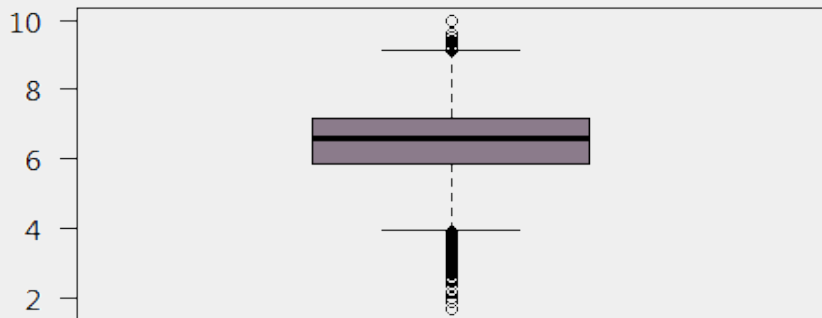
Es posible observar distintos comportamientos en las curvas de densidad dependiendo de cierto sesgo o inclinación que pudiera existir en base a los datos:



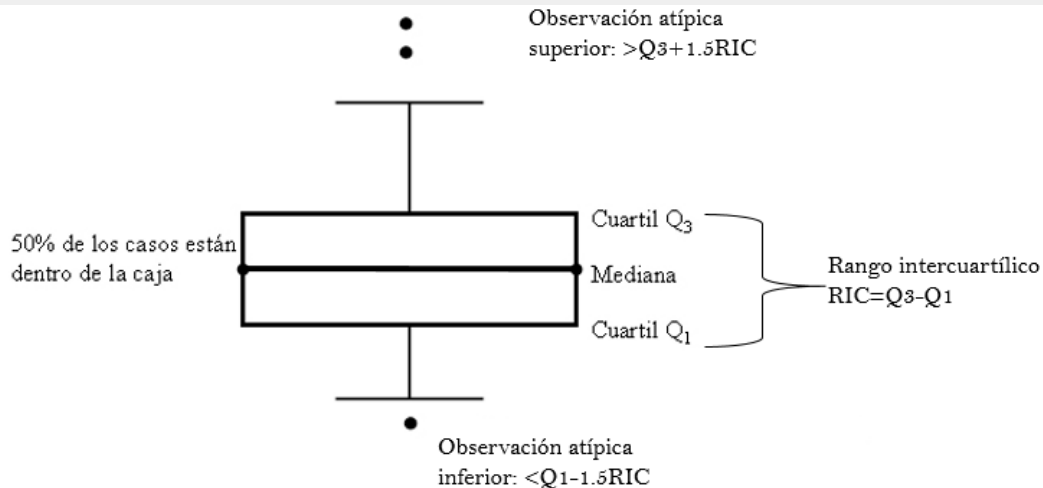
Boxplot o Diagrama de caja

Otro gráfico interesante es el boxplot o diagrama de caja y bigote. Este gráfico se obtiene fácilmente con el comando `boxplot()`.

Boxplot de ratings de anime



¿Qué información contiene el boxplot?



¿Cuándo usar un boxplot?

El boxplot también puede usarse para variables discretas, en el fondo, utilizar siempre que tenga sentido hablar de cuartiles.

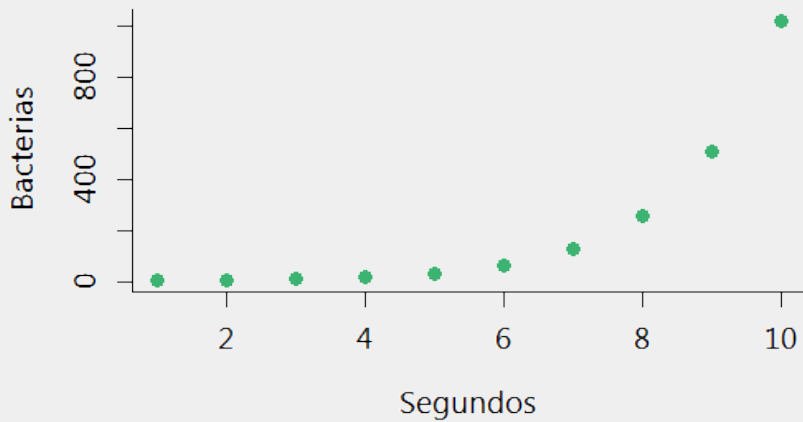
Es ampliamente utilizado realizar boxplots para comparar categorías, de tal modo, de comparar estadísticas para cada categoría.

Ejemplo: Crecimiento de bacterias en el tiempo

- a) Un equipo de trabajo registró durante 10 segundos el crecimiento de bacterias en una Placa de Petri. Los valores registrados fueron 2, 4, 8, 16, 32, 64, 128, 256, 512 y 1024. La académica a cargo de la investigación sugiere que una curva exponencial pudiera ajustar bien la tendencia creciente de las bacterias. Realice un gráfico y comente.

Ejemplo: Crecimiento de bacterias en el tiempo

Crecimiento de bacterias en placa de petri

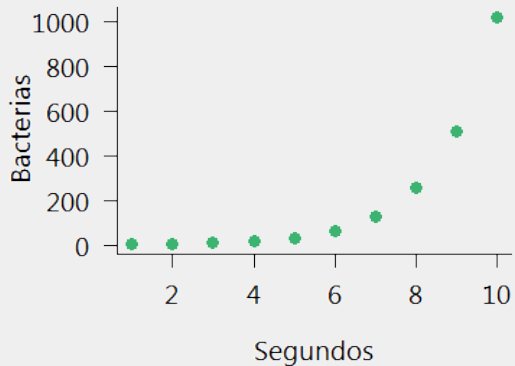


Ejemplo: Crecimiento de bacterias en el tiempo

- b) Estime con los datos el parámetro λ . Realice 10 simulaciones de $Exp(\hat{\lambda})$ de modo de obtener una simulación del crecimiento de bacterias. Compare con los datos reales. Comente.

Ejemplo: Crecimiento de bacterias en el tiempo

Crecimiento real



Crecimiento simulado

