La normal

J. Abellán

7/11/2019

La normal o gaussiana

Hemos visto un 'mecanismo' por el que la función de distribución *normal* o *gaussiana* aparece por doquier: el teorema del límite central.

Vamos ahora a familiarizarnos con ella y con su función acumulada y a ver como contestamos a diferentes preguntas sobre probabilidades.

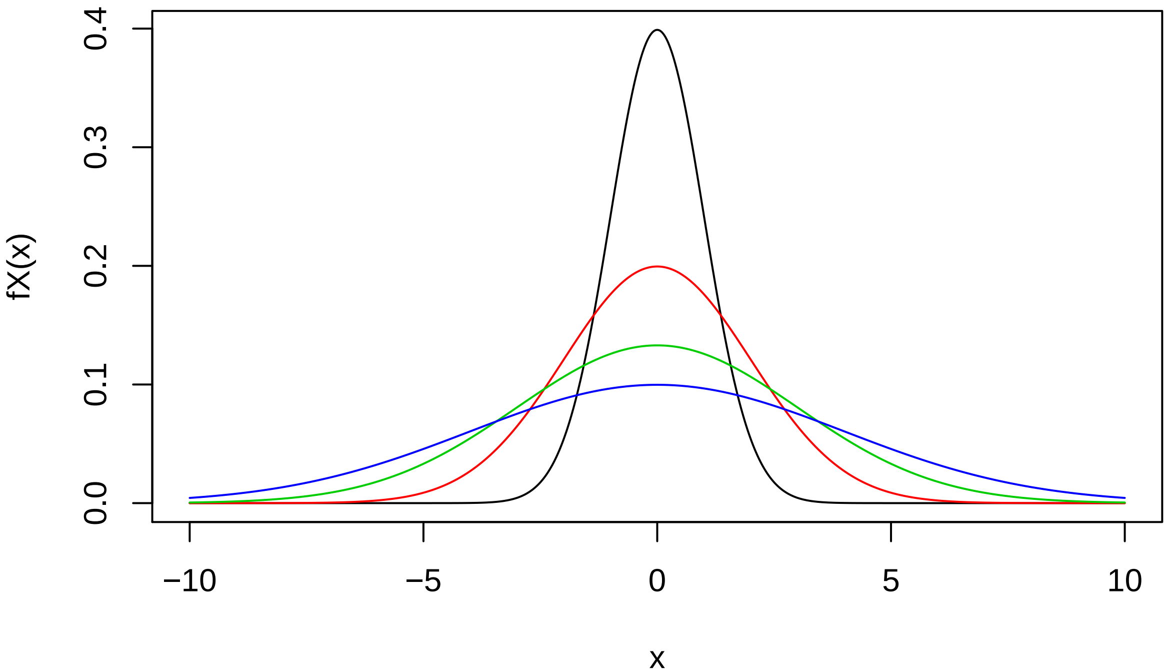
La función de distribución $f_X(x)$

La función de distribución normal de parámetros μ, σ se define:

$$f_X(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Dibujamos superpuestas las funciones de distribución $f_X(x)$ de diferentes variables aleatorias X gaussianas de igual media $\mu_X = 0$ pero diferentes desviaciones $\sigma_X = 1, 2, 3, 4$ para ver la relación de éstas con el *ancho* de la gaussiana:

```r
#library("latex2exp", lib.loc="~/R/i686-pc-linux-gnu-library/3.2")

x <- seq( - 10, 10, .01 )

plot( x, dnorm( x, 0, 1 ),

      # xlab = latex2exp("$x$"),
      xlab = "x",

      # ylab = latex2exp("$\\f_X(x)$"),
      ylab = "fX(x)",

      type = "l",

      #main = latex2exp("$\\N(0,\\sigma)$")

    )

lines( x, dnorm( x, 0, 2 ), col = 2 )

lines( x, dnorm( x, 0, 3 ), col = 3 )

lines( x, dnorm( x, 0, 4 ), col = 4 )
```



A continuación dibujamos las funciones acumuladas $F_X(x)$

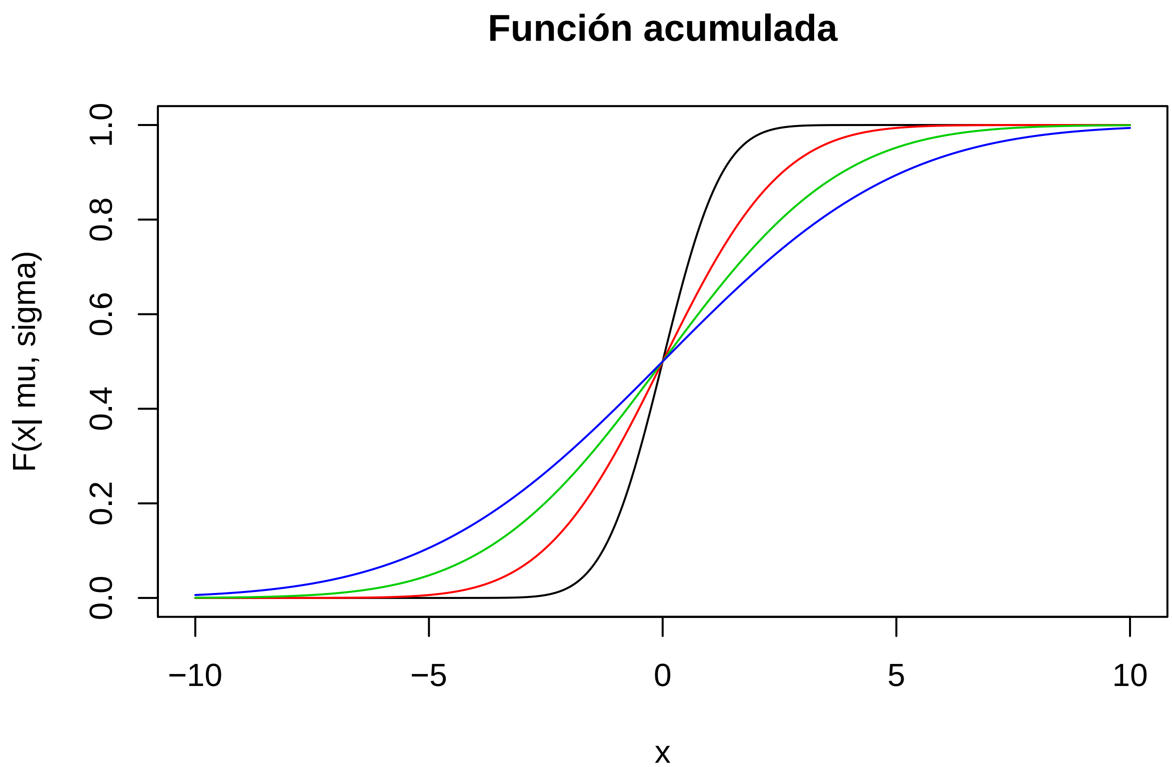
La función acumulada $F_X(x;\mu,\sigma)$

La función acumulada se define:

$$F_X(x|\mu,\sigma) \equiv \int_{-\infty}^{x} dx' f_X(x'|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} dx' e^{-\frac{1}{2}\frac{(x'-\mu)^2}{\sigma^2}}$$

Como no existe la primitiva de la normal, el cálculo de $F_X(x;\mu,\sigma)$ se hace numéricamente. Dibujamos las acumuladas de las anteriores funciones:

```
plot( x, pnorm( x, 0, 1 ),

      #ylab = latex2exp("$\\F_X(x) = P(X<x)$"),
      ylab = "F(x| mu, sigma)",

      type = "l",

      main = "Función acumulada"

    )

lines( x, pnorm( x, 0, 2 ), col = 2 )

lines( x, pnorm( x, 0, 3 ), col = 3 )

lines( x, pnorm( x, 0, 4 ), col = 4 )
```



Probabilidades.

Supongamos que el coeficiente intelectual de una población es una variable aleatoria normal de media $\mu = 100$ y desviación $\sigma = 15$.

- ¿Cuál es la probabilidad de que al escoger una individuo al azar tenga un coeficiente intelectual de 100? Respuesta: **¡cero!**
- ¿Cuál es la probabilidad de escoger un individuo al azar que tenga un coeficiente intelectual entre 95 y 105?

$$P(95 < X < 105) = \int_{95}^{105} dx' f_X(x') = \int_{-\infty}^{105} dx' f_X(x') - \int_{-\infty}^{95} dx' f_X(x') = F_X(105|100, 15) - F_X(95|100, 15)$$

```
Pb <- pnorm( 105, 100, 15 )

Pa <- pnorm( 95, 100, 15 )

Px <- Pb - Pa
```

Respuesta: $P(95 < X < 105) = 0.2611173$

- ¿Cuál es la probabilidad de escoger un individuo al azar que tenga un coeficiente intelectual entre $100 - 15$ y $100 + 15$? Es decir, cuál es la probabilidad de estar a menos de un σ de la media:

Respuesta: 0.6826895, es decir, aproximadamente un 68%.

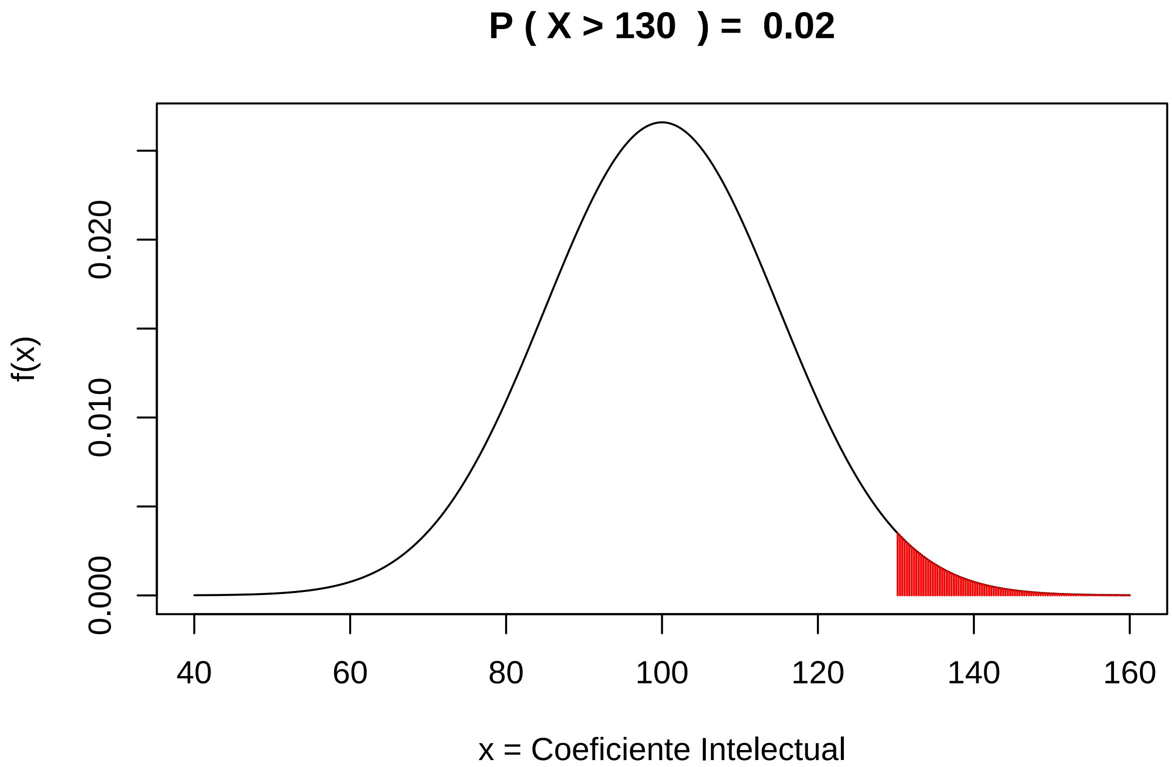
- ¿Cuál es la probabilidad de escoger cinco individuos al azar y que la media de sus coeficientes intelectuales esté entre $100 - 15$ y $100 + 15$?

Respuesta: 0.9746527. Recuerde que la variable aleatoria X_m es distinta de X. La relación entre sus desviaciones es $\sigma_{X_m} = \sigma_X/\sqrt{n}$, donde n es el número de medidas o individuos de la muestra.

- ¿Cuál es la probabilidad de escoger un individuo al azar y que tenga un coeficiente intelectual superior a 2σ?

Respuesta: $P(X > 2\sigma) = 0.0227501$. Para verlo mejor:

```
#Parámetros: media y desviación
muX <-100 ; sigmaX <- 15

#Valor escogido
xe <- muX + 2 * sigmaX

fxe <- dnorm( xe, muX, sigmaX )

#Probabilidad de que X sea menor que mu+3 sigma
Fmenor <- pnorm( xe, muX, sigmaX )

#Por tanto
Fmayor <- 1 - Fmenor

#Dibujo la normal del coeficiente intelectual
x1 <- muX - 4 * sigmaX

x2 <- muX + 4 * sigmaX

x <- seq( x1, x2, len = 400 )

fX <- dnorm( x, muX, sigmaX )

plot( x, fX,

      xlab = "x = Coeficiente Intelectual",

      #ylab = latex2exp(" $ f_X ( x ) $"),
      ylab = "f(x)",

      type = "l",

      main = paste("P ( X >", xe, " ) = ", round( Fmayor, 2 ) )

    )

segments( x[ x > xe ], rep( 0, length( x[ x > xe ] ) ),

          x[ x > xe ], fX[ x > xe ],

          col = 2

        )
```



El área en rojo es la respuesta que nos piden.