PREWORK
SESIÓN 4

Introducción

En este prework, conocerás algunas definiciones y resultados importantes de la estadística inferencial. En el work trabajarás con datos y los cálculos serán muy rápidos y fáciles de hacer con algunas funciones de R.

Objetivos

- Conocer algunas funciones de densidad de variables aleatorias muy útiles y comunes.
- Conocer lo que dice el teorema central del límite
- Conocer procedimientos para llevar a cabo algunos contrastes de hipótesis

Temas

- Experimentos aleatorios, probabilidad
- Funciones de distribución para variables discretas y continuas
- Distribución normal

Experimentos aleatorios y probabilidad

En la vida cotidiana, existen una gran cantidad de fenómenos que son aleatorios, es decir, que su resultado no puede predecirse de forma exacta con total seguridad, por lo tanto están sujetos a cierta incertidumbre, la cual puede modelarse para determinar la magnitud de la posibilidad de cierto evento.

En estadística, un experimento del cual no se puede conocer su resultado con certeza, pero del cual se pueden enlistar todos los posibles resultados, es llamada con **experimento aleatorio**. Al conjunto de todos estos posibles resultados se le conoce como el **espacio muestral**. Por ejemplo:

1) si un experimento consiste en lanzar un dado perfectamente balanceado, entonces el espacio muestral queda como:

$$S = \{1, 2, 3, 4, 5, 6\}$$

2) si un experimento consiste en lanzar dos monedas balanceadas, entonces el espacio muestral queda como:

$$S = \{(A, A), (A, S), (S, A), (S, S)\}$$

Donde *A* representa águila y *S* representa sol.

3) si un experimento consiste en medir el monto gastado por un persona al entrar a una tienda departamental, entonces el espacio muestral queda como:

$$S = \{x \; tal \; que \; 0 \leq x < inf)$$

Dado el espacio muestral de un experimento aleatorio, podemos comenzar a pensar en un **evento** particular. Por ejemplo,

1) Para nuestro experimento de lanzar un dado perfectamente balanceado, supongamos que estamos interesados en el evento de que el evento sea obtener como resultado un número par. El conjunto de nuestro evento queda como:

$$E = \{2, 4, 6\}$$

2) Para nuestro experimento de lanzar dos monedas balanceadas, supongamos que estamos interesados en el evento de que obtengamos como resultado al menos un sol. El conjunto de nuestro evento queda como:

$$E = \{(A, S), (S, A), (S, S)\}$$

Con esto en mente, podemos definir de forma intuitiva la probabilidad de un evento, como el número de veces que ocurre un evento entre el total de eventos posibles. Para ello, necesitamos la cardinalidad de un conjunto, que no es más que el número de elementos que posee dicho conjunto.

Por ejemplo, la probabilidad de que ocurra el evento de obtener un resultado par al lanzar un dado perfectamente balanceado es

$$P = \frac{|E|}{|S|} = \frac{|\{2, 4, 6\}|}{|\{1,2,3,4,5,6\}|} = \frac{3}{6} = 0.5$$

Esta definición, aunque intuitiva, tiene ciertos aspectos teóricos que son objeto de debate, sin embargo, para lo que estamos buscando estudiar es suficiente.

Funciones de distribución para variables discretas y continuas

En la práctica, podemos encontrar valores o datos, que de acuerdo a la teoría estadística pueden ser considerados como provenientes de algunas poblaciones o conjuntos de datos más grandes cuyas distribuciones están bien estudiadas.

Ejemplos de tales distribuciones son la distribución binomial y la normal. Todas las distribuciones tienen ciertas funciones asociadas llamadas funciones de masa, para el caso de variables discretas, y de densidad, para el caso de variables continuas, que nos ayudan a obtener probabilidades relacionadas con la obtención nuevos datos o valores futuros provenientes de las poblaciones.

Las funciones de masa de probabilidad cumplen con algunas propiedades, por ejemplo:

- La suma total de las masas de probabilidad es igual a 1 siempre
- $P(X = c)$ para cualquier valor c para el que se encuentre definida la función de masa es igual a la probabilidad de que ocurra c
- Si X es la variable aleatoria que genera los valores de la población, entonces la probabilidad de que X se encuentre en determinado intervalo *(a, b)* es igual a la suma de las masas para los eventos que se encuentran entre los puntos *a y b*.

Ejemplos de funciones de masa de probabilidad son las distribuciones:

- Uniforme
- Binomial
- Geométrica y hipergeométrica
- Poisson

Por su parte, las funciones de densidad de probabilidad cumplen con algunas propiedades, por ejemplo:

- El área total bajo la curva de la función de densidad de probabilidad es igual a 1 siempre
- Si X es la variable aleatoria que genera los valores de la población, entonces la probabilidad de que X se encuentre en determinado intervalo *(a, b)* es igual al área bajo la curva entre los puntos *a y b*.
- $P(X = c) = 0$ para cualquier valor c para el que se encuentre definida la función de densidad.

Ejemplos de funciones de densidad de probabilidad son las distribuciones:

- Uniforme (¡exacto! otra vez esta distribución)
- Exponencial
- Beta y Gamma
- Normal

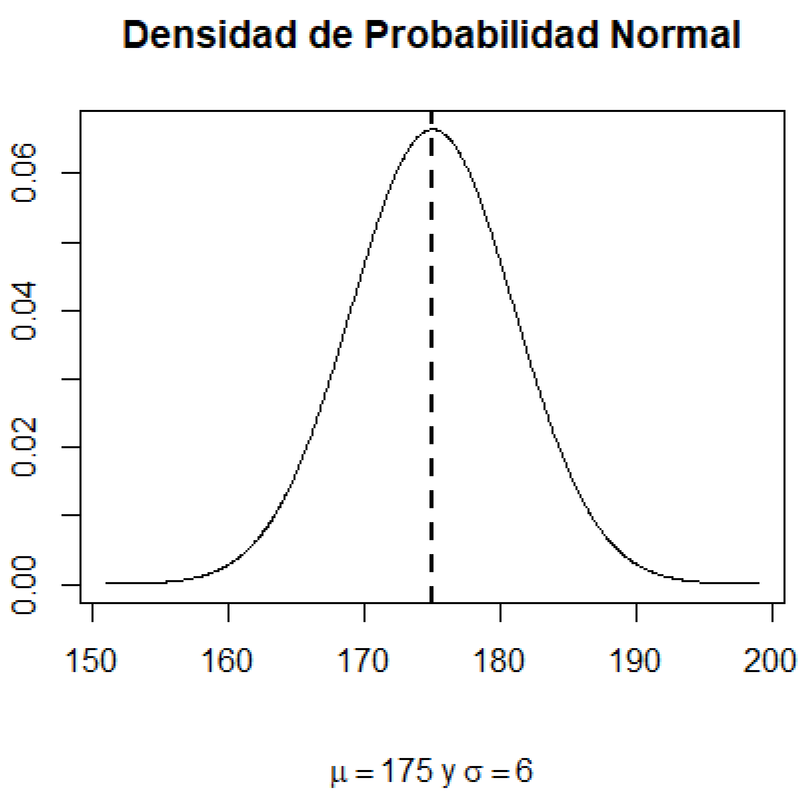
Distribución normal

En R, podemos obtener valores de la función de densidad normal con la función dnorm. Por ejemplo, supongamos que la variable aleatoria X representa la estatura de una persona elegida aleatoriamente de una población, donde esta población tiene una distribución normal con media 175 centímetros y desviación estándar de 6 centímetros. Los valores de la función de densidad correspondiente para cada posible valor x de X, los podemos obtener como dnorm(x, mean = 175, sd = 6). Así, para un conjunto de valores diferentes de X, digamos x <- seq(-4, 4, 0.01)*6 + 175, podemos obtener sus valores correspondientes y <- dnorm(x, mean = 175, sd = 6) y pintar la gráfica de la función de densidad que nos ayudará a obtener probabilidades relacionadas con futuras observaciones de X

```
plot(x, y, type = "l", xlab = "", ylab = "")
```

```
title(main = "Densidad de Probabilidad Normal", sub = expression(paste(mu == 175, " y ", sigma == 6)))
```

```
abline(v = 175, lwd = 2, lty = 2) # La media es 175
```



Ejemplo. Las calificaciones para un examen de admisión a una universidad están normalmente distribuidas con media de 75 y desviación estándar 10. ¿Qué fracción de las calificaciones se encuentra entre 80 y 90?

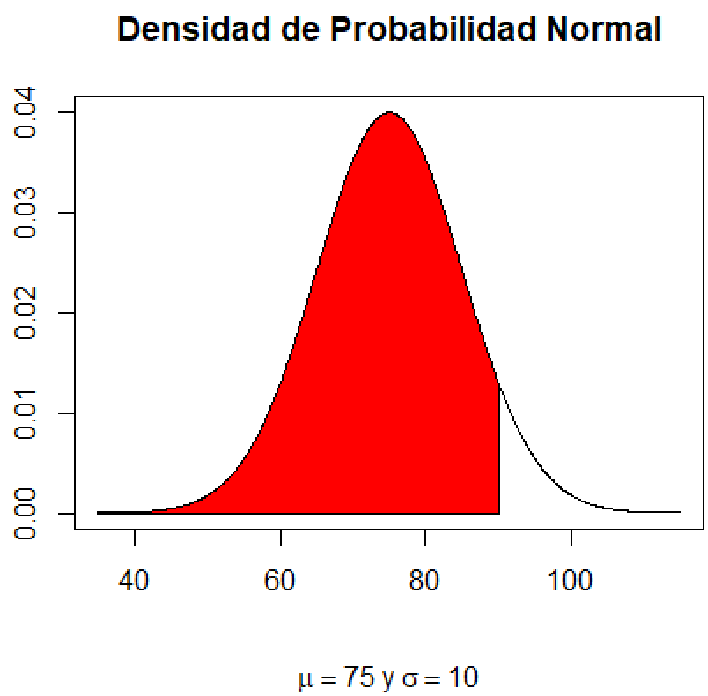
Solución.

Si X representa una calificación elegida al azar, entonces para obtener $P(X \leq 90)$, es decir, la probabilidad de que X tome un valor menor o igual a 90, ejecutamos:

```
pnorm(q = 90, mean = 75, sd = 10)
```

```
0.9331928
```

Observemos el área que corresponde a esta probabilidad en la siguiente gráfica en color rojo:

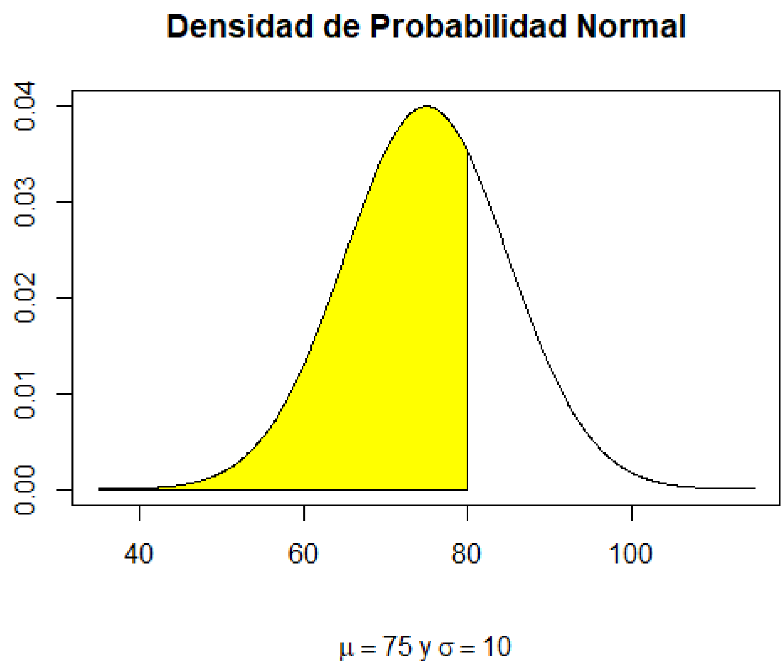


Para obtener $P(X \leq 80)$, es decir, la probabilidad de que X tome un valor menor o igual a 80, ejecutamos:

```
pnorm(q = 80, mean = 75, sd = 10)
```

```
0.6914625
```

Observemos el área que corresponde a esta probabilidad en la siguiente gráfica en color amarillo:

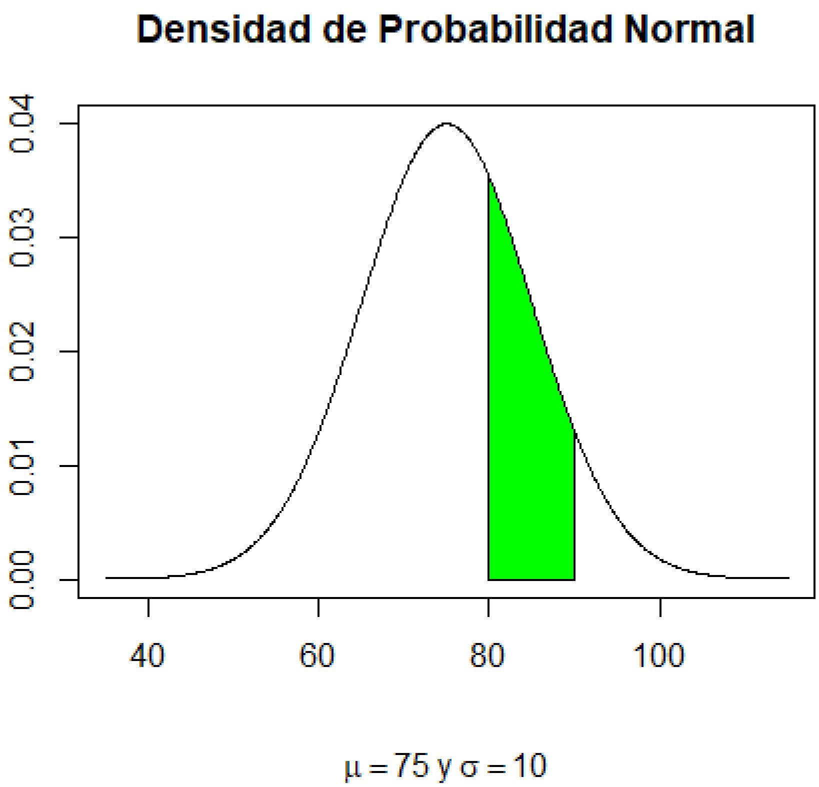


Para obtener $P(80 \leq X \leq 90)$, es decir, la probabilidad de que X tome un valor mayor o igual a 80 y menor o igual a 90, debemos correr

```
pnorm(q = 90, mean = 75, sd = 10) - pnorm(q = 80, mean = 75, sd = 10)
```

```
0.2417303
```

Observemos el área que corresponde a esta probabilidad en la siguiente gráfica en color verde:



Quiz

1. Considera el experimento de lanzar dos monedas balanceadas, una tras otras, y el evento de obtener sólo un Sol. ¿Cuál es la probabilidad de que ocurra dicho evento?
    a. 0.5
    b. 0.75
    c. 1
    d. 0.25
2. Considera el experimento de tirar un dado balanceado y el evento de que el resultado sea mayor a 5. ¿Cuál es la probabilidad de que ocurra dicho evento?
    a. 0.3333
    b. 0.1667
    c. 0.6667
    d. 0
3. Tenemos la variable aleatoria “Número de personas que deciden tramitar una tarjeta de crédito en el banco ABC”. ¿Qué tipo de función de distribución tiene?
    a. Función de distribución de probabilidad
    b. Función de densidad de probabilidad

    c. Función de distribución normal
    d. Función de masa de probabilidad
4. Tenemos la variable aleatoria "Gasto federal en salud pública". ¿Qué tipo de función de distribución tiene?
    a. Función de distribución de probabilidad
    b. Función de densidad de probabilidad
    c. Función de distribución exponencial
    d. Función de masa de probabilidad

BIBLIOGRAFÍA UTILIZADA

- Wackerly, D. et al. (2010). Estadística Matemática con Aplicaciones. Cengage Learning Editores, S.A. de C.V.