

Tarea_02

Nota: 10

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --

## v ggplot2 3.3.2      v purrr  0.3.4
## v tibble  3.0.4      v dplyr  1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(rio)
library(knitr)
library(DescTools)
library(data.table)

##
## Attaching package: 'data.table'

## The following object is masked from 'package:DescTools':
##
##      %like%

## The following objects are masked from 'package:dplyr':
##
##      between, first, last

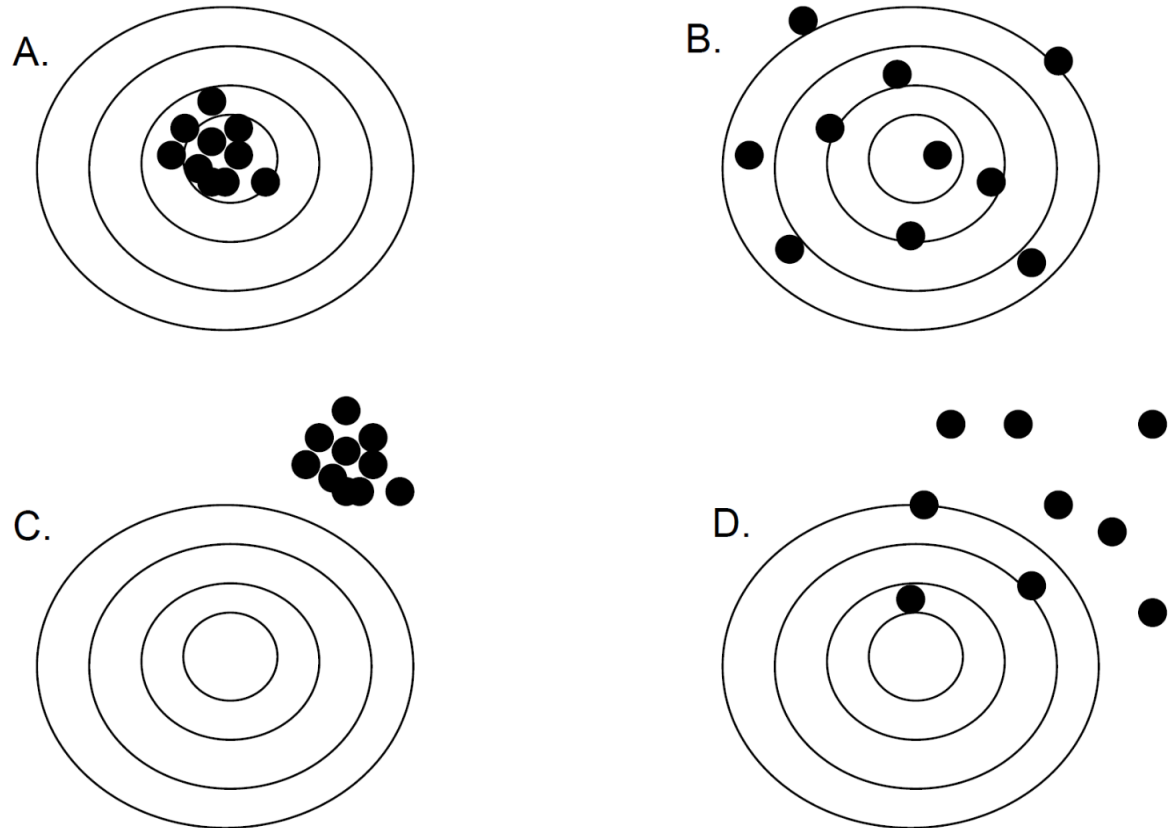
## The following object is masked from 'package:purrr':
##
##      transpose

library(modelbased)
library(emmeans)
```

Nombre del estudiante: Tania Picado Pérez

1. Identifica los conceptos de precisión, sesgo, y exactitud en los diagramas siguientes. Una pista, el diagrama B es no sesgado pero impreciso = inexacto.

```
include_graphics("figura.png")
```



A: no sesgado pero preciso = exacto

B: no sesgado pero impreciso = inexacto

C: sesgado pero preciso = inexacto

D: sesgado pero impreciso = inexacto

2. Un investigador esta investigando algunas características morfométricas, tales como el peso corporal (g) y la longitud del pico (mm), de 2 especies de de pinzones africanos.

Datos: PinzonesAfricanos.csv

```
Datos_pinzones = import("PinzonesAfricanos.csv")
```

```
head(Datos_pinzones)
```

```
##      Especie Peso LargoDePico
## 1 WB.SPARG  40         10.6
## 2 WB.SPARG  43         10.8
```

```
## 3 WB.SPARG 37 10.9
## 4 WB.SPARG 38 11.3
## 5 WB.SPARG 43 10.9
## 6 WB.SPARG 33 10.1
```

¿Cuál de las especies tiene el pico más largo?

- a. Analiza el resultado desde el punto de vista del valor de P.

```
t.test(LargoDePico ~ Especie, data = Datos_pinzones, var.equal = TRUE)
```

```
##
## Two Sample t-test
##
## data: LargoDePico by Especie
## t = -22.4, df = 28, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -3.666094 -3.051763
## sample estimates:
## mean in group CRU.WAXB mean in group WB.SPARG
## 7.378571 10.737500
```

De acuerdo al estadístico $t = -22.4$, se rechaza la hipótesis nula. Con respecto al valor de $p\text{-value} < 2.2e-16$ existe una diferencia estadísticamente significativa ($p < 0,05$) entre el largo de pico de las dos especies de Pinzones

- b. Analiza el resultado desde el punto de vista del tamaño del efecto.

```
MeanDiffCI(LargoDePico ~ Especie, data = Datos_pinzones)
```

```
## meandiff lwr.ci upr.ci
## -3.358929 -3.667242 -3.050615
```

$WB.SPARG - CRU.WAXB = 3.4$ (IC 95% = 3,0 - 3,7) Se puede inferir con un 95% de confianza que la verdadera diferencia en el largo de pico para las especies (WB.SPARG,CRU.WAXB) de Pinzones Africanos se encuentra entre 3,0 - 3,7cm

3. Muchas personas creen que para lograr un estimado preciso de la media poblacional es necesario muestrear una fracción sustancial de la población. Esta pregunta esta desarrollada para probar si tal aseveración es cierta o no.

- a. Para una población con desvío estándar 50, encuentra el error estándar de los siguientes valores de N (tamaño de población) y n (tamaño de muestra). Coloca los errores estándares calculados en las celdas vacías de la tabla.

N =	100	1000	10000	1000000	10000000
n =	10	100	1000	10000	1000000
Error estándar =	15.1	4.7	1.5	0.4	0.1

```
Error_E_100 <- (50/sqrt(10))*(sqrt(1-(10/100)))*(sqrt(100/(100-1)))
```

```
Error_E_100
```

```
## [1] 15.07557
```

```
Error_E_1000 <- (50/sqrt(100))*(sqrt(1-(100/1000)))*(sqrt(1000/(1000-1)))
```

```
Error_E_1000
```

```
## [1] 4.74579
```

```
Error_E_10000 <- (50/sqrt(1000))*(sqrt(1-(1000/10000)))*(sqrt(10000/(10000-1)))
```

```
Error_E_10000
```

```
## [1] 1.500075
```

```
Error_E_100000 <- (50/sqrt(10000))*(sqrt(1-(10000/100000)))*(sqrt(100000/(100000-1)))
```

```
Error_E_100000
```

```
## [1] 0.474344
```

```
Error_E_1000000 <- (50/sqrt(100000))*(sqrt(1-(100000/1000000)))*(sqrt(1000000/(1000000-1)))
```

```
Error_E_1000000
```

```
## [1] 0.1500001
```

b. Se te ocurre alguna explicación de los resultados obtenidos.

Con un desvío estandar igual, si se aumenta el tamaño de la muestra se disminuye el error estándar, esto debido a que en este caso está aumentando el tamaño de la muestra pero la desviación estándar sigue igual, por lo cual esta desviación disminuye en relación con los datos de N y n.