

## Tarea\_02

Nota: 9,0

```
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

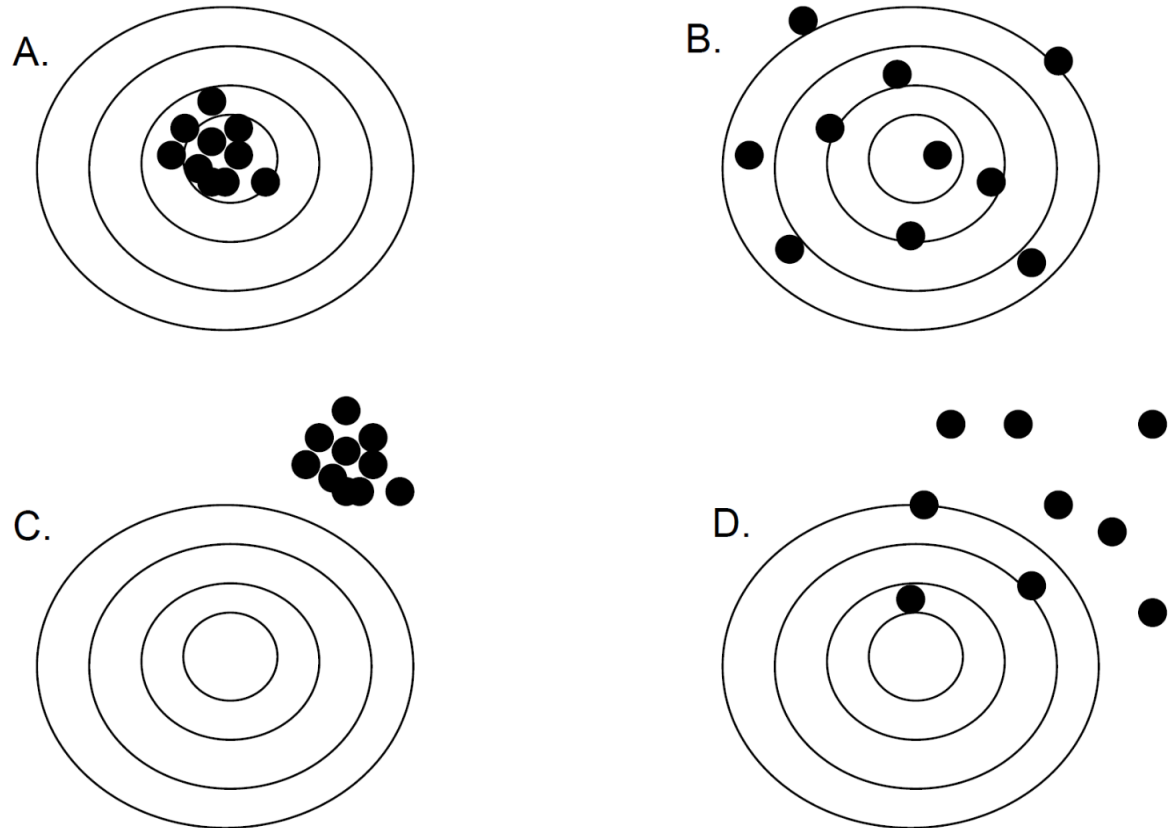
## The following object is masked from 'package:purrr':
##
##      transpose

library(modelbased)
library(emmeans)
```

Nombre del estudiante: María José Cordero Vega

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, preciso y exacto

B: no sesgado pero impreciso = inexacto

C: Sesgado, preciso e inexacto

D: Sesgado, impreciso e inexacto

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

Datos: PinzonesAfricanos.csv

```
pina <- import("PinzonesAfricanos.csv")
```

```
pina %>%
  group_by(Especie) %>%
  summarize(mean(LargoDePico))
```

```
## 'summarise()' ungrouping output (override with '.groups' argument)
```

```
## # A tibble: 2 x 2
##   Especie 'mean(LargoDePico)'
##   <chr>          <dbl>
## 1 CRU.WAXB          7.38
## 2 WB.SPARG          10.7
```

```
t.test (LargoDePico ~ Especie, data = pina, 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
```

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

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

CRU.WAXB - WB.SPARG = 3,35 (IC 95% = 3,05 - 3,67)

### ¿Cuál de las especies tiene el pico más largo?

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

La especie WB.SPARG tiene la longitud de pico más grande que la especie CRU.WAXB con diferencias estadísticas significativas, por tanto, el valor del pico no es igual. El estadístico T se encuentra por encima del -1,96, por tanto, se tienen elementos para rechazar la hipótesis nula. Además, el valor de  $p < 0.05$ , por lo que se rechaza la hipótesis nula de que son iguales.

Desde el punto de vista del tamaño del efecto, lo que yo encontré en los datos es que la diferencia es de 3,35. Y puedo decir con un 95% de confianza que la verdadera diferencia en el largo del pico en ambas poblaciones de pinzones africanos se encuentra entre 3,05 y 3,67 cm.

**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 está desarrollada para probar si tal aseveración es cierta o no.**

- 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 =	15,8	5	1,6	0,5	0,05

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

Al aumentar la muestra poblacional, el tamaño de la muestra es mayor. Cuando se aumenta la muestra poblacional puede ser necesario aumentar la muestra muestral. Sin embargo, la proporción a muestrear va a depender de la variabilidad en la población. En este caso preciso, con una desviación estándar de 50, es necesario aumentar la muestra muestral para disminuir el error estándar.

No existe la muestra poblacional, no se a que te refieres