

## Tarea\_02

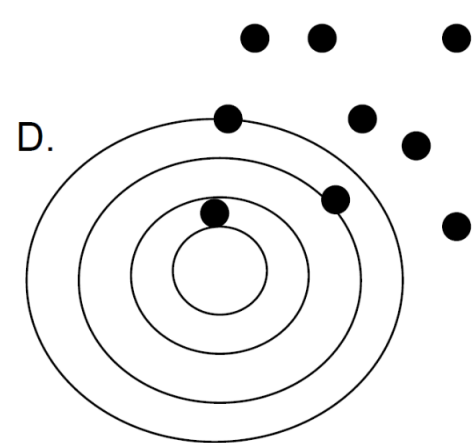
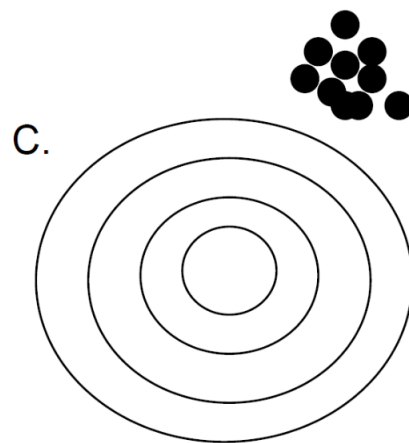
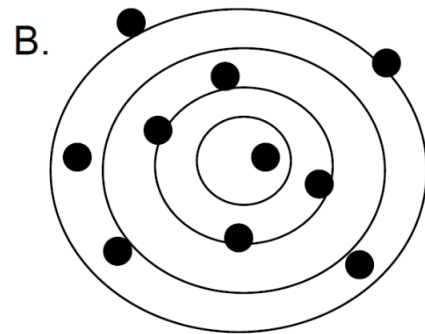
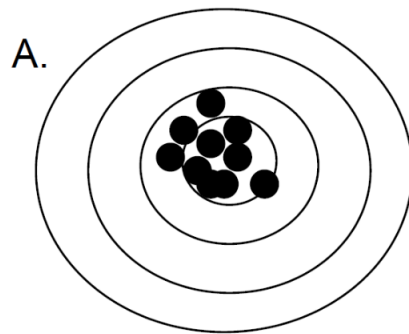
Nota: 10

```
library(knitr)
library(rio)
library(tidyverse)
library(rio)
library(DescTools)
library(emmeans)
library(modelbased)
library(see)
library(parameters)
library(survey)
library(SDaA)
library(radiant)
```

**Nombre del estudiante: Randalth Cristobal Medina Flores**

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 y preciso = exacto

B: no sesgado pero impreciso = inexacto

C: sesgado pero preciso = inexacto

D: sesgado e 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

#####¿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.

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

```
pinz$Especie <- factor(pinz$Especie)
```

```
pinz <- na.omit(pinz)
```

```
head (pinz)
```

```
##      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
```

```
medias <- pinz %>%
  group_by(Especie) %>%
  summarize(medias = mean(LargoDePico))
```

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

```
medias
```

```
## # A tibble: 2 x 2
##   Especie medias
##   <fct>     <dbl>
## 1 CRU.WAXB  7.38
## 2 WB.SPARG 10.7
```

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

```
alfa = 0.05
```

La especie con el pico mas largo es WB.SPARG, que en promedio tiene 10.737500, a diferencia de CRU.WAXB, que alcanza un largo de pico en promedio de 7.378571. El valor de  $p < 0.05$  por lo que se infiere que la diferencia entre el largo del pico entre las especies es significativo.

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

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

El tamaño del efecto de la diferencia del largo de pico entre las especies es de **3.358929** pudiendo inferir con un 95% de confianza que la verdadera diferencia entre las especies se encuentra entre: 3.667242 - 3.050615

####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.

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

## The following objects are masked from 'package:radiant.data':
##
##      month, wday

## The following objects are masked from 'package:lubridate':
##
##      hour, isoweek, mday, minute, month, quarter, second, wday, week,
##      yday, year

## 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
```

N =	100	1000	10000	1000000	10000000
n =	10	100	1000	10000	1000000
Error estándar =	15.81139	5	1.581139	0.5	0.05

c

```
##           N =      100 1000      10000 1000000 10000000
## 1:         n =       10  100       1000   10000   1000000
## 2: Error estándar = 15.81139    5 1.581139    0.5    0.05
```

```
ee_n10 <- c%>%
  summarise(ee = 50/sqrt(10))
```

```
ee_n10
```

```
##          ee  
## 1 15.81139
```

```
ee_n100 <- c%>%  
  summarise(ee = 50/sqrt(100))
```

```
ee_n100
```

```
##          ee  
## 1 5
```

```
ee_n1000 <- c%>%  
  summarise(ee = 50/sqrt(1000))
```

```
ee_n1000
```

```
##          ee  
## 1 1.581139
```

```
ee_n10000 <- c%>%  
  summarise(ee = 50/sqrt(10000))
```

```
ee_n10000
```

```
##          ee  
## 1 0.5
```

```
ee_n1000000 <- c%>%  
  summarise(ee = 50/sqrt(1000000))
```

```
ee_n1000000
```

```
##          ee  
## 1 0.05
```

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

A mayor precisión habrá un menor error estándar. En general, con tamaños de muestra grandes se logran estimados más precisos que con tamaños de muestra pequeños, y al no considerar la población en el cálculo del error estándar, este irá decreciendo a medida que aumente el número de muestras.