Evaluación 2

Diana Laura Javier Garcia

18/10/2020

1. Usted está encargado de determinar si un nuevo método de análisis es mejor que el método tradicional.

a) Presentar (en Geogebra o en R) una descripción estadística de las mediciones con los dos métodos (medidas de tendencia, varianza y graficas, histograma y boxplot)

En este trabajo se analizaran los datos correspondientes al documento "catalizador2". El cual consiste en dos muestras m_1 y m_2, cada una con 36 datos, los cuales se leen con las siguientes instrucciones, y además, se presentan la lectura de los datos de la población.

En este caso, la muestra 1 (m_1) se refiere a las mediciones tomada con un método tradicional, mientras que la muestra 2 (m_2) se refiere a datos tomados por un método nuevo.

```
m1<-read.table("catalizador2_m1.txt",header = F)
attach(m1)

m2<-read.table("catalizador2_m2.txt",header = F)
attach(m2)

## The following object is masked from m1:
##
##     V1

poblacion<-read.table("poblacion.txt",header = F)
attach(poblacion)

## The following object is masked from m2:
##
##     V1

## The following object is masked from m1:
##
##     V1
```

A continuación se presenta la descripción estadistica de las mediciones para la población, y cada una de las muestras.

Las medidas de tendencia central fueron calculadas con la instrucción:

```
{r} # summary(x) #
```

Mientrás, que las medidas de dispersión fuerón calculadas de la siguiente manera:

Varianza:

```
{r} # var(x) #
```

Desviación estandar:

```
{r} # sd(x) #
```

O bien:

```
{r} # sapply(x,sd) #
```

1. Población

Para estudiar la población se estudiaron la dos muestras obtenidas por ambos métodos en un solo archivo. Los parametros estadisticos de la población se presentan a continuación

```
summary(poblacion)

##        V1
##  Min.   :59.20
##  1st Qu.:70.60
##  Median :75.15
##  Mean   :75.61
##  3rd Qu.:80.15
##  Max.   :94.10
```

Las medidas de dispersión de la población son las siguientes:

Varianza:

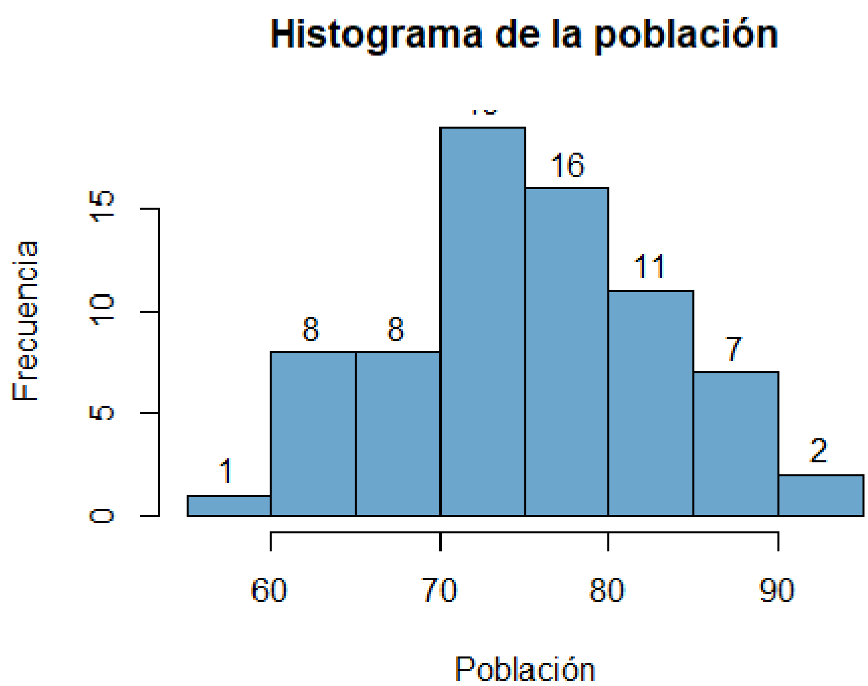
```
 var(poblacion)

##           V1
## V1 60.93574
```

Desviación estandar:

```
sapply(poblacion,sd)

##       V1
## 7.806135
```

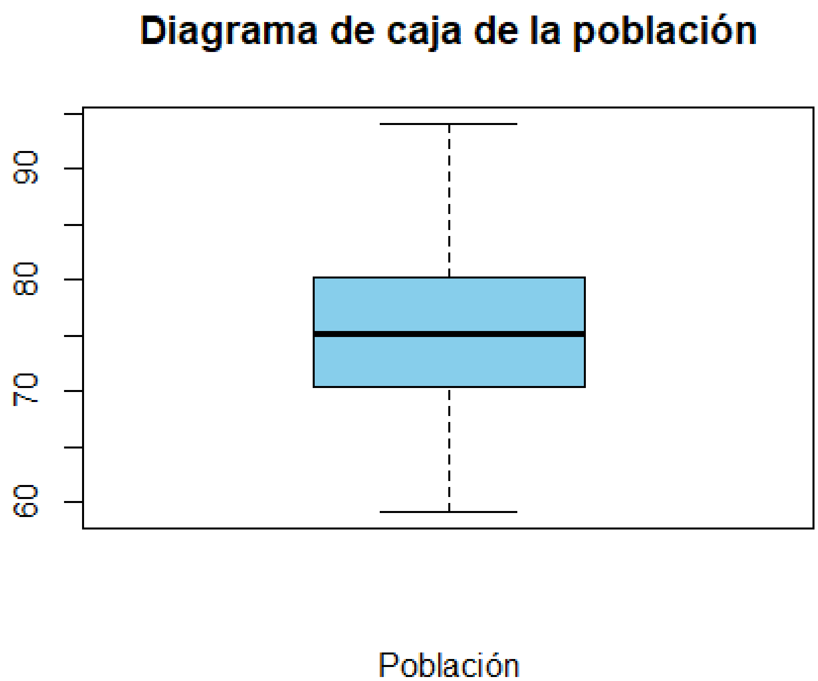
Por otra parte, el histograma de la población se generó con la siguiente instrucción:

```
hist(poblacion$V1, main = "Histograma de la población", ylab =
"Frecuencia", xlab = "Población",col = "skyblue3", labels = TRUE)
```



Mientras que el diagrama de caja, fue obtenido de la siguiente manera:

```
boxplot(poblacion, main="Diagrama de caja de la población",
xlab="Población",col = "skyblue", labels=TRUE)
```



De los parámetros presentados en el resumen de la población se observa que la media de los datos es 75.61 con una varianza importante de 60.93. Partiendo de que el mínimo de los datos es 59.20 y el máximo es 94.10, se infiere que el 25% de los datos se encuentran en entre 59.20 y 70.60, el 50% de los datos se encuentran entre 59.20 y 75.1, y el 75% se encuentran entre 59.20 y 80.15. Se observa además que la mediana y la media difieren por decimas, por lo que se pueden considerar como iguales. Del histograma presentado, con ocho clases se observa que los datos se conccentran en la región entre 70 y 80, además a partir de diagrama de caja se observa que existe una dispersión de datos importante de aproximadamente siete unidades.

2. Muestra 1 (referentes al método tradional)

Con respecto a los datos de la muestra 1, obtenidos por el método tradicional se obtuvieron los parámetros estadísticos:

```
summary(m1)
```

```
##        V1
##  Min.   :62.50
##  1st Qu.:68.65
##  Median :72.60
##  Mean   :73.02
##  3rd Qu.:78.47
##  Max.   :89.00
```

Con una varianza y desviación estandar de:

Varianza:

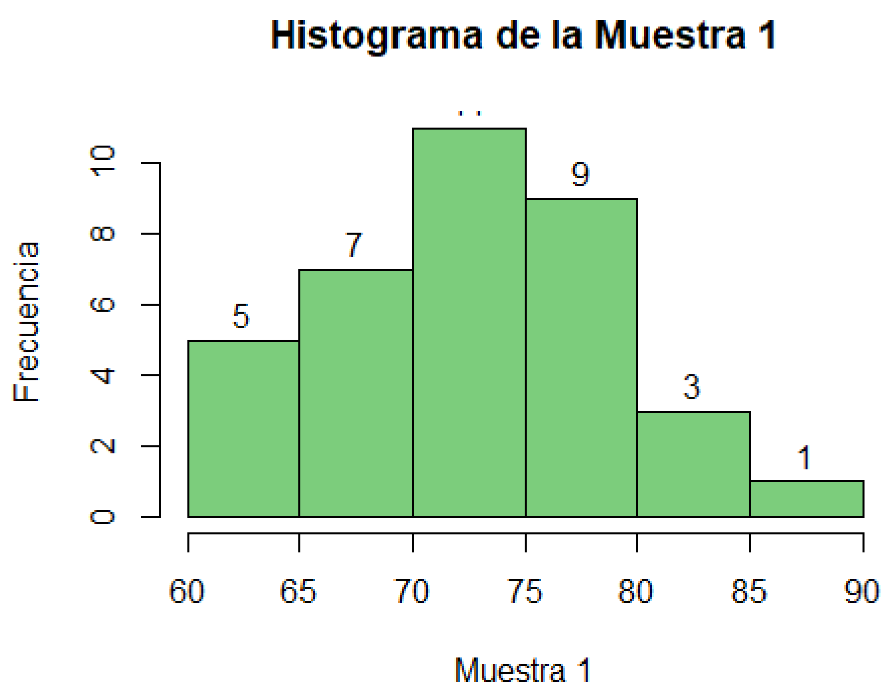
```r
var(m1)

##          V1
## V1 38.79075
```

Desviación estandar:

```r
sapply(m1,sd)

##       V1
## 6.228222
```

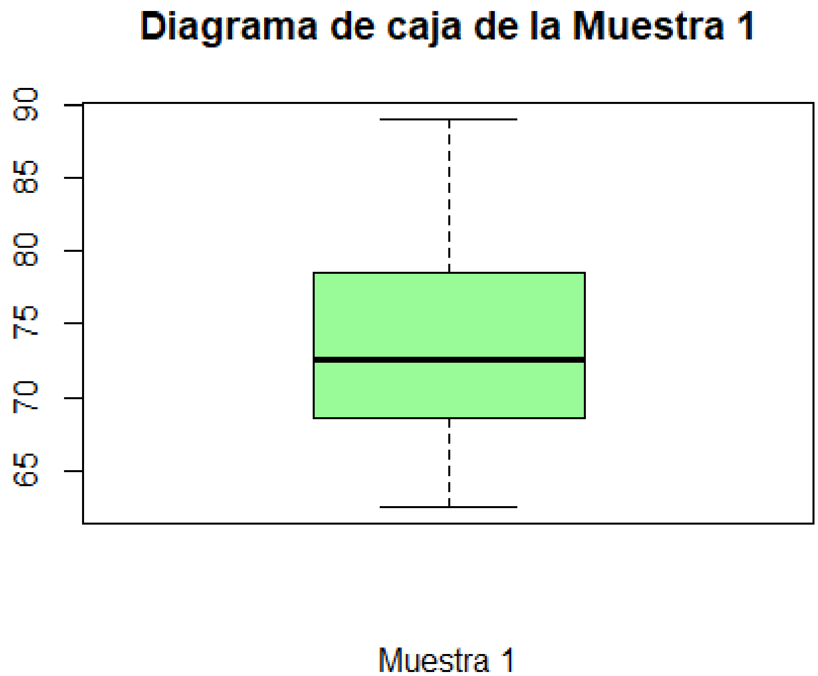
El histograma de la muestra 1 se generó con la siguiente instrucción:

```r
hist(m1$V1, main = "Histograma de la Muestra 1", ylab = "Frecuencia",
xlab = "Muestra 1",col = "palegreen3", labels = TRUE)
```



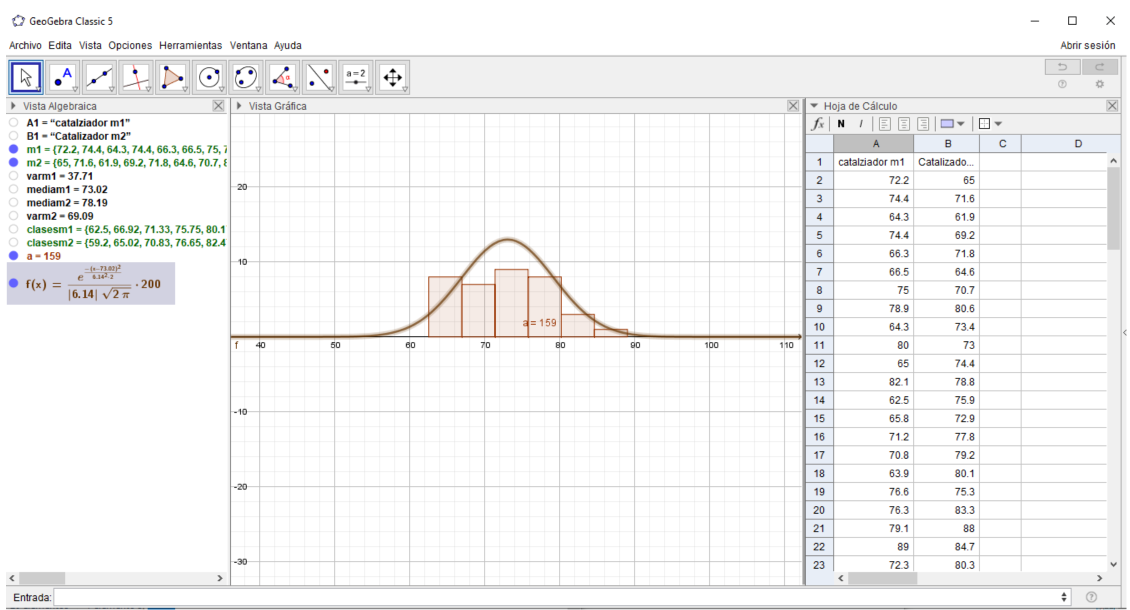
Mientras que el diagrama de caja, fue obtenido de la siguiente manera:

```r
boxplot(m1, main="Diagrama de caja de la Muestra 1", xlab="Muestra 1",col
= "palegreen", labels=TRUE)
```



Finalmente se ajusto una curva normal en geogebra:

```r
knitr::include_graphics("normalm1.png" )
```



De los parámetros estadisticos de la muestra 1 correspondientes al método tradicional se tienen las siguientes observaciones:

La media es de 73.020 con una varianza de 38.79 y una desviación estandar de 6.22. Dado que el minimo es 62.50 y el máximo es 89.00, se infiere que el 25% de los datos se encuentran entre 62.50 y 68.65, el 50% entre 62.50 y 72.60 y el 75% de los datos ente 62.50 y 78.47. Por otra parte, el histograma presentado muestra que la clase ente 70 y 75 es la que mayor frecuencia tiene, y además le corresponde una distribución normal.

3. Muestra 2 (referentes al método nuevo)

Con respecto a la muestra dos, obtenidos por el método nuevo se obtuvieron las parámetros estadísticos:

```
summary(m2)

##        V1
##  Min.   :59.20
##  1st Qu.:72.97
##  Median :79.35
##  Mean   :78.19
##  3rd Qu.:84.55
##  Max.   :94.10
```

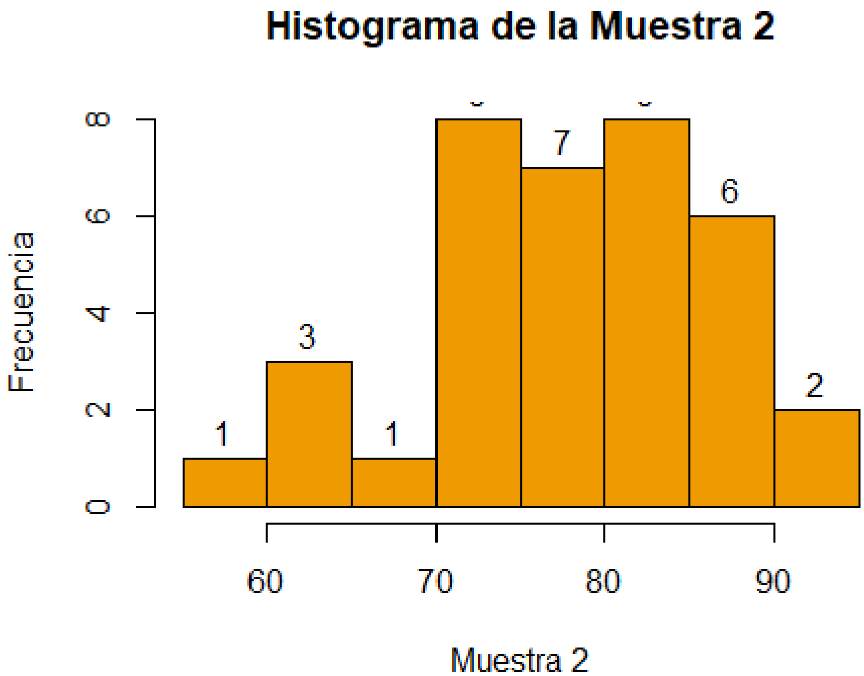
Con una varianza

```
var(m2)

##          V1
## V1 71.06364

sapply(m2,sd)

##       V1
## 8.429925
```

Con una desviación estandar:

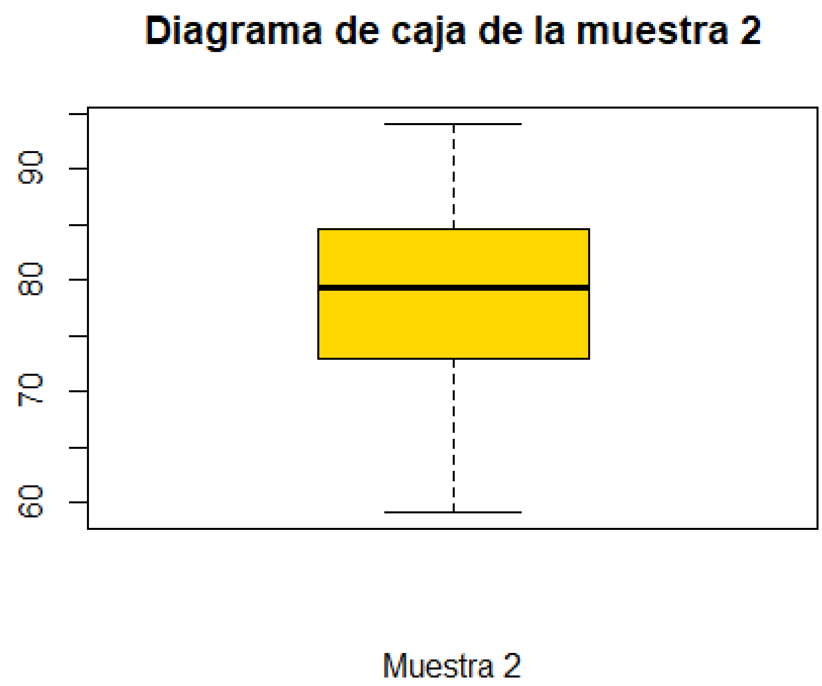
Por otra parte, el histograma de la muestra 2 se genero con la siguiente instrucción:

```
hist(m2$V1, main = "Histograma de la Muestra 2", ylab = "Frecuencia",
xlab = "Muestra 2",col = "orange2", labels = TRUE)
```



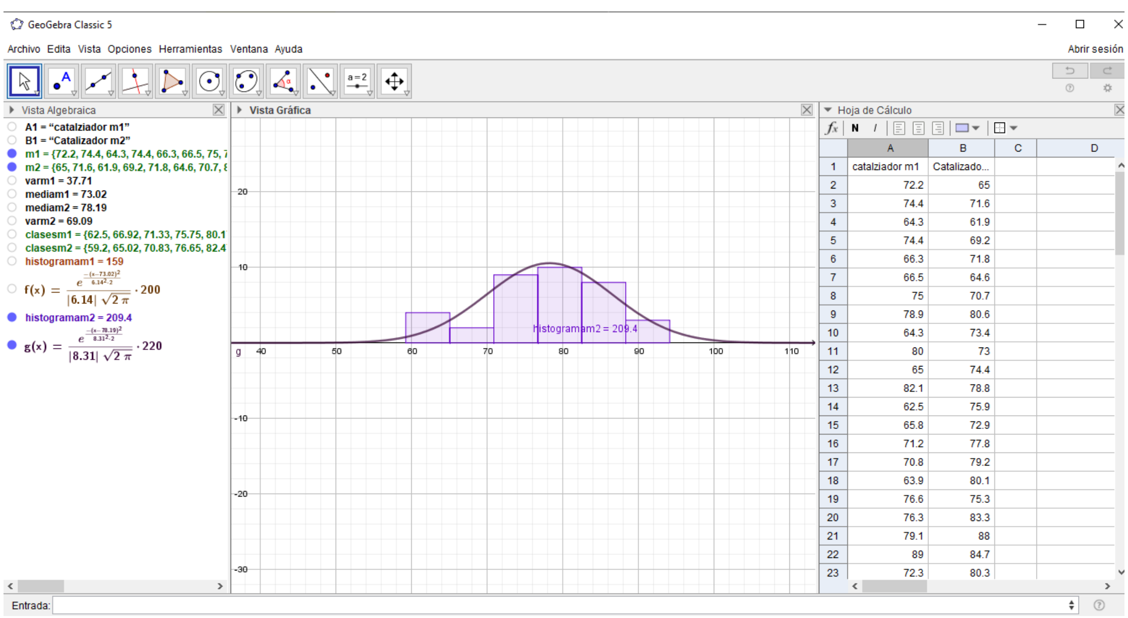
Mientras que el diagrama de caja, fue obtenido de la siguiente manera:

```
boxplot(m2, main="Diagrama de caja de la muestra 2", xlab="Muestra 2",col
= "gold", labels=TRUE)
```



El ajuste de la distribución normal se muestra a continuación:

```r
knitr::include_graphics("normalm2.png" )
```



A partir de los parametros obtenidos para la muestra 2, se tienen las siguientes observaciones:

La muestra dos, correspondiente a un método nuevo, tiene una media de 78.19, con una varianza y una desviación estandar de 71.06 y 8.42 respectivamente. El 25% de los datos se concentra entre 59.20 y 72.97, el 50% entre 59.20 y 79.35, y el 75% entre 59.20 y 84.55. A partir del histograma se observa que los datos se concentran entre 70 y 90, sin embargo, para seis numeros de clase, la distribución normal se ajusta.

Los datos se resumen en la siguiente tabla:

Tabla 1. Descripción estadistica de la población, la muestra 1 (método tradicional) y la muestra 2 (método nuevo)

Parametros	**Población**	**Muestra 1**	**Muestra 2**
Mínimo	59.20	62.50	59.20
1er cuartil	70.60	68.65	72.97
Mediana	75.15	72.60	79.35
Media	75.61	73.02	78.19
3er cuartil	80.15	78.47	84.55
Máximo	94.10	89.00	94.10
Varianza	60.93	38.79	71.06
Desviación estandar	7.80	6.22	8.42

Como se puede observar en la Tabla anterior, los valores de la población y las muestras difieren por dos y cinco unidades entre población y muestras, y lo mismo ocurre con los demás parámetros estadisticos, salvo la muestra 2 que presenta una varianza mayor comparada con la de la muestra 1. Se observó además que a cada muestra le corresponde una distribución normal. De manera general, se puede concluir que aunque ambos métodos proporcionan una media similar, la desviación estandar de la muestra dos es más grande, dicho de otra manera, el empleo del método nuevo genera tienes desviaciones importantes en las mediciones.

b) En R: Para cada muestra hacer una estimación puntual del promedio de la población y una estimación por intervalo del promedio poblacional a un nivel de confianza del 95 (suponer que no se conoce la desviación estándar ni la varianza de la población). Compararar los resultados usando una Distribución Normal y la Distribución t Student. Presentar los resultados en una tabla ¿Cuáles son las conclusiones?

Definiciónes:

1. Un estimador puntual está determinado por un número. El estimador del promedio, o de la media, esta dado por: $\overline{x} = \frac{\Sigma x_i}{n}$ donde n es el número de datos de la muestra.

2. Un estimador de intervalo está determinado por dos números entre los cuales se encuentra el parametro poblacional. A continuación se describe el intervalo de confianza para la media μ con σ desconocida para una población con variable aleatoria X con distribución $X \rightarrow N(\mu, \sigma)$.

Tomando como estimador de la media poblacional la media muestral, el estadistico muestral Ut,v sigue un modelo t de Student con $n-1$ grado de libertad:

$$U_{t,v} = \frac{\overline{X}-\mu}{\frac{S}{\sqrt{n}}} \rightarrow t_{n-1}$$

Dado un nivel de confianza $(1-\alpha)$ el intervalo de confianza para la media a este nivel es:

$$[\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}]$$

Donde S es la cuasivarianza muestral:

$$s = \sqrt{\frac{(\Sigma(x_i) - \overline{X})^2}{n-1}}$$

Es importante mencionar que cuando el tamaño muestral es grande, la t de Student se aproxima a una normal.

Si además, se supone que la población sigue una variable aleatoria X con distribución $X \rightarrow N(\mu, \frac{\sigma}{\sqrt{n}})$. La variable estandarizada sigue un modelo N(0,1):

$Z= N(0,1) $

Dado un nivel de confianza $(1-\alpha)$ podemos encontrar en la distribución los valores que encierran el centro de la distribución un área igual $(1-\alpha)$:

$P(-z_{\alpha/2}< Z < z_{\alpha/2}})=P(-z_{\alpha/2} < \frac{\bar{X}-\mu} \frac{\sigma}{\sqrt{n}} <z_{\alpha/2}})$

$$[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

Para la muestra 1 se obtuvo lo siguiente:

a) Estimador puntual del promedio de la población (muestra 1)

Dado que en el inciso anterior se obtuvieron los estadisticos muestrales, como el promedio, entonces el estimador puntial del promedio de la población para la muestra 1 $\overline{x_1}$ es el siguiente:

```
mean(m1$V1)
```

```
## [1] 73.01944
```

$\Rightarrow \overline{x_1} = \frac{\Sigma x_i}{n} = 73.01$

b) Estimador de intervalo del promedio poblacional a 95%

El enunciado indica que no se conoce la desviación estándar ni la varianza de la población, por lo tanto las ecuaciones que se deben emplear son las siguientes:

Cuasivarianza muestral s: $s = \sqrt{\frac{\Sigma(x_i - \overline{X})^2}{n-1}}$

$S = \sqrt{\frac{n}{n-1} Var_{muestra}}$$

Intervalo de confianza: $[\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}]$

Para el nivel de confianza del 95%, se tiene el intervalo de confianza para la media al nivel:

$(1-\alpha) = 1 - 0.95 = 0.05$

Además:

$$S = \sqrt{\frac{n}{n-1} Var_{muestra}} = \sqrt{\frac{36}{36-1}(38.79)} = 6.31$$

Para evaluar $t_{\frac{\alpha}{2},\nu}$ se debe tomar en cuenta la Distribución t de Student, la cual es simétrica,en este caso $\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$ con $\nu = n - 1$ grados de libertad, es decir: $\nu = 36 - 1 = 35$. De esta manera, se busca en la tabla la intersección de la fila $\nu = 35$ y $\frac{\alpha}{2} = 0.025$.

```
knitr::include_graphics("Student.png" )
```

Tabla t-Student

Grados de libertad	0.25	0.1	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.8165	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.7407	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923

Así: $t_{\frac{\alpha}{2},\nu}$ =2.0301

Por lo tanto, el intervalo de confianza: $[\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}]$

$$[73.01 - 2.0301\frac{6.31}{\sqrt{36}}, 73.01 + 2.0301\frac{6.31}{\sqrt{36}}]$$

$$[73.01 - 2.1349, 73.01 + 2.1349] = [70.8751, 75.144]$$

Además, se calculo el intervalo de confianza a partir de la Distribución Normal.

$[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$

$[73.01 - (1.96)\frac{6.31}{\sqrt{36}}, 73.01 + (1.96)\frac{6.31}{\sqrt{36}}]$

$[73.01 - 2.0612, 73.01 + 2.0612]$

$\Rightarrow [70.9488, 75.0712]$

Para la muestra 2 se obtuvo lo siguiente:

a) Estimador puntual del promedio de la población (muestra 2)

De la misma manera que para la muestra 1, en el inciso anterior se obtuvieron los estadisticos muestrales, entre ellos el promedio, entonces el estimador puntual del promedio de la población para la muestra 1 $\overline{x_2}$ es el siguiente:

```
mean(m2$V1)

## [1] 78.19167
```

$\Rightarrow \overline{x_2} = \frac{\Sigma x_i}{n} = 78.19$

b) Estimador de intervalo del promedio poblacional a 95%

El enunciado indica que no se conoce la desviación estándar ni la varianza de la población, por lo tanto las ecuaciones que se deben emplear son las siguientes:

$S = \sqrt{\frac{n}{n-1} Var_{muestra}}$$

Intervalo de confianza: $[\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}]$

Para el nivel de confianza del 95%, se tiene el intervalo de confianza para la media al nivel:

$(1 - \alpha) = 1 - 0.95 = 0.05$

Además:

$S = \sqrt{\frac{n}{n-1} Var_{muestra}} = \sqrt{\frac{36}{36-1}(71.06)} = 8.549$

Para evaluar $t_{\frac{\alpha}{2},\nu}$ se debe tomar en cuenta la Distribución t de Student, la cual es simétrica,en este caso $\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$ con $\nu = n - 1$ grados de libertad, es decir: $\nu = 36 - 1 = 35$. De esta manera, se busca en la tabla la intersección de la fila $\nu = 35$ y $\frac{\alpha}{2} = 0.025$.

Así: $t_{\frac{\alpha}{2},\nu}$ =2.0301

Por lo tanto, el intervalo de confianza: $[\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}]$

$$[78.19 - 2.0301\frac{8.549}{\sqrt{36}}, 78.19 + 2.0301\frac{8.549}{\sqrt{36}}]$$

$$[78.19 - 2.8925, 78.19 + 2.8925] = [75.3, 81.08]$$

Además, se calculo el intervalo de confianza a partir de la Distribución Normal.

$$[\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

$$[78.19 - (1.96)\frac{8.549}{\sqrt{36}}, 78.19 + (1.96)\frac{8.549}{\sqrt{36}}]$$

$$[78.19 - 2.7916, 78.19 + 2.7916]$$

$$\Rightarrow [75.399, 80.981]$$

Los resultados se resumen en la tabla 2:

Tabla 2. Estimación puntual y por intervalo del promedio de la población de las muestras

Estimador	**Muestra 1**	**Muestra 2**
Estimador puntual	73.01	78.19
Intervalo t Student	[70.875,75.144]	[75.3,81.08]
Intervalo D. Normal	[70.948,75.071]	[75.399,80.981]

Con se puede observar en la Tabla anterior, las estimaciones por intervalo del promedio con Distribución de Student y Normal no difieren entre sí en cada una de las muestras, lo cual sugiere que se puede emplear cualquiera de estas dos distribuciones. Es importante notar además que, los intervalo de la muestra 2 toman valores mayores que los de la muestra 1, esto se puede deber a que la muestra tiene una varianza mayor.

c) En R. Hacer una estimación por intervalo de la desviación estándar y de la varianza de cada muestra a un nivel de confianza del 95%.

Definición

Intervalo de confianza para la varianza de una población: Sea una población sobre la que se observa una variable aleatoria X con distribución $x \rightarrow N(\mu, \sigma)$, tomando como estimador de la varianza poblacional la cuasivarianza muestra. El estadistico muestral Uchi sigue un modelo Chi-cuadrdo con n-1 grados de libertad.

$$U_{chi} = \frac{(n-1)S^2}{\sigma^2} \to \chi^2_{n-1}$$

Donde $S^2 = \frac{\Sigma(X_i - \overline{X})^2}{n-1}$

Dado un nivel de confianza $(1-\alpha)$ podemos encontrar en la distribución los valores que encierran en el centro de la distribución un área (probablidad) igual a $(1-\alpha)$. De esta manera, el intervalor de confianza para la varianza al nivel $(1-\alpha)$ es el siguiente:

$$[\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}]$$

a) Muestra 1

En el inciso anterior se calculó la cuasivarianza muestral s: $S = \sqrt{\frac{n}{n-1} Var_{muestra}} = \sqrt{\frac{36}{36-1}(38.79)} = 6.31$

$\Rightarrow S^2 = 39.816$

Entonces, el intervalo para un nivel de confianza del 95% con $n = 36$, está dado de la siguiente manera:

$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$

$$[\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}]$$

Donde $\chi^2_{\alpha/2} = \chi^2_{0.05/2} = \chi^2_{0.025} = 20.5693$

$\chi^2_{1-\alpha/2} = \chi_{1-0.05/2} = \chi_{0.975} = 53.2033$

```
qchisq(0.025,35)
## [1] 20.56938
qchisq(0.975,35)
## [1] 53.20335
```

$$\Rightarrow [\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}]$$

$$\Rightarrow [\frac{(36-1)*39.81}{53.2033}, \frac{(36-1)*39.81}{20.5993}]$$

$\Rightarrow [26.1891,67.6406]$

Por lo tanto, S^2 de la muestra 1 con un nivel de confianza del 95% se encuentra en el intervalo [26.1891,67.6406].

b) Muestra 2

Al igual que en el inciso anterior, se evaluaron $\chi^2_{\alpha/2}$ y $\chi^2_{1-\alpha/2}$ para $\nu = n - 1 = 36 - 1 = 354$. Donde:

$1 - \alpha = 0.95 \Rightarrow \alpha = 0.05$

S está dada por:

$$S = \sqrt{\frac{n}{n-1} Var_{muestra}} = \sqrt{\frac{36}{36-1}(71.06)} = 8.549$$

$\Rightarrow S^2 = 73.090$

El intervalo está dado por:

$$[\frac{(n-1)S^2}{\chi^2_{\alpha/2}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}]$$

Donde $\chi^2_{\alpha/2} = \chi^2_{0.05/2} = \chi^2_{0.025} = 20.5693$

$\chi^2_{1-\alpha/2} = \chi_{1-0.05/2} = \chi_{0.975} = 53.2033$

```
qchisq(0.025,35)

## [1] 20.56938

qchisq(0.975,35)

## [1] 53.20335
```

$$\Rightarrow [\frac{(36-1) * 73.090}{53.2035}, \frac{(36-1) * 73.090}{20.5993}]$$

$\Rightarrow [48.082,124.186]$

Por lo tanto, S^2 de la muestra 2 con un nivel de confianza del 95% se encuentra en el intervalo [48.082,124.186].

Comparando los dos intervalos de confianza de la varianza de la muestra y de la muestra 2, se observa que el intervalo de la muestra 2 tiene una amplitud mayor comparada con la muestra 1, esto se debe a que, como se presentó en el primer inciso, la muestra 2 correspondiente al método nuevo, tiene una mayor varianza.

Definición

El intervalo de confianza correspondiente para la desviación estandar σ está dado por las raices cuadradas de los límites de confianza para la varianza:

$$\sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2}}} < \sigma < \sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}}}$$

a) Para la muestra 1 con $\nu = 36 - 1 = 35$ grados de libertad y $s = 6.31$ con intervalo de la varianza S^2:

$26.1891 < S^2 < ,67.6406$

Calculando la raíz cuadrada en el intervalo:

$\sqrt{26.1891} < S < \sqrt{67.6406}$

$\Rightarrow 5.117 < s < 8.224$

Por lo tanto, la muestra 1 con desviación estandar 6.31 tiene un intervalo de confianza: $5.117 < s < 8.224$.

b) Para la muestra 2 con $\nu = 36 - 1 = 35$ grados de libertad y $s = 8.541$ e intervalo de la varianza S^2:

$[48.082 < S^2 < 124.186]$

Se obtiene el intervalo de la desviación estandar calculando las raices cuadradas del intervalo de la varianza:

$\sqrt{48.082} < S < \sqrt{124.186}$

$\Rightarrow 6.934 < S < 11.143$

Por lo tanto, la muestra 2 con desviación estandar 8.541 tiene un intervalo de confianza $6.934 < S < 11.143$.

Como se puede observar, al igual que el intervalo de varianza, el intervalo de la desviación estandar del método dos es más amplió que el del método uno, esto se debe a que está relacionado directamente con la varianza.

A manera de conclusión se tienen las siguientes observaciones: la media obtenido en ambos métodos es similar, aunque difiere en cinco unidades, sin embargo, la varianza del método dos o nuevo es mucho mayor. De igual manera, los intervalos de confianza de la media, la varianza y la desviación estandar de la muestra dos son más amplios que los de la muestra una, por lo tanto se sugiere mantener las mediciones por el método tradicional.