

Tarea 3: Estimación

Instrucciones: Resuelva los siguientes ejercicios, si es necesario que haga ciertas suposiciones debe establecerlas desde el principio. Recuerde que en los ejercicios no sólo se solicita el cálculo, sino que debe hacer la interpretación de los resultados en función de la variable caracterizada en cada caso y entre ellos.

1. Se llevaron a cabo mediciones de la edad ósea en población pediátrica de la Ciudad de Querétaro por dos diferentes médicos pediatras expertos en el seguimiento de la maduración esquelética en dos diferentes tipos de imágenes: RX y DXA. (Valor 5 puntos)

Suposiciones y teoría previa para los cálculos:

Tenemos muestras de una población de niños y niñas de tamaño $n = 216$. Para la realización de los cálculos primero se especifica que los datos se toman como si tuvieran una distribución normal. El Teorema de Límite Central estipula que para muestras grandes con una distribución no normal se puede tomar como normal. Las muestras no dicen la información de las edades óseas de los niños y niñas. La edad ósea es un parámetro que nos dice la maduración de los huesos de un niño [1] o joven y a menudo se compara con la edad cronológica porque no son iguales en ocasiones.

Los estudios de imagen que se realizaron en este caso fueron RX y DXA.

Los rayos X son radiación electromagnética de alta energía (ionizante) que en este caso nos sirve para captar el contraste de los huesos [2].

Mientras que la técnica de radioabsorciometría de doble energía (DXA) consiste en emitir dos tipos de rayos X y, de acuerdo a la absorción de rayos, estudiar la densidad ósea del paciente [3].

Dichos estudios nos sirven para evaluar la salud ósea de un paciente.

La media y la varianza de las edades cronológicas son:

Dato	Edad cronológica	Intervalo de confianza 95%
Media	11.4875	10.96029 12.01471
Varianza	12.55899	10.29312 15.66943

La media y la varianza de las edades cronológicas en intervalos son:

- (De 5 a 9 años) 112 datos:

Dato	Edad cronológica	Intervalo de confianza 95%
Media	7.760982	7.495301 8.026663
Varianza	2.01335	1.573282 2.668926

- (De 10 a 17 años) 104 datos:

Dato	Edad cronológica	Intervalo de confianza 95%
Media	14.00769	13.60093 14.41446
Varianza	4.374789	3.388639 5.866573

- Calcule los dos intervalos de confianza para la media de las mediciones de cada observador para cada técnica de imagen. Comparé y discuta que es lo que observa.

Para el cálculo de los intervalos se realizó un programa en RStudio, el cual viene adjunto en la tarea. Se realizó de este modo para aprovechar toda la cantidad de datos y tener un error más pequeño. De igual modo, porque las tablas existentes por practicidad no tienen hasta tal grado de libertad.

Primero se instalaron las paqueterías readr, fdth, DescTools, readxl y EnvStats y se subió la tabla de Excel de los datos_EO. En mi caso, mir archivos son EO que contiene las medidas de los dos evaluadores para RX y DXA. Y otros dos archivos para los intervalos que se piden a en el segundo punto llamados E059 y E01017.

Luego se pasaron los datos de cada columna al tipo de arreglos.

Obtuvimos las varianzas y las medias para cada evaluador y cada técnica de imagen, dando lo siguiente:

Valor	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
Media	10.71019	11.44569	10.89949	11.00593
Varianza	18.88657	23.12252	10.89949	20.26273

Estos valores no son muy confiables a la hora de analizar los datos entonces necesitamos los intervalos de confianza. Para el IC de la media usamos la función t.test() con $p\text{-value} < 2.2e-16$. Quedando como resultado lo siguiente:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Media	10.12735 11.29303	10.80080 12.09059	10.30106 11.49792	10.40223 11.60963
Longitud del intervalo	1.16568	1.28979	1.19686	1.2074

El tamaño de los intervalos sobrepasa y es cercano a 1. Son más grandes los de DXA que los de RX.

Estos intervalos nos dicen la probabilidad del 95 % de tener a la media muestral en ese intervalo.

- Divida los datos en 2 intervalos de edad: de 5 a 9 años y de 10 a 17 años. Calcule los intervalos de confianza para la media de dichas observaciones. ¿Qué observa?

Para 5 a 9 años:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Media	6.816187 7.675599	7.057793 8.153814	6.856783 7.782324	6.964497 7.936395
Longitud del intervalo	0.859412	1.096021	0.925541	0.971898

El tamaño de las longitudes es pequeño, no pasan de 1 la mayoría. Además, son más grandes en el caso de DXA que de RX.

Para 10 a 17 años:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Media	13.93969 14.94223	15.11525 16.04667	14.27710 15.23252	14.33956 15.33025
Longitud del intervalo	1.00254	0.93142	0.95542	0.99069

El tamaño de las longitudes es menor que 1 y cercano a 1 aproximadamente. Además, son más grandes en el caso de DXA que de RX.

- Ahora calculé los intervalos de confianza para la diferencia de medias de las mediciones entre ambos observadores para cada técnica de imagen. Discuta sus hallazgos.

En este caso para la media, podemos restar elemento a elemento el arreglo de RX1 menos el de RX2 y lo mismo con DXA, y esto sacarle el intervalo de confianza (ya que la diferencia de medias es igual a la media de las diferencias). O bien, con la función `with(data, t.test(x,y,paired=TRUE))`.

Dato	Evaluador 1 RX y Evaluador 2 RX	Evaluador 2 DXA Y Evaluador 2 DXA
IC Diferencia de medias	-0.33840840 -0.04020271	0.2725562 0.6069809

El 0 no está en el intervalo de ninguna por lo que se puede considerar que sí hay diferencia entre las medias. Está más adentro el 0 el de RX que el de DXA.

- Si sus intervalos de confianza son grandes proponga un tamaño de muestra teórico para hacerlos más estrechos y viceversa. ¿Qué diría de la muestra total que le fue proporcionada?

Proponiendo un tamaño de muestra 100, cuyos datos son escogidos al azar, tenemos que:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Media	10.12735 11.29303	11.35988 12.82966	10.94357 12.28370	10.95207 12.31872
Longitud del intervalo	1.16568	1.46978	1.34013	1.36665

Vemos que si tomamos menos personas el intervalo se vuelve ligeramente más grande. Como el tamaño de los intervalos no cambia significativamente la muestra de 100 es representativa. Sin embargo, no podríamos tener tanta certeza al diagnosticar de anomalías no severas de la edad ósea de la muestra total.

Si en cambio, proponemos una muestra de 30 datos a azar obtenemos:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Media	9.002711 12.169955	10.45601 14.30532	9.056611 12.182056	8.266922 11.932411
Longitud del intervalo	3.167244	3.84931	3.125445	3.665489

La longitud del intervalo se hace demasiado grande si nos reducimos a 30 y es una muestra sesgada no elegible para sacar conclusiones certeras.

2. Se llevaron a cabo mediciones de la edad ósea en población pediátrica de la Ciudad de Querétaro por dos diferentes médicos pediatras expertos en el seguimiento de la maduración esquelética en dos diferentes tipos de imágenes: RX y DXA. (Valor 5 puntos)

- Calculé los dos intervalos de confianza para la varianza de las mediciones de cada observador para cada técnica de imagen. Comparé y discuta que es lo que observa.

Para el IC de la varianza usamos la función varTest().

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Varianza	15.76921 23.03457	19.30599 28.20084	16.62400 24.28319	16.91823 24.71297
Longitud del intervalo	7.26536	8.89485	7.65919	7.79474

- Divida los datos en 2 intervalos de edad: de 5 a 9 años y de 10 a 17 años. Calculé los intervalos de confianza para la varianza de dichas observaciones. ¿Qué observa?

Para 5 a 9 años:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Varianza	4.115566 6.981672	6.693668 11.355183	4.773291 8.097441	5.263417 8.928895
Longitud del intervalo	2.866106	4.661515	3.32415	3.665478

Para 10 a 17 años:

Dato	Evaluador 1 RX	Evaluador 1 DXA	Evaluador 2 RX	Evaluador 2 DXA
IC Varianza	5.146171 8.909295	4.441872 7.689978	4.673804 8.091510	5.025274 8.699991

- Ahora calculé los intervalos de confianza para el cociente de las varianzas de las mediciones entre ambos observadores para cada técnica de imagen. Discuta sus hallazgos.

Para esto, se usó la función with(data, var.test(x,y,paired)).

Dato	Evaluador 1 RX y Evaluador 2 RX	Evaluador 2 DXA y Evaluador 2 DXA
IC Cociente de varianzas	0.7255427 1.2401828	0.8728221 1.4919301

En este intervalo se encuentra el 1 por lo que se puede decir que las varianzas son iguales para los dos evaluadores en la respectiva técnica de imagen.

- Comparé y discuta todos sus hallazgos en relación con lo observado en los cálculos del ejercicio.

El hecho de que los intervalos en RX sean más pequeños que los de DXA nos puede decir que los de DXA son más “diversos” como es en la realidad e incluso puede revelarnos más información que el RX.

También pudimos observar que la varianza se redujo cuando dividimos los intervalos, y la edad cronológica es un factor que sí se debería considerar a la hora de evaluar los datos.

Si bien el intervalo de la diferencia de medias no incluyó a cero (pero sí cercano) en ninguno de los casos, el intervalo del cociente de varianzas dio 1. Por lo tanto, sí podemos sacar conclusiones a partir de los datos

En el caso de nuestra muestra, el intervalo de edad cronológica coincide con el de la edad ósea aproximadamente, es decir no difieren de la edad cronológica de 1 o más años lo cual es indicativo de la salud [4] y podemos concluir que la muestra de esta población no presenta alguna anomalía ósea del crecimiento en promedio. Si bien hay casos en los que sí difiere por más de un año, no es una constante en toda la muestra.

Referencias

[1] Medina, Martín & López. (2014). “*El uso de la edad ósea en la práctica clínica*”. Vol. 12. Núm. 6. DOI: 10.1016/S1696-2818(14)70204-5.

[2] Wikipedia. (2023). “*Rayos X*”. Recuperado el 14 de marzo de 2023 de https://es.wikipedia.org/wiki/Rayos_X#M%C3%A9dicas

[3] UCSF. (s.f.). Exploración de densidad ósea. Recuperado el 14 de marzo de 2023 de <https://radiology.ucsf.edu/sites/radiology.ucsf.edu/files/wysiwyg/patientcare/services/DXA/Bone-Density-Scan-DXA-espanol.pdf>

[4] Felicidad, S. (2016). “*Edad ósea: ¿qué es?, ¿para qué nos puede ser útil?*” Recuperado el 14 de marzo de 2023 de <https://www.familiaysalud.es/medicinas/pruebas-complementarias/edad-osea-que-es-para-que-nos-puede-ser-util#:~:text=La%20edad%20%C3%B3sea%20muestra%20la,hay%20un%20problema%20de%20crecimiento>

Apéndice

Código tarea 3 Estimación Bioestadística en R

library(readr)

library(fdth)

library(DescTools)

library(readxl)

library(EnvStats)

library(PairedData)

#####

EO <- read_excel("C:/Users/Admin/Downloads/EO.xlsx")

#View(EO)

Poner los datos en forma de listas

```

data<-data.frame(E0)
# Mostrar la variable Data
data
# Datos columna Evaluador 1 RX
RX1<-data$a1
# Datos columna
DXA1<-data$b1
# Datos columna
RX2<-data$a2
# Datos columna
DXA2<-data$b2
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 1 RX
t.test(RX1)
varTest(RX1)
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 1 DXA
t.test(DXA1)
varTest(DXA1)
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 2 RX
t.test(RX2)
varTest(RX2)
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 2 DXA
t.test(DXA2)
varTest(DXA2)
# Primer método propuesto para e IC de medias
# IC Diferencia de medias para RX
c <- RX1-RX2
t.test(c)
# IC Diferencia de medias para DXA
d <- DXA1-DXA2
t.test(d)
# IC para la diferencia de medias de RX
with(data, t.test(RX1,RX2,paired=TRUE))
# IC para la diferencia de medias de DXA

```

```

with(data, t.test(DXA1,DXA2,paired=TRUE))

# IC para el cociente de varianzas de RX

with(data, var.test(RX1,RX2,paired=TRUE))

# IC para el Cociente de varianzas de DXA

with(data, var.test(DXA1,DXA2,paired=TRUE))

#####

# Datos para el intervalo de 5 a 9 años

E059 <- read_excel("C:/Users/Admin/Downloads/E059.xlsx")

#View(E059)

# División en intervalos

# Poner los datos en forma de listas

data1<-data.frame(E059)

# Mostrar la variable Data

data1

# Datos columna Evaluador 1 RX intervalo 5-9 años

RX1m<-data1$ma1

# Datos columna Evaluador 1 DXA intervalo 5-9 años

DXA1m<-data1$mb1

# Datos columna Evaluador 2 RX intervalo 5-9 años

RX2m<-data1$ma2

# Datos columna Evaluador 2 DXA intervalo 5-9 años

DXA2m<-data1$mb2

# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 1 RX 5-9

t.test(RX1m)

varTest(RX1m)

# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 1 DXA 5-9

t.test(DXA1m)

varTest(DXA1m)

# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 2 RX 5-9

t.test(RX2m)

varTest(RX2m)

# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 2 DXA 5-9

```

```

t.test(DXA2m)
varTest(DXA2m)
#####

# Datos para el intervalo de 10 a 17 años
EO1017 <- read_excel("C:/Users/Admin/Downloads/EO1017.xlsx")
#View(EO1017)
data2<-data.frame(EO1017)
# Mostrar la variable Data
data2
# Datos columna Evaluador 1 RX intervalo 5-9 años
RX1M<-data2$Ma1
RX1M
# Datos columna Evaluador 1 DXAX intervalo 5-9 años
DXA1M<-data2$Mb1
# Datos columna Evaluador 2 RX intervalo 5-9 años
RX2M<-data2$Ma2
# Datos columna Evaluador 2 DXA intervalo 5-9 años
DXA2M<-data2$Mb2
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 1 RX 10 17
t.test(RX1M)
varTest(RX1M)
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 1 DXA 10 17
t.test(DXA1M)
varTest(DXA1M)
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 2 RX 10 17
t.test(RX2M)
varTest(RX2M)
# Intervalo de confianza para la media y la varianza al 95%. Para Evaluador 2 DXA 10 17
t.test(DXA2M)
varTest(DXA2M)
#####
# Para la edad cronológica
edad <- data$E

```


Intervalo de confianza para la media y la varianza al 95%. Para la edad

t.test(edad)

varTest(edad)

Edad cronológica en intervalos

De 5 a 9

edad5a9 <- data1\$E

edad5a9

mean(edad5a9)

var(edad5a9)

t.test(edad5a9)

varTest(edad5a9)

De 10 a 17

edad10a17 <- data2\$E

mean(edad10a17)

var(edad10a17)

t.test(edad10a17)

varTest(edad10a17)

#####

#Tamaño de muestra teórico 100 y 30.

azarRX1 <- sample(RX1, size = 30)

azarRX2 <- sample(RX2, size = 30)

azarDXA1 <- sample(DXA1, size = 30)

azarDXA2 <- sample(DXA2, size = 30)

IC para la media de la muestra aleatoria de las muestras.

t.test(azarRX1)

t.test(azarRX2)

t.test(azarDXA1)

t.test(azarDXA2)

IC para la varianza de la muestra aleatoria de las muestras.

varTest(azarRX1)

varTest(azarRX2)

varTest(azarDXA1)

varTest(azarDXA2)