

PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

FACULTAD DE CIENCIAS E INGENIERÍA



PC1

Curso:

ESTADÍSTICA PARA INGENIERÍA

Horario:

501

Elaborado por:

(20200315)	Camayo Arias, Jose Félix	20200315@pucp.edu.pe
(20195861)	Iman Loja, Luis Angel Luis	luis.imanl@pucp.edu.pe

Docente:

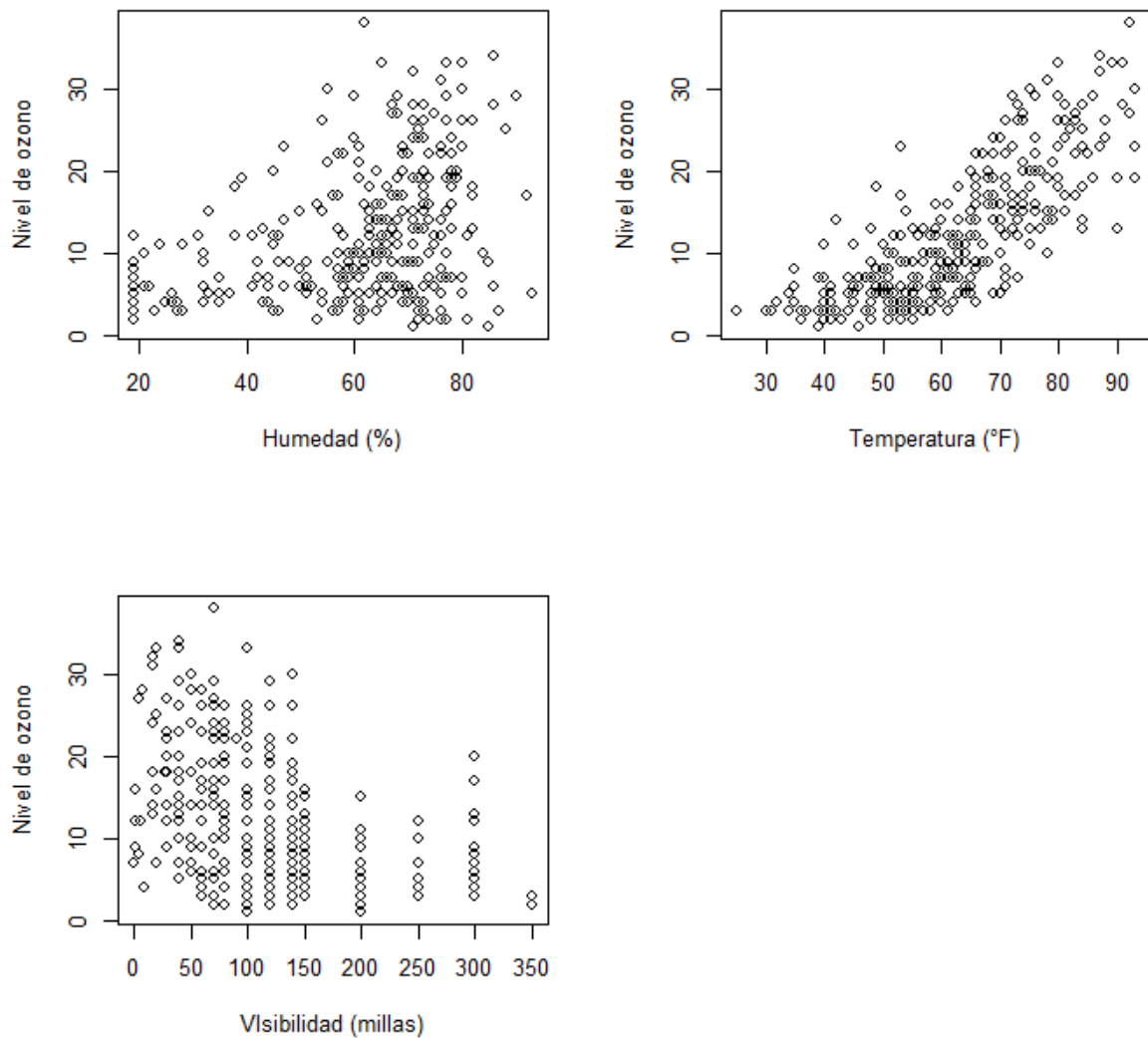
TARAZONA VARGAS , ENVER GERALD

Lima, 24 de enero del 2023

CUESTIONARIO:

A:

Grafica 1: Diagrama de dispersión del nivel de ozono con respecto a las variables meteorológicas



De acuerdo al Gráfico N° 1, podemos notar que existe una relación positiva entre el nivel de ozono y humedad, puesto a que a mayores niveles de humedad hay mayor concentración de datos con niveles de ozono altos; sin embargo, esta asociación da la impresión de ser débil. Asimismo, se puede resaltar que hay una relación directa entre el nivel de ozono y la temperatura, además, se aprecia visualmente una asociación lineal fuerte. Cabe destacar la presencia de observaciones inusuales ya que hubo días con temperaturas no tan altas donde el

nivel de ozono fue mayor. Por último, pareciese no haber una relación entre el nivel de ozono y la visibilidad puesto a que los datos se encuentran agrupados.

Tabla N° 1: Coeficientes de correlación de las distintas variables meteorológicas con respecto al nivel de ozono

	Humedad	Temperatura	Visibilidad
Coeficiente de Correlación	0.44922399	0.780702835	-0.44098946

A partir de la tabla N° 1, se puede apreciar el coeficiente de correlación de Pearson del nivel de ozono y las variables meteorológicas: Humedad, Temperatura y Visibilidad. Con respecto al primero, muestra una asociación positiva débil puesto a que su valor se aleja del grado máximo de correlación. De forma similar, el valor absoluto de la correlación con respecto a la tercera variable se aleja del grado máximo, por lo tanto, la relación es débil, cabe mencionar que la asociación con esta variable es inversa. Por el contrario, hay una relación lineal positiva con la temperatura la cual es más fuerte en comparación a las anteriores variables, este representaría una mejor relación directa y sería más útil para predecir el nivel de ozono con un modelo lineal.

B:

Se escogió el modelo lineal del nivel de ozono con respecto a la temperatura.

Modelo de regresión lineal para predecir el nivel medio de ozono:

$$\hat{Y} = 0.43257 X - 14.93745$$

\hat{Y} : Nivel de ozono
 X : Temperatura (°F)

De acuerdo al modelo mostrado, el coeficiente “-14,93745” es el valor medio esperado del nivel de ozono cuando la temperatura es 0 °F; sin embargo, este pareciese no tener sentido físico para su aplicación además que 0 °F esta fuera del dominio de interpolación del modelo obtenido, por ende, el comportamiento del nivel de ozono no puede ser explicado por este modelo alrededor de dicha temperatura. Por otro lado, la pendiente señala que, por cada variación de 1°F en la temperatura, el valor medio esperado del nivel de ozono cambia en 0.43257 unidades siguiendo una relación directa.

C:

Tabla N° 2: Coeficientes de correlación del nivel de ozono con respecto a la temperatura y al ITH

	Temperatura	ITH
Coeficiente de Correlación	0.78070283	0.79456849

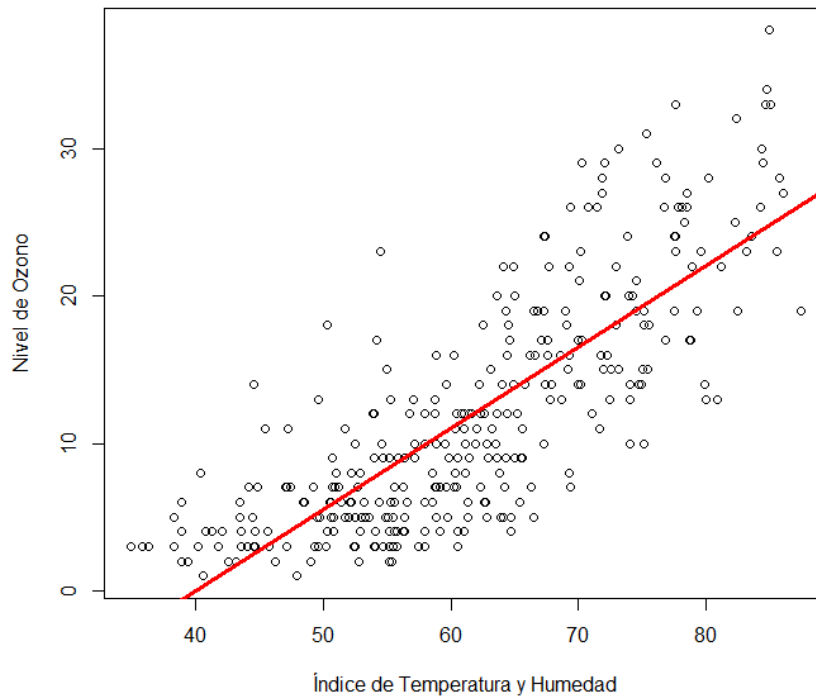
De acuerdo a los resultados obtenidos, se puede notar que tanto la temperatura como el Índice de Temperatura y Humedad (ITH) tienen una asociación lineal positiva con el nivel de ozono. Sin embargo, la correlación con el ITH tiene un valor ligeramente mayor que con la temperatura, esto significaría que una relación lineal es más fuerte en comparación con el modelo anteriormente estimado. Asimismo, cabe destacar que el ITH es un parámetro ambiental recomendado para predecir el nivel de ozono, puesto a que encuentra en función de la temperatura y humedad relativa, las cuales son factores que influyen en la formación de ozono. Por lo tanto, se debería utilizar el modelo que asocie estas variables para la estimación del nivel de este.

Modelo de regresión lineal para predecir el nivel medio de ozono:

$$\hat{Y} = 0.55238 X - 22.15511$$

\hat{Y} : *Nivel de ozono*
 X : *ITH*

Grafica N°2: Modelo de regresión lineal del nivel de ozono y el ITH



D:

Como se mencionó el mejor modelo es el que relaciona el nivel de ozono y el ITH, por lo tanto, para predecir el valor medio del nivel de ozono se utilizara dicho modelo para un día con las siguientes condiciones:

Humedad = 51%

Temperatura = 54°F

Visibilidad = 100 millas

A partir de las condiciones anteriores se calculó el ITH, cuyo valor se encuentra en el rango de valores normales de ITH:

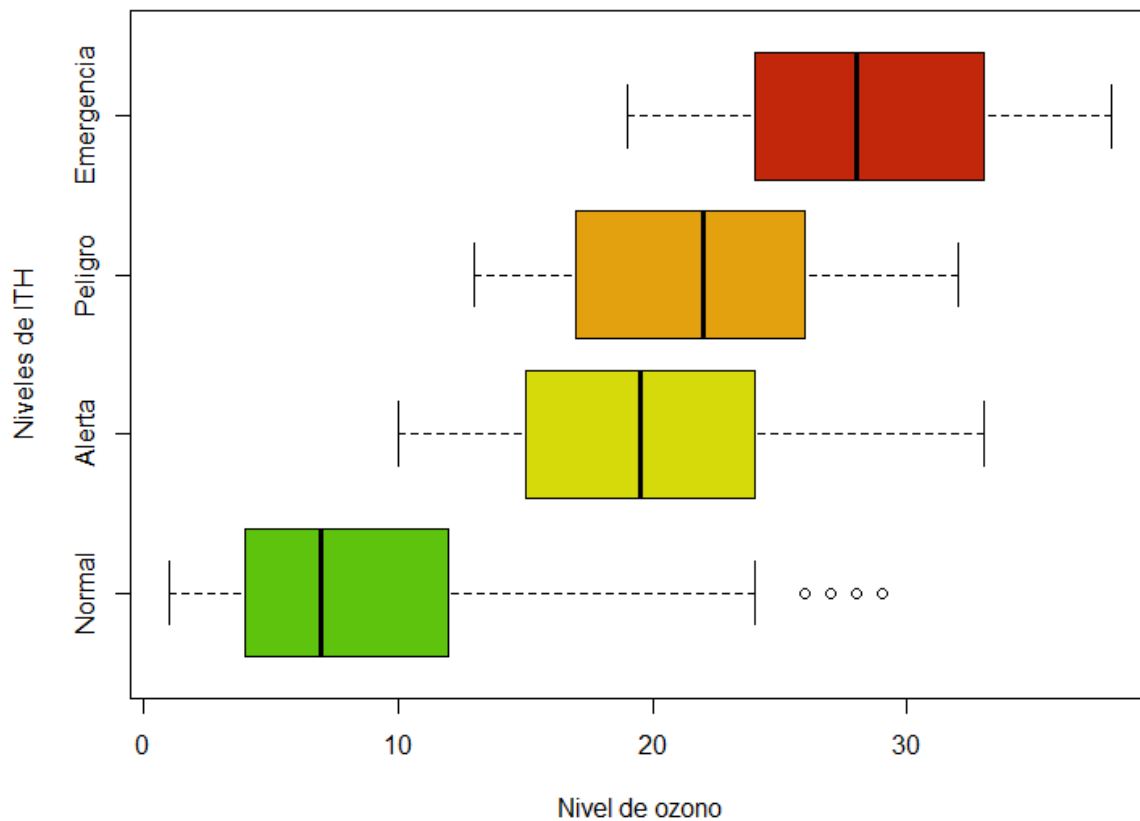
$$\therefore ITH = 55.078$$

Con el resultado anterior se estimó utilizando la función `predict` de R:

$$\rightarrow \text{Nivel medio de ozono estimado} = 8.269131$$

E:

Gráfica N°3: Nivel de ozono según los niveles de ITH



Del grafico N° 3, se puede observar que la mediana del nivel medio de ozono aumenta según el nivel de ITH. Esto significa, según el grafico, que el promedio del nivel de ozono es mayor según más crítico sea el nivel de ITH. Por otro lado, el rango intercuartílico pareciese ser similar para los todos los grupos de observaciones, es decir, la concentración del 50% de los datos centrales tiene similar distancia entre sus extremos; no obstante, según el rango, la distancia entre el mínimo y máximo nivel de ozono, se puede resaltar que la variabilidad de los datos disminuye conforme aumenta el nivel de ITH.

Para los días con nivel de ITH normal, se puede resaltar que hay una asimetría positiva, el cual sugiere que valores elevados del nivel de ozono son inusuales, por ello tiene sentido la presencia de observaciones atípicas. Para los días con nivel de ITH alerta, se aprecia una ligera asimetría positiva donde niveles muy elevados de ozono tienden a ser atípicos en estos días. Para los días con nivel ITH peligro, no se puede definir la asimetría solamente con el grafico, Por último, para días con nivel ITH emergencia existe una asimetría ligeramente positiva, por ende, los valores mayores tienden a ser inusuales; pero, debido a que la mediana del nivel de ozono es

elevada para esta categoría, se podría considerar que niveles de ozono son extremadamente elevados.

F:

Tabla N° 2: Medidas estadísticas del nivel de ozono para días de ITH normal

ESTADÍSTICOS	Nivel de ozono en días de ITH normal
Media	9.095
Mediana	7
Asimetría	1.097
Rango	28
Desviación estándar	5.93
Rango inter cuartil	8
Coficiente de variación	0.6519

Como se mencionó anteriormente, existe una asimetría positiva notable en las observaciones del nivel de ozono para los días con ITH normal, evidenciado en la presencia de valores atípicos a la derecha de la distribución. Por lo tanto, la mediana sería una medida de tendencia central más robusta para representar a las observaciones. Asimismo, el rango inter cuartil sería una medida de dispersión más adecuada para representar la variabilidad de los datos, ya que el rango se ve fuertemente afectado por la presencia de valores atípicos extremos.

ANEXO:

CODIGO R UTILIZADO PARA LA GENERACIÓN DE RESULTADOS

CODIGO PC2

2023-01-24

```
d=read.csv("C:\\Users\\thexn\\Downloads\\ESTADISTICA\\PC2\\ozono.csv",sep
= ";",
          stringsAsFactors = T)

attach(d)
library(agricolae)
library(EnvStats)

## A)

## Humedad (%)
OvsH<-cor(O3,humidity)

## Temperatura (°F)
OvsT<-cor(O3,temp)

## Visibilidad (millas)
OvsV<-cor(O3,vis)

##Graficas de dispersión

par(mfrow=c(2,2))

plot(humidity,O3,xlab = "Humedad (%)", ylab="Nivel de ozono")
plot(temp,O3,xlab = "Temperatura (°F)", ylab="Nivel de ozono")
plot(vis,O3,xlab = "Visibilidad (millas)", ylab="Nivel de ozono")

par(mfrow=c(1,1))
```

#Comparación entre coeficientes de correlación

```
tabla_a<-as.table(cbind(OvsH,OvsT,OvsV))
colnames(tabla_a)<-c("Humedad","Temperatura","Visibilidad")
rownames(tabla_a)<- "Coef. correlación"
tabla_a

##              Humedad Temperatura Visibilidad
## Coef. correlación 0.4492240    0.7807028   -0.4409895

write.table(x=tabla_a, file = "Tabla_a.txt", sep = ",",
           row.names = F, col.names = T)##PARA EXPORTAR TABLA
```


B) Modelo de O3 y La temperatura

```
rltemp<-lm(O3~temp)
summary(rltemp)

##
## Call:
## lm(formula = O3 ~ temp)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.9939  -3.8202  -0.1796   3.1951  15.0112
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -14.93745     1.21247  -12.32  <2e-16 ***
## temp         0.43257     0.01912   22.63  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.014 on 328 degrees of freedom
## Multiple R-squared:  0.6095, Adjusted R-squared:  0.6083
## F-statistic: 511.9 on 1 and 328 DF,  p-value: < 2.2e-16
```

C)

##Variable Indice de Temperatura y Humedad

```
ITH<-temp-(0.55-0.55*humidity/100)*(temp-58)
OvsITH<-cor(O3,ITH)
```

##Comparación de correlaciones del nivel de ozono con respecto al ITH y a la

##temperatura

```
tabla_c<-as.table(cbind(OvsT,OvsITH))
colnames(tabla_c)<-c("Temperatura","ITH")
rownames(tabla_c)<-c("Coef. correlación")
write.table(x=tabla_c, file = "Tabla_c.txt", sep = ",",
            row.names = F, col.names = T)##PARA EXPORTAR TABLA
```

Estimación del modelo lineal

```
rlITH<-lm(O3~ITH)
summary(rlITH)

##
## Call:
## lm(formula = O3 ~ ITH)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
##  -9.5892  -3.7121  -0.1893   3.0467  15.0736
##
## Coefficients:
```

```
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) -22.15511    1.45656  -15.21  <2e-16 ***
## ITH          0.55238     0.02331   23.70  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.872 on 328 degrees of freedom
## Multiple R-squared:  0.6313, Adjusted R-squared:  0.6302
## F-statistic: 561.7 on 1 and 328 DF,  p-value: < 2.2e-16

plot(ITH,O3,xlab = "Índice de Temperatura y Humedad", ylab="Nivel de Ozono",
     main="Grafica N°2: Modelo de regresión lineal del nivel de ozono \n y el ITH")
abline(r1ITH,col="red",lwd=3)
```

D)

```
est_ITH<-54-(0.55-0.55*51/100)*(54-58)
est_ITH<-data.frame(ITH=c(est_ITH))

predict(object = r1ITH, newdata = est_ITH)

##           1
## 8.269131
```

E)

```
#variable de Los Umbrales

umbralITH<-ITH

##Definition de Los umbrales

umbralITH[72>=ITH]="Normal"
umbralITH[ITH>72 & ITH<=78]="Alerta"
umbralITH[ITH>78 & ITH<=83]="Peligro"
umbralITH[ITH>83]="Emergencia"

##Jerarquia de Los niveles de ITH

orden<-c("Normal","Alerta","Peligro","Emergencia")
umbralITH <- factor(umbralITH,levels = orden,ordered = T)

##Diagrama de cajas del nivel de ozono según el nivel de ITH

boxplot(O3~umbralITH, horizontal = T, xlab= "Nivel de ozono",
        ylab = "Niveles de ITH",main="Gráfica N°3: Nivel de ozono según los niveles de ITH",
        col=c("#5EC30D", "#D6DA0A", "#E4A10F", "#C3270B"))
```

##F

```
tabla_f<-summaryFull(O3[umbralITH=="Normal"])  
write.table(x=tabla_f, file = "Tabla_f.txt", sep = ",",  
            row.names = T, col.names = T)##PARA EXPORTAR TABLA
```

```
knitr::opts_chunk$set(echo = TRUE)
```