

# REGRESIÓN MÚLTIPLE: COLINEALIDAD



Dra. Purificación Galindo  
Villardón

## PRÁCTICA: REGRESIÓN MÚLTIPLE RESUELTA

*The Federal Trade Commision* estudia cada año diferentes marcas de tabaco, y analiza su contenido en alquitrán, nicotina y monóxido de carbono. *The United States Surgeon General*, considera cada una de estas sustancias peligrosas para la salud de los fumadores. Estudios ya realizados, ponen de manifiesto que incrementos en el contenido de alquitrán y nicotina de los cigarrillos vienen acompañados por incrementos en el monóxido de carbono emitido al fumar.

Los datos que se presentan aquí han sido tomados de *Mendenhall and Sincich* (1992).

Para más información, ver el artículo “Using Cigarette Data for an Introduction to Multiple Regression”, by Lauren McIntyre in Volume 2, Number 1, of the *Journal of Statistics Education*.

Los datos son los siguientes:

CLASE	ALQUITRAN	NICOTINA	PESO	MONÓXIDO
RUBIO	4,10	0,860	0,9853	13,60
NEGRO	16,00	1,060	1,0938	16,60
NEGRO	29,80	2,030	1,1650	23,50
RUBIO	8,00	0,670	0,9280	10,20
RUBIO	4,10	0,400	0,9462	5,40
NEGRO	15,00	1,040	0,8885	15,00
RUBIO	8,80	0,750	1,0267	9,00
RUBIO	12,40	0,950	0,9225	12,30
NEGRO	16,60	1,120	0,9372	16,30
RUBIO	14,90	1,020	0,8858	15,40
RUBIO	13,70	1,010	0,9643	13,00
RUBIO	15,10	0,900	0,9316	14,40
RUBIO	7,80	0,570	0,9705	10,00
RUBIO	11,40	0,780	1,1240	10,20
RUBIO	9,00	0,740	0,8517	9,50
RUBIO	1,00	0,130	0,7851	1,50
NEGRO	17,00	1,260	0,9186	18,50
RUBIO	12,80	1,080	1,0395	12,60
NEGRO	15,80	0,960	0,9573	17,50
RUBIO	4,50	0,420	0,9106	4,90
NEGRO	14,50	1,010	1,0070	15,90
RUBIO	7,30	0,610	0,9806	8,50
NEGRO	8,60	0,690	0,9693	10,60
NEGRO	15,20	1,020	0,9496	13,90
RUBIO	12,00	0,820	1,1184	14,90

La variable dependiente es ***Y: contenido en monóxido de carbono***

Las regresoras (variables explicativas) son:

***X<sub>1</sub>: contenido de alquitrán***

***X<sub>2</sub>: contenido en nicotina***



*X<sub>3</sub>: peso del cigarrillo*  
*X<sub>4</sub>: clase de tabaco*

## 1) PRIMERA TAREA: REGRESIÓN SIMPLE

La primera tarea consiste en ajustar un modelo de regresión simple con cada una de las 3 variables explicativas que son cuantitativas (*X<sub>1</sub>: contenido de alquitrán*, *X<sub>2</sub>: contenido en nicotina*, *X<sub>3</sub>: peso del cigarrillo*) e interpretar los coeficientes de regresión resultantes. Escriba los modelos y los coeficientes de determinación correspondientes e interprete brevemente los resultados más relevantes.

Para cada modelo, hacer una regresión lineal:  
 Expresar el modelo y el coeficiente de Determinación ( $R^2$ )

- Modelo *Monóxido de carbono/ Alquitrán*:

$$\text{Monóx.Carbono} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Alquitrán}$$

Resumen del modelo		
Modelo	R Coef. Correlación	R cuadrado Coef. Determinación
1	0,907 <sup>a</sup>	0,823

Coeficientes			
Modelo		Coeficientes no estandarizados	
		B	Error estándar de B
1	(Constante)	3,883	,930
	Alquitrán	0,732	,071

Recogiendo la información anterior tenemos:

$$\text{Monóx.Carbono} = 3,88 + 0,73 \cdot \text{Alquitrán}$$

$R^2 = 0,82$

El valor  $R^2$  vale 0,82, lo que indica que si conocemos el contenido de *Alquitrán* de un cigarrillo, podemos mejorar en un 82% nuestros pronósticos sobre el contenido en *Monóxido de carbono*; es decir, si en lugar de utilizar como pronóstico el valor medio del contenido en *Monóxido de carbono*, basamos nuestro pronóstico en el contenido en *Alquitrán*, mejora la estimación en un 82%.

Interpretando el término independiente y el coeficiente de regresión para la variable *Alquitrán* podemos afirmar:

- Si el contenido en *Alquitrán* de los cigarrillos fuese 0 mg, cabría esperar una cantidad de *Monóxido de carbono* de 3,88 mg.

- Por cada incremento de 1 mg en el contenido de *Alquitrán* cabe esperar que la cantidad de *Monóxido de carbono* producido se incremente en 0,73 mg.
- Modelo *Monóxido de carbono/ Nicotina*:

$$\text{Monóxido de Carbono} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Nicotina}$$

Resumen del modelo		
Modelo	R Coef. Correlación	R cuadrado Coef. Determinación
1	0,926	0,858

Coeficientes			
Modelo		Coeficientes no estandarizados	
		B	Error estándar de B
1	(Constante)	1,668	,990
	alquitrán	12,397	1,050

Recogiendo la información anterior tenemos:

$$\text{Monóxido de Carbono} = 1,67 + 12,40 \cdot \text{Nicotina}$$

$R^2 = 0,86$

El valor  $R^2$  vale 0,86, lo que indica que si conocemos el contenido de *Nicotina* de un cigarrillo, podemos mejorar en un 86% nuestros pronósticos sobre el contenido en *Monóxido de carbono*; es decir, si en lugar de utilizar como pronóstico el valor medio del contenido en *Monóxido de carbono*, basamos nuestro pronóstico en el contenido en *Nicotina*, mejora la estimación en un 86%.

Interpretando el término independiente y el coeficiente de regresión para la variable *Nicotina* podemos afirmar:

- Si el contenido en *Nicotina* de los cigarrillos fuese 0 mg, cabría esperar una cantidad de *Monóxido de carbono* de 1,67 mg.
- Por cada incremento de 1 mg en el contenido de *Nicotina* cabe esperar que la cantidad de *Monóxido de carbono* producido se incremente en 12,40 mg.

- Modelo *Monóxido de carbono/ Peso del cigarrillo*:

$$\text{Monóxido de Carbono} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Peso}$$

Resumen del modelo		
Modelo	R	R cuadrado

	Coef. Correlación	Coef. Determinación
1	0,464	0,215

Coeficientes			
Modelo		Coeficientes no estandarizados	
		B	Error estándar de B
1	(Constante)	-11,795	9,722
	Alquitrán	25,068	9,980

Recogiendo la información anterior tenemos:

$$\boxed{Monóx.Carbono = -11,80 + 25,07 \cdot Peso}$$

$$R^2 = 0,46$$

El valor  $R^2$  vale 0,46, lo que indica que si conocemos el *Peso* del cigarrillo, podemos mejorar en un 46% nuestros pronósticos sobre el contenido en *Monóxido de carbono*; es decir, si en lugar de utilizar como pronóstico el valor medio del contenido en *Monóxido de carbono*, basamos nuestro pronóstico en el contenido en el *Peso* del cigarrillo, mejora la estimación en un 46%.

Interpretando el término independiente y el coeficiente de regresión para la variable *Peso* del cigarrillo podemos decir:

- Por cada incremento de 1 mg en el *Peso* del cigarrillo cabe esperar que la cantidad de *Monóxido de carbono* producido se incremente en 25,07 mg.

### Nota:

Es posible utilizar variables cualitativas en la regresión, pero siempre como paso previo hemos de codificarlas como cuantitativas (variables *dummy*) y a partir de ese momento pensar en ellas como si fueran métricas. Por ejemplo, la variable categórica, *Tipo de tabaco*, ha de recodificarse como variable dicotómica; es decir, una de las categorías se toma como referencia, por ejemplo *Negro*, y para cada paquete estudiado se graba un “1” si el paquete era de tabaco *Negro* y “0” si era *Rubio*. El coeficiente de regresión que sale indica la relación *No Negro/Negro*.

La **media** de la variable *dummy* coincide con la proporción (p) de paquetes de la clase elegida.

Si la variable categórica tuviera varias categorías en vez de dos como en este caso se crearían tantas variables ficticias (*dummy*) como número de categorías menos uno. Por tanto, si tiene  $k$  categorías se definirían  $k-1$  variables ficticias.

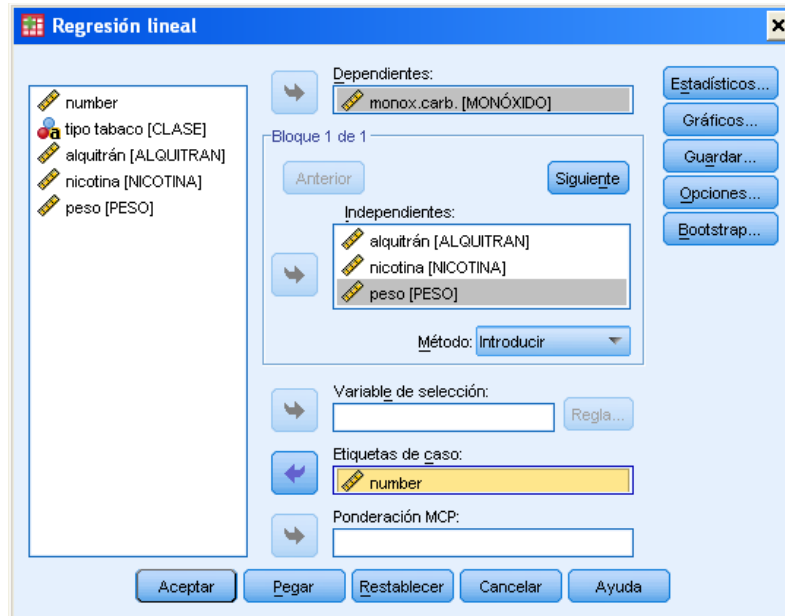
Sintetizando, cuando tengamos variables nominales u ordinales que queramos incorporar a un modelo de regresión, se convierte la variable categórica en tantas variables cuantitativas como número de categorías menos uno y se introducen en el análisis de regresión como si fueran cuantitativas.

## 2) SEGUNDA TAREA: REGRESIÓN MÚLTIPLE

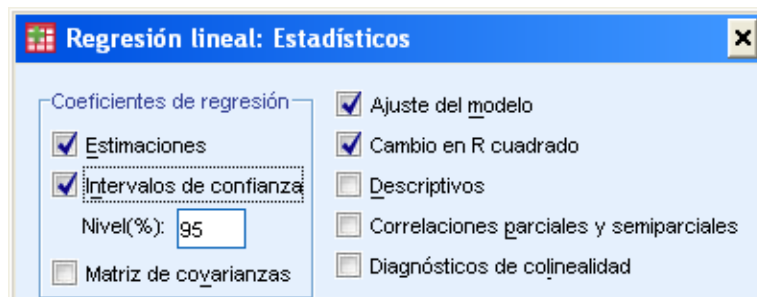
La segunda tarea consiste en ajustar el modelo:

$$\text{Monóx.Carbono} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Alquitrán} + \hat{\beta}_2 \cdot \text{Nicotina} + \hat{\beta}_3 \cdot \text{Peso}$$

Seleccionamos como variables independientes las tres citadas anteriormente:



En el botón Estadísticos..., seleccionamos la siguiente información:



En la Tabla siguiente observamos cómo el modelo es bueno, pues tenemos un coeficiente de correlación R (en los libros de estadística se llama  $r$  y en las salidas de ordenador  $R$ ) múltiple muy alto y coeficiente de determinación (R cuadrado) muy próximo a 1. En este caso es más conveniente interpretar el Coeficiente de Determinación Ajustado (R cuadrado corregida) ya que hay varias variables en el modelo.

#### Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida
1	0,931	0,868	0,849

Variables predictoras: (Constante), peso, alquitrán, nicotina  
Variable dependiente: monox.carb.

En la tabla siguiente aparecen los coeficientes del modelo de regresión múltiple. Aparecen los coeficientes de regresión brutos (no estandarizados) y los coeficientes estandarizados. Los coeficientes de regresión calculados con los valores de las variables sin transformar (es lo que se hace más frecuentemente) se interpretan como el efecto de la variable  $X_i$  en la variable dependiente  $Y$ , cuando el resto de las variables se mantienen constantes.

#### Coeficientes

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	Sig.	Intervalo Confianza B al 95%	
		B	Error típ.	Beta		Limite Inferior	Limite Superior
	(Constante)	1,838	4,388		0,680	-7,287	10,964
	alquitrán	0,239	0,196	0,296	0,237	-,169	0,647
	nicotina	8,612	3,332	0,644	0,017	1,682	15,542
	peso	0,334	4,954	0,006	0,947	-9,968	10,637

Variable dependiente: monox.carb.

#### Nota:

Los coeficientes de regresión parcial estandarizados son los coeficientes que definen la ecuación de regresión cuando ésta se obtiene tras estandarizar (restar su media y dividir por su desviación típica) las variables originales:

$$\beta_{1st.} = \beta_1 (s_x / s_y)$$

Cada coeficiente estandarizado se interpreta como el cambio, en unidades de desviación típica, en la variable dependiente por cada cambio en una desviación típica en la variable independiente  $X_i$ , manteniendo el resto de las variables independientes constantes. La interpretación es menos intuitiva pero es más adecuada cuando las variables vienen expresadas en unidades no comparables.

El modelo resultante es:

$$\text{Monóx. Carbono} = 1,838 + 0,239 \cdot \text{Alquitrán} + 8,612 \cdot \text{Nicotina} + 0,334 \cdot \text{Peso}$$

La interpretación es la siguiente:

- Las variables *Alquitrán*, *Nicotina* y *Peso* del cigarrillo explican el 84.9% de la información.

- La variable *Nicotina* es la que presenta un coeficiente de regresión parcial más alto, 8.612, luego a la vista de estos datos parece ser que es la que tiene mayor relación con el contenido en *Monóxido de carbono* y por tanto, la que tiene mayor capacidad informativa. (El coeficiente de regresión parcial es similar al coeficiente de regresión pero informa sobre lo que aporta una variable cuando las otras se mantienen constantes).  
La relación con esta variable es estadísticamente significativa ya que el correspondiente p-valor es  $0.017 < 0.05$ . Las demás (*Alquitrán* y *Peso*) no aportan información estadísticamente significativa (una vez que la variable contenido en *Nicotina* está en el modelo).
- La relación entre contenido en *Nicotina* y contenido en *Monóxido de carbono* es directa ya que el signo del coeficiente de regresión es positivo: es decir si aumenta la cantidad de *Nicotina* aumenta la cantidad de *Monóxido de Carbono*.
- Si la cantidad de *Alquitrán* y el *Peso* de los cigarrillos fuera constante para los diferentes valores de *Nicotina* (en otras palabras, si el contenido en *Alquitrán* y el *Peso* fueran independientes de la *Nicotina*), podríamos afirmar que por cada incremento unitario en la *Nicotina*, se produce un incremento de 8.61 unidades en el contenido en *Monóxido de Carbono*.

Si las tres variables están relacionadas eso puede no ser cierto. Es necesario hacer un diagnóstico del grado de relación entre ellas, **COLINEALIDAD**, para poder saber si estos resultados son realmente fiables.

### 3) TERCERA TAREA: ANÁLISIS DE LA COLINEALIDAD

Cuando la correlación entre las variables independientes es alta se puede producir el fenómeno de la **colinealidad** el cual afecta a la estabilidad de los coeficientes de regresión haciendo que éstos no sean interpretables.

En presencia de colinealidad los estimadores de regresión son inestables pudiendo aparecer incluso con signo contrario al que deberían de tener, debido a la gran variabilidad que en este caso presentan los estimadores.

#### 3.1) ANÁLISIS DE LOS SÍNTOMAS DE COLINEALIDAD

Síntomas más característicos:

1. Altas correlaciones entre al menos un par de variables.
2. Aparece como no significativa alguna variable que el investigador sabe que es importante.
3. Aparece alguna variable con signo contrario al esperado en el modelo.
4. Los errores estándar de los estimadores son anormalmente grandes, disminuyendo drásticamente al eliminar una o varias variables regresoras.

Teniendo en cuenta que el problema aparece por la posible relación entre las variables explicativas comenzaremos analizando las pistas que nos da la matriz de correlaciones.

#### 1. Análisis de la matriz de correlaciones entre las variables regresoras (explicativas):

		Correlaciones			
		Monóxido carbono	Alquitrán	Nicotina	Peso
Coeficiente Correlación Pearson	Monóxido carbono	1,000	0,907	0,926	0,464
	Alquitrán	0,907	1,000	0,945	0,461
	Nicotina	0,926	0,945	1,000	0,499
	Peso	0,464	0,461	0,499	1,000
Sig. (unilateral)	Monóxido carbono	.	0,000	0,000	0,010
	Alquitrán	0,000	.	0,000	0,010
	Nicotina	0,000	0,000	.	0,006
	Peso	0,010	0,010	0,006	.

La matriz de correlaciones pone claramente de manifiesto que hay una alta correlación directa ( $r = +0.945$ ) y estadísticamente significativa ( $p\text{-valor} = 0.000$ ) entre el contenido en *Alquitrán* y el contenido en *Nicotina*. También existen otras relaciones significativas (ver tabla más arriba) pero no son tan fuertes.

Por tanto el *Alquitrán* puede estar distorsionando el coeficiente de regresión (parcial) que aparece en el modelo acompañando a la *Nicotina*.



## 2. Análisis de la significación de los coeficientes de regresión (parcial):

Vimos al estudiar los modelos simples que la variable *Alquitrán* tenía relación estadísticamente significativa ( $p\text{-valor} < 0.05$ ) con el contenido en *Monóxido de carbono*, sin embargo en el modelo de regresión múltiple aparece como no significativa.

## 3. Análisis del signo de los coeficientes de regresión (parcial):

En este caso el signo encontrado es el esperado, luego aquí no encontramos pistas sobre la posible colinealidad.

## 4. Análisis de los errores estándar de los coeficientes de regresión (parcial):

Retomando aquí la tabla ya analizada más arriba, que volvemos a escribir debajo, observamos:

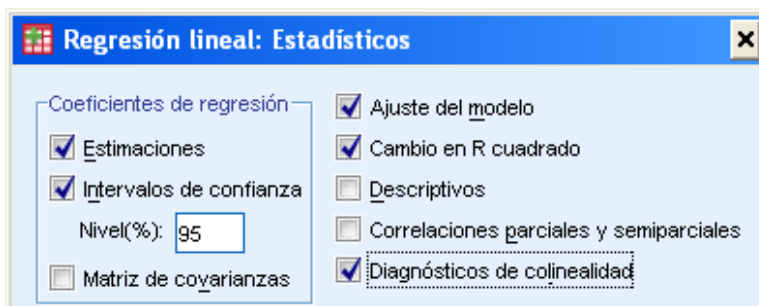
	Coeficientes no estandarizados	
	B	Error típ.
(Constante)	1,838	4,388
alquitrán	0,239	0,196
nicotina	8,612	3,332
peso	0,334	4,954

El error estándar del coeficiente de la *Nicotina* es muy grande y muchísimo más aún es el error estándar del coeficiente de regresión de la variable *Peso*. Un síntoma más de que el fenómeno de colinealidad puede estar presente.

Debemos, pues, pasar a realizar un diagnóstico exhaustivo. (*Es conveniente que esta parte sea realizada por un estadístico experto en el tema, no todos lo son ya que dentro de los estadísticos hay especialidades lo mismo que entre los médicos*).

## 3.2) DIAGNÓSTICO DE COLINEALIDAD

Para realizar un estudio de colinealidad debemos marcarlo en el cuadro de entrada:



Tener en cuenta los siguientes criterios:

Diagnostico Colinealidad:	
	Analizar
¿Hay colinealidad?	Condition Number
¿Hasta donde alcanza?	Condition Index
¿Qué coeficientes están afectados?	VIF
¿Qué regresoras están involucradas?	Contribuciones de cada componente al factor de inflación de varianza

### ¿HAY COLINEALIDAD?

Analizamos el "**condition number**" que es un índice para evaluar la inestabilidad global de los coeficientes de regresión. Un *condition number* en torno a 30 podría resultar peligroso.

(Se calcula como la raíz cuadrada de la proporción entre el valor propio (*Eigenvalue*) más grande y el más pequeño de la matriz de covarianzas entre las variables regresoras. Los valores propios (*Eigenvalue*) informan sobre cuántas dimensiones o factores diferentes subyacen en el conjunto de variables independientes analizadas. La presencia de varios valores propios cercanos a cero indica que las variables independientes están muy relacionadas entre sí (*colinealidad*) ya que si las variables están relacionadas, las dimensiones latentes (*Eigenvectors*) verdaderas son menos que las variables existentes y por tanto algún valor propio tiene que ser cero o casi cero).

En nuestro caso el *condition number* = 34,509, luego estamos en presencia de colinealidad. Ahora hay que localizar las variables implicadas y las afectadas. Ver tabla en el apartado siguiente.

### ¿HASTA DONDE ALCANZA?

Para saber la magnitud del problema de la colinealidad se calcula el "**condition index**" (*índice de condición*) que nos indica cuántas dimensiones latentes están asociadas a valores propios muy pequeños y por tanto si existe una o más relaciones que afecten a las estimaciones de los coeficientes de regresión.

En nuestro estudio hay dos *condition index* muy altos, el correspondiente a la dimensión 3 y el correspondiente a la dimensión 4. El mayor de los *condition index* es el *condition number*.

Diagnósticos de colinealidad			
Modelo	Dimensión	Autovalor	Índice de condición
1	1	3,834	1,000
	2	0,154	4,993
	3	0,009	21,190
	4	0,003	34,509

Variable dependiente: monox.carb.

Cualquier *condition index* superior a 15 debe ser tenido en cuenta. En nuestro caso hay dos, tal como ya habíamos señalado.

### ¿QUÉ COEFICIENTES ESTÁN AFECTADOS?

Se calculan los *factores de inflación* de varianza que indican cuánto se han desestabilizado los coeficientes de regresión como resultado de la presencia de colinealidad.

La salida del SPSS proporciona esta información al lado del cuadro que recoge los estimadores y sus errores estándar, si en el cuadro de diálogo correspondiente se ha marcado *Diagnósticos de colinealidad*.

La *Tolerancia* es igual al inverso del *FIV* luego ambas columnas (las dos de la derecha) proporcionan exactamente la misma información.

**Coeficientes<sup>a</sup>**

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Intervalo de confianza para B al 95%		Estadísticos de colinealidad	
	B	Error típ.	Beta			Límite inferior	Límite superior	Tolerancia	FIV
1 (Constante)	1,838	4,388		,419	,680	-7,287	10,964		
alquitran	,239	,196	,296	1,217	,237	-,169	,647	,107	9,382
nicotina	8,612	3,332	,644	2,584	,017	1,682	15,542	,102	9,841
peso	,334	4,954	,006	,067	,947	-9,968	10,637	,750	1,334

a. Variable dependiente: monox.carb.

Los factores de inflación se calculan como:

$$\frac{1}{1 - R_j^2}$$

Eso implica que cuando las variables explicativas están muy relacionadas, como el coeficiente de correlación múltiple va a estar próximo a uno, tendríamos que dividir por una cantidad muy pequeña y el cociente sería muy grande. Eso es lo que significa que se han inflado.

Por ejemplo, para el *Alquitrán*:

$$VIF_{Alquitran} = \frac{1}{1 - R_{Alquitran}^2} = \frac{1}{1 - 0,893} = 9,35 \approx 9,382$$

De lo anterior se deduce que tanto la *Nicotina* como el *Alquitrán* tienen coeficientes que están afectados por el problema de la colinealidad.

Factores de inflación de varianza mayores de 10 son ya muy peligrosos y el problema debe atajarse con valores bastante más bajos.

## ¿QUÉ REGRESORAS ESTÁN INVOLUCRADAS?

Las variables más correlacionadas con las dimensiones más cortas (las asociadas a los *autovalores* más pequeños son las responsables de la inflación detectada en los coeficientes de regresión. En la salida del SPSS leemos “*Proporciones de la varianza*” indicando cuánto aportan a la variabilidad (varianza) de los coeficientes.

Diagnósticos de colinealidad

Modelo	Dimensión	Autovalor	Índice de condición	Proporciones de la varianza			
				Constante	Alquitrán	Nicotina	Peso
1	1	3,834	1,000	0,00	0,00	0,00	0,00
	2	0,154	4,993	0,01	0,05	0,01	0,01
	3	0,009	21,190	0,02	0,95	0,94	0,00
	4	0,003	34,509	0,97	0,00	0,04	0,99

Variable dependiente: monox.carb.

De esta información se deduce que el *Alquitrán* y la *Nicotina* son las responsables de la colinealidad. De la información anterior sabemos, además, que son las afectadas.

La síntesis de todo lo anterior aparece a continuación.

**DATOS**

CLASE	ALQUITRAN	NICOTINA	PESO	MONÓXIDO
RUBIO	4,10	0,860	0,9853	13,60
NEGRO	16,00	1,060	1,0938	16,60
NEGRO	29,80	2,030	1,1650	23,50
RUBIO	8,00	0,670	0,9280	10,20
RUBIO	4,10	0,400	0,9462	5,40
NEGRO	15,00	1,040	0,8885	15,00
RUBIO	8,80	0,750	1,0267	9,00
RUBIO	12,40	0,950	0,9225	12,30
NEGRO	16,60	1,120	0,9372	16,30
RUBIO	14,90	1,020	0,8858	15,40
RUBIO	13,70	1,010	0,9643	13,00
RUBIO	15,10	0,900	0,9316	14,40
RUBIO	7,80	0,570	0,9705	10,00
RUBIO	11,40	0,780	1,1240	10,20
RUBIO	9,00	0,740	0,8517	9,50
RUBIO	1,00	0,130	0,7851	1,50
NEGRO	17,00	1,260	0,9186	18,50
RUBIO	12,80	1,080	1,0395	12,60
NEGRO	15,80	0,960	0,9573	17,50
RUBIO	4,50	0,420	0,9106	4,90
NEGRO	14,50	1,010	1,0070	15,90
RUBIO	7,30	0,610	0,9806	8,50
NEGRO	8,60	0,690	0,9693	10,60
NEGRO	15,20	1,020	0,9496	13,90
RUBIO	12,00	0,820	1,1184	14,90

**SPSS**

Analyze > Regression > Linear...

Dependent: monox.carb. [MONÓXIDO]

Independent(s): alquitrán [ALQUITRAN], nicotina [NICOTINA], peso [PESO]

Regression Coefficients: ☒ Estimates, ☒ Confidence intervals, ☐ Covariance matrix, ☐ Part and partial correlations, ☒ Collinearity diagnostics

**¿Hay colinealidad?**

Collinearity Diagnostics

Mod el	Dime nsion	Eigenvalue	Condition Index	Variance Proportions
1	1	3,834	1,000	(Constant) ,00, alquitrán ,00, nicotina ,00, peso ,00
	2	,154	4,993	,01, ,05, ,01, ,01
	3	,009	21,190	,02, ,95, ,94, ,00
	4	,003	34,509	,97, ,00, ,04, ,99

a. Dependent Variable: monox.carb. **Cond Number>30**

**Regresoras involucradas**

**¿Afectadas? VIF>10**

**Coefficients\***

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error				Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	1,838	4,388		,419	,680	-7,287	10,964		
	alquitrán	,239	,196	,296	1,217	,237	-,169	,647	,107	9,382
	nicotina	8,612	3,332	,644	2,584	,017	1,682	15,542	,102	9,841
	peso	,334	4,954	,006	,067	,947	-9,968	10,637	,750	1,334

a. Dependent Variable: monox.carb.

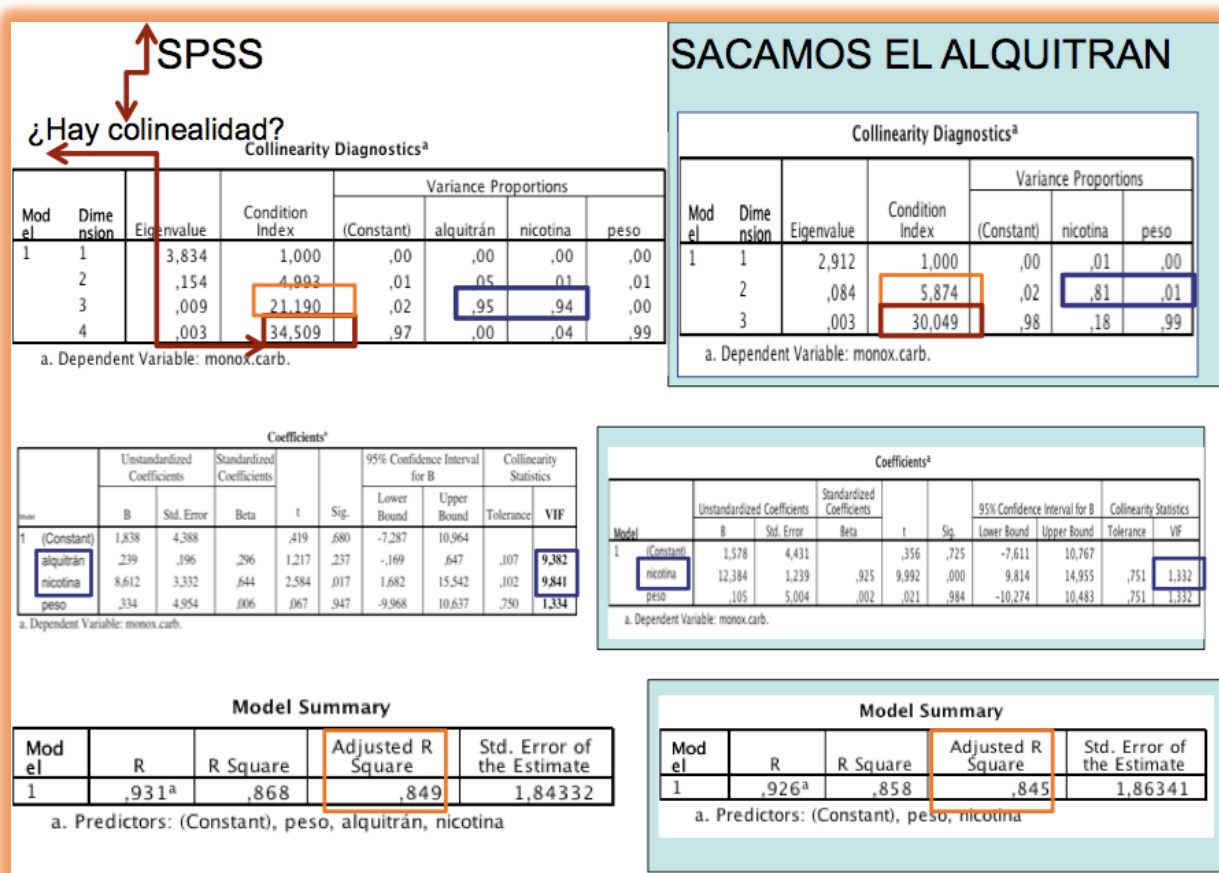
Como conclusión señalamos que UNA DE LAS DOS VARIABLES DEBE SALIR DEL MODELO.

Como las dos variables, *Nicotina* y *Alquitrán* aportan la misma proporción de varianza, una de las dos es superflua. Sacamos el *Alquitrán* aunque los resultados serían similares si sacásemos la *Nicotina*.

El modelo a ajustar, será pues:

$$\text{Monóx.Carbono} = \hat{\beta}_0 + \hat{\beta}_1 \cdot \text{Nicotina} + \hat{\beta}_2 \cdot \text{Peso}$$

Repetiendo el estudio obtenemos lo siguiente:



Modelo

$$\text{Monóx.Carbono} = 1,578 + 12,384 \cdot \text{Nicotina} + 0,105 \cdot \text{Peso}$$

El modelo está libre de colinealidad y explica el 85.8% de la información.