

# Regresión Lineal Múltiple

Predicción, diagnosis del modelo, selección de variables

Curso 2011-2012

## Predicción en el modelo de regresión lineal múltiple

### Modelo de regresión

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \varepsilon$$

Al igual que en el modelo de regresión lineal simple, debemos distinguir dos problemas diferentes:

- ▶ Estimar el valor esperado (valor medio) de  $Y$  dado que  $X_1 = x_1, \dots, X_p = x_p$ .

$$\mathbb{E}(Y/X_1 = x_1, \dots, X_p = x_p)$$

**Ejemplo:** ¿Cuál es el número medio estimado de parásitos en las localizaciones con un 85 % de humedad y temperatura de 20°C?

- ▶ Predecir el valor de  $Y$  en un individuo para el que se conoce que  $X_1 = x_1, \dots, X_p = x_p$ .

$$Y/X_1 = x_1, \dots, X_p = x_p$$

**Ejemplo:** En una zona de clima tropical del sur de Estados Unidos, la temperatura es de 20°C y su humedad del 85 %, ¿cuál es el número predicho de parásitos en dicha localización?

## Predicción en el modelo de regresión lineal múltiple

### Modelo de regresión ajustado

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p$$

Al igual que en el modelo de regresión lineal simple, debemos distinguir dos problemas diferentes:

- ▶ El estimador de la media de  $Y$  dado que  $X_1 = x_1, \dots, X_p = x_p$  será:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p$$

**Ejemplo:** ¿Cuál es el número medio estimado de parásitos en las localizaciones con un 85 % de humedad y temperatura de 20°C? El número medio estimado será

$$25.7115 + 1.5818 \cdot 20 + 1.5424 \cdot 85 = 188.45 \text{ parásitos.}$$

- ▶ Predecimos el valor de  $Y$  en un individuo para el que se conoce que  $X_1 = x_1, \dots, X_p = x_p$  mediante:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_p x_p$$

**Ejemplo:** En una zona de clima tropical del sur de Estados Unidos, la temperatura es de 20°C y su humedad del 85 %, ¿cuál es el número predicho de parásitos en dicha localización? Predecimos que en dicha localización habrá

$$25.7115 + 1.5818 \cdot 20 + 1.5424 \cdot 85 = 188.45 \text{ parásitos.}$$

## Predicción en el modelo de regresión lineal múltiple

### Modelo de regresión ajustado

$$Y = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_p X_p$$

Al igual que en el modelo de regresión lineal simple, debemos distinguir dos problemas diferentes:

- ▶ El intervalo de confianza para la media de  $Y$  dado que  $X_1 = x_1, \dots, X_p = x_p$  de nivel  $\alpha$  será:

$$\left( \hat{Y} - \hat{\sigma} \sqrt{h(x_1, \dots, x_p)} t_{n-(p+1)}^{\alpha/2}, \hat{Y} + \hat{\sigma} \sqrt{h(x_1, \dots, x_p)} t_{n-(p+1)}^{\alpha/2} \right)$$

- ▶ El intervalo de confianza para el valor de  $Y$  en un individuo para el que se conoce que  $X_1 = x_1, \dots, X_p = x_p$  será:

$$\left( \hat{Y} - \hat{\sigma} \sqrt{1 + h(x_1, \dots, x_p)} t_{n-(p+1)}^{\alpha/2}, \hat{Y} + \hat{\sigma} \sqrt{1 + h(x_1, \dots, x_p)} t_{n-(p+1)}^{\alpha/2} \right)$$

# Predicción en el modelo de regresión lineal múltiple

Regresión lineal

Dependientes:  
Recuento

Independientes:  
Temperatura  
Humedad

Método: Introducir

Variable de selección:

Etiquetas de caso:

Ponderación MCP:

Aceptar Pegar Restablecer Cancelar Ayuda

Regresión lineal: Guardar

Valores pronosticados:  
 No tipificados  
 Tipificados  
 Correídos  
 E.T. del genérico promedio

Residuos:  
 No tipificados  
 Tipificados  
 Método de Student  
 Eliminados  
 Eliminados estandarizados

Distancias:  
 Máximas  
 De Coef.  
 Valores de influencia

Intervalos de pronóstico:  
 Media  Intervalos  
Intervalo de confianza: 95 %

Estadísticos de los coeficientes:  
 Crear un nuevo conjunto de datos  
Nombre del conjunto de datos:  
  
 Escribir un nuevo archivo de datos  
Archivo:

Exportar información del modelo a un archivo XML:  
 Examinar  
 Incluir la matriz de covarianzas

Continuar Cancelar Ayuda

"peraso.sav [Conjunto de datos15] - PASW Statistics Editor de datos

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda

1. LMCI\_1 152.82330157097864

	Temperatura	Humedad	Recuento	PRE_1	RES_1	LMCI_1	LMCI_1	LIC_1	UICI_1
1	15	70	156	157.41015	-1.41015	152.82330	161.39059	144.88238	169.93791
2	16	65	157	151.27973	5.72027	145.64240	156.31705	138.33039	164.22906
3	24	71	177	173.18896	3.81104	167.73621	178.64171	160.31891	186.05901
4	13	64	145	144.99183	.00817	138.48698	151.49567	131.64196	158.34169
5	21	84	197	188.49533	8.50467	184.23801	192.75264	176.08442	200.90623
6	16	86	184	183.67113	.32887	177.39637	189.94589	170.43186	196.91041
7	22	72	172	171.56777	4.3223	167.29614	175.83940	159.15195	183.98359
8	18	84	187	183.74967	3.25013	178.93219	188.56755	171.13576	196.26398
9	20	71	157	166.86169	-9.86169	162.91867	170.80470	154.55506	179.16632
10	16	75	169	166.70421	2.29579	162.88175	170.52662	154.43568	178.97273
11	28	84	200	199.56805	4.3195	193.21643	205.91968	186.29217	212.84293
12	27	79	193	190.27399	2.72601	184.53137	196.01662	177.27848	203.26951
13	13	80	167	169.67099	-2.67099	163.96365	175.77833	156.51024	182.83174
14	22	76	170	177.73756	-7.73756	174.23843	181.23676	165.56588	189.90504
15	23	88	192	197.82875	-5.82875	192.27211	203.38540	184.91434	210.74317
17									

Visa de datos Visa de variables

PSPSS Statistics Processor está listo

Intervalos de confianza para la media

# Predicción en el modelo de regresión lineal múltiple

Regresión lineal

Dependientes:  
Recuento

Independientes:  
Temperatura  
Humedad

Método: Introducir

Variable de selección:

Etiquetas de caso:

Ponderación MCP:

Aceptar Pegar Restablecer Cancelar Ayuda

Regresión lineal: Guardar

Valores pronosticados:  
 No especificados  
 Tipificados  
 Correídos  
 E.T. del genérico promedio

Residuos:  
 No tipificados  
 Tipificados  
 Método de Student  
 Eliminados  
 Eliminados estandarizados

Distancias:  
 Matriz de covarianzas  
 De Cook  
 Valores de influencia

Intervalos de pronóstico:  
 Media y intervalos  
Intervalo de confianza: 95 %

Estadísticos de los coeficientes:  
 Crear estadístico de los coeficientes  
Crear un nuevo conjunto de datos:  
Nombre de conjunto de datos:  
Escribir un nuevo archivo de datos:  
Archivo:

Exportar información del modelo a un archivo XML:  
Incluir la matriz de covarianzas

Continuar Cancelar Ayuda

"perasno.sav (conjunto\_de\_datos15) - PASW Statistics Editor de datos

1	LIQI_1	Temperatura	Humedad	Recuento	PRE_1	RES_1	LMCI_1	LMCI_1	LIQI_1	LIQI_1
1	15	70	156	157.41015	-1.41015	152.82330	161.99659	144.88238	169.93791	
2	16	65	157	151.27973	5.72027	145.64240	156.91705	138.33039	164.22906	
3	24	71	177	173.18896	3.81104	167.73621	178.64171	160.31891	186.05901	
4	13	64	145	144.99183	.00817	138.48698	151.49667	131.64196	158.34169	
5	21	84	197	188.49533	8.50467	184.23801	192.75264	176.08442	200.90623	
6	16	86	184	183.67113	.32887	177.39637	189.94589	170.43186	196.91041	
7	22	72	172	171.56777	.43223	167.29614	175.83940	159.15195	183.80359	
8	18	84	187	183.74967	3.25013	178.93219	188.56755	171.13576	196.26398	
9	20	71	157	166.86169	-9.86169	162.91867	170.80470	154.55506	179.16832	
10	16	75	169	166.70421	2.29579	162.88175	170.52662	154.43568	178.97273	
11	28	84	200	199.56805	4.3195	193.21643	205.91968	186.29217	212.84393	
12	27	79	193	190.27399	2.72601	184.53137	196.01662	177.27848	203.26951	
13	13	80	167	169.67099	-2.67099	163.56365	175.77833	156.51024	182.83174	
14	22	76	170	177.73756	-7.73756	174.23843	181.23670	165.56588	189.90504	
15	23	88	192	197.82875	-5.82875	192.27211	203.38540	184.91434	210.74317	
17										

Visa de datos Visa de variables

PASW Statistics Processor está listo

Intervalos de confianza para los valores de Y

## Diagnosis y validación del modelo de regresión lineal múltiple

- ▶ Valorar la calidad del ajuste a través, por ejemplo, del coeficiente de determinación no es lo mismo que valorar el cumplimiento de las hipótesis básicas del modelo.
- ▶ Las labores de inferencia que realizamos bajo el modelo lineal, tienen sentido suponiendo que los datos proceden de tal modelo, tal y como se ha formulado, con todas sus hipótesis básicas
- ▶ Si las hipótesis no se corresponden con la realidad, se pueden estar cometiendo graves errores en las conclusiones obtenidas de la inferencia basada en el modelo

## Diagnosis y validación del modelo de regresión lineal múltiple

Para llevar a cabo una buena interpretación de un modelo de regresión debemos acompañar siempre nuestro estudio de la diagnosis y validación del modelo. Dicha diagnosis consiste en analizar si se verifican las **hipótesis básicas del modelo**:

- ▶ **Linealidad**: los parámetros y su interpretación carecen de sentido si en realidad los datos no proceden de un modelo lineal, situación en la que además las predicciones pueden ser completamente equivocadas.
- ▶ **Normalidad de los errores**: El modelo de regresión lineal asume que la distribución de los errores es Normal.
- ▶ **Homocedasticidad**: La varianza del error es constante.
- ▶ **Independencia de los errores**: Las variables aleatorias que representan los errores son mutuamente independientes.
- ▶ Las variables explicativas  $X_1, X_2, \dots, X_p$ , son linealmente independientes.



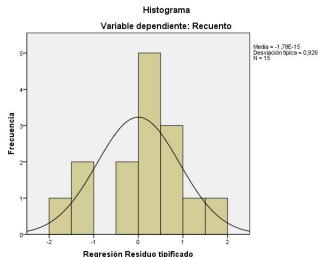
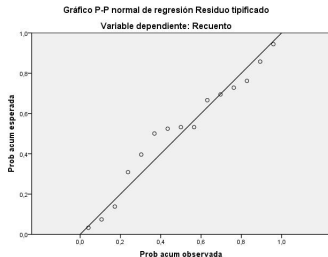
# Diagnos y validación del modelo de regresión lineal múltiple

## Hipótesis de normalidad

El modelo de regresión lineal múltiple asume que la distribución de los errores es Normal. Para estudiar si se verifica esta hipótesis se pueden realizar gráficos de normalidad o aplicar contrastes de normalidad.

- ▶ Gráficos de normalidad: histograma, QQ-plot,...
- ▶ Contrastes de normalidad: contraste de Kolmogorov-Smirnov, contraste Chi-cuadrado,...

La falta de normalidad influye en el modelo en que los estimadores no son eficientes y los intervalos de confianza de los parámetros del modelo y los contrastes de significación no son exactos.



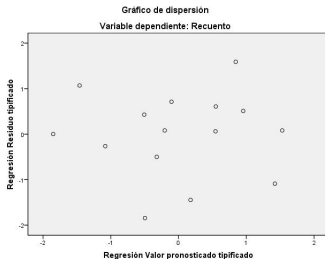
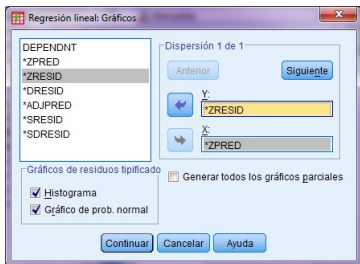
# Diagnos y validación del modelo de regresión lineal múltiple

## Homocedasticidad

El modelo de regresión lineal múltiple asume que la varianza de los errores es constante (homocedasticidad). Para estudiar si se verifica esta hipótesis se pueden realizar gráficos de residuos.

- ▶ Gráfico de residuos frente a las predicciones.

La falta de homocedasticidad hace que no puedan aplicarse los contrastes de significación.



# Diagnóstico y validación del modelo de regresión lineal múltiple

## Independencia de los errores

El modelo de regresión lineal múltiple asume que los errores son independientes.

- ▶ La sospecha sobre un incumplimiento de la independencia suele venir avalada por la presencia de varias observaciones sobre un mismo individuo o un mismo colectivo, o por observaciones consecutivas en el tiempo.
- ▶ En este caso existen métodos de contraste específicos para detectar una posible correlación secuencial entre los errores.

La falta de independencia afecta a la calidad de los estimadores, a los contrastes individuales sobre los parámetros ( $t$  de Student), a la calidad de las predicciones,...

# Diagnóstico y validación del modelo de regresión lineal múltiple

## Multicolinealidad

El modelo de regresión lineal múltiple asume que las variables explicativas  $X_1, X_2, \dots, X_p$ , son linealmente independientes.

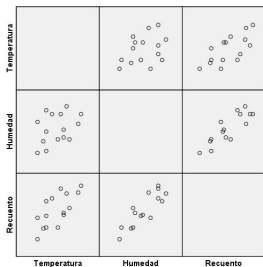
- ▶ Recuerda, que  $\hat{\beta} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y}$ . Si existe relación lineal exacta entre variables explicativas, entonces  $\mathbf{X}^t \mathbf{X}$  no tiene inversa y no se puede estimar de manera única el vector de parámetros  $\beta$ .
- ▶ En la práctica, aunque no se suele dar una relación lineal exacta entre variables explicativas, sí que suele ocurrir que existe una cierta relación entre ellas (multicolinealidad). En ese caso se puede ajustar y estimar el modelo de regresión lineal, pero con mucha variabilidad.
- ▶ El problema de multicolinealidad se resuelve eliminando del modelo las variables explicativas dependientes.

# Diagnosis y validación del modelo de regresión lineal múltiple

## Multicolinealidad

Se puede detectar multicolinealidad a partir de:

- ▶ Gráfico de dispersión
- ▶ Matriz de correlaciones



Correlaciones

		Temperatura	Humedad	Recuento
Temperatura	Correlación de Pearson	1	,356	,697**
	Sig. (bilateral)		,193	,004
	N	15	15	15
Humedad	Correlación de Pearson	,356	1	,860**
	Sig. (bilateral)	,193		,000
	N	15	15	15
Recuento	Correlación de Pearson	,697**	,860**	1
	Sig. (bilateral)	,004	,000	
	N	15	15	15

\*\* La correlación es significativa al nivel 0,01 (bilateral).

# Diagnos y validación del modelo de regresión lineal múltiple

## Multicolinealidad

Se puede detectar multicolinealidad a partir de los resultados de los contrastes:

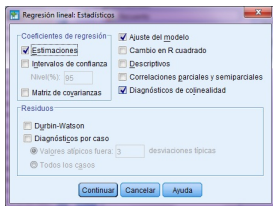
Contraste global ( $F$ )	Contrastes individuales ( $t$ )	Conclusión
Modelo explicativo	Todas las $X_i$ explicativas	Nos quedamos con todas las $X_i$
Modelo explicativo	Algunas $X_i$ explicativas	Nos quedamos con las $X_i$ explicativas
Modelo explicativo	Ninguna $X_i$ explicativa	Puede indicar multicolinealidad
Modelo no explicativo	Alguna o todas las $X_i$ explicativas	Puede indicar multicolinealidad
Modelo no explicativo	Ninguna $X_i$ explicativa	Modelo no adecuado

# Diagnos y validación del modelo de regresión lineal múltiple

## Multicolinealidad

Se puede detectar multicolinealidad a partir de diagnósticos específicos:

- ▶ Factores de inflación de la varianza (FIV): Como regla general nos deberían preocupar valores de FIV mayores de 10.
- ▶ Índice de condición: Si no hay multicolinealidad, estos índices no deben superar el valor 15. Índices mayores que 15 indican posible existencia de multicolinealidad. Si los índices superan en valor 30, hay un problema severo de multicolinealidad.



Modelo		Coefficients no estandarizados		Coefficientes tipificados	t	Sig.	Estadísticos de colinealidad	
		B	Error típ.	Beta			Tolerancia	FIV
1	(Constante)	25,712	14,372		1,789	,099		
	Temperatura	1,582	,320	,447	4,939	,000	,873	1,145
	Humedad	1,542	,200	,700	7,731	,000	,873	1,145

a. Variable dependiente: Recuento

Modelo	Dimensión	Autovalores	Índice de condición	Proporciones de la varianza		
				(Constante)	Temperatura	Humedad
1	1	2,964	1,000	,00	,01	,00
	2	,032	9,648	,06	,96	,03
	3	,005	25,638	,94	,03	,97

a. Variable dependiente: Recuento

## Selección de variables en el modelo de regresión lineal múltiple

- ▶ Generalmente en la práctica disponemos de un conjunto grande de posibles variables explicativas.
- ▶ Se trata de determinar qué variables deben entrar en el modelo y cuáles no deben entrar en el modelo de regresión.
- ▶ Considerar todos las posibles combinaciones de variables explicativas es inviable en muchas situaciones pero existen procedimientos para seleccionar las variables explicativas que deben entrar en el modelo.

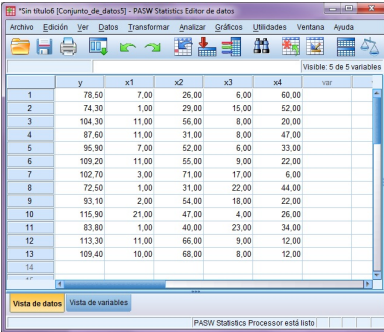


# Selección de variables en el modelo de regresión lineal múltiple

## Ejemplo: Calor emitido por el fraguado de cemento

Un estudio analiza la relación entre la composición de un cemento tipo Portland y el calor desprendido durante la fase de fraguado. La muestra está formada por 13 cementos.<sup>1</sup>

- ▶  $Y$  es la cantidad de calor desprendido (cals/gr).
- ▶ Las variables  $X_1$ ,  $X_2$ ,  $X_3$  y  $X_4$  representan el contenido (%) de cuatro ingredientes activos.

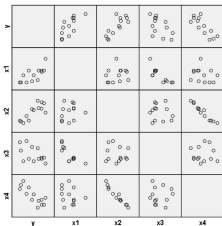


	y	x1	x2	x3	x4	var
1	78,50	7,00	26,00	6,00	60,00	
2	74,30	1,00	29,00	15,00	52,00	
3	104,30	11,00	56,00	8,00	20,00	
4	87,60	11,00	31,00	8,00	47,00	
5	95,90	7,00	52,00	6,00	33,00	
6	109,20	11,00	55,00	9,00	22,00	
7	102,70	3,00	71,00	17,00	6,00	
8	72,50	1,00	31,00	22,00	44,00	
9	93,10	2,00	54,00	18,00	22,00	
10	115,90	21,00	47,00	4,00	26,00	
11	83,80	1,00	40,00	23,00	34,00	
12	113,30	11,00	66,00	9,00	12,00	
13	109,40	10,00	68,00	8,00	12,00	
14						

<sup>1</sup>Woods, H., Steinour, H.H., Starke, H.R. (1932) Effect of composition of Portland cement on heat evolved during hardening. Industrial Engineering and Chemistry, 24, 1207-1214.

# Selección de variables en el modelo de regresión lineal múltiple

## Ejemplo: Calor emitido por el fraguado de cemento



**Correlaciones**

		y	x1	x2	x3	x4
y	Correlación de Pearson	1	,731**	,816**	-,535	-,821**
	Sig. (bilateral)		,005	,001	,060	,001
	N	13	13	13	13	13
x1	Correlación de Pearson	,731**	1	,229	-,824**	-,245
	Sig. (bilateral)	,005		,453	,001	,419
	N	13	13	13	13	13
x2	Correlación de Pearson	,816**	,229	1	-,139	-,973**
	Sig. (bilateral)	,001	,453		,650	,000
	N	13	13	13	13	13
x3	Correlación de Pearson	-,535	-,824**	-,139	1	,030
	Sig. (bilateral)	,060	,001	,650		,924
	N	13	13	13	13	13
x4	Correlación de Pearson	-,821**	-,245	-,973**	,030	1
	Sig. (bilateral)	,001	,419	,000	,924	
	N	13	13	13	13	13

\*\* La correlación es significativa al nivel 0,01 (bilateral).

**Variables introducidas/eliminadas<sup>b</sup>**

Modelo	Variables introducidas	Variables eliminadas	Método
1	x4, x3, x1, x2 <sup>a</sup>	.	Introducir

a. Todas las variables solicitadas introducidas.

b. Variable dependiente: y

**Resumen del modelo**

Modelo	R	R cuadrado	R cuadrado corregido	Error tip. de la estimación
1	,991 <sup>a</sup>	,982	,974	2,44601

a. Variables predictoras: (Constante), x4, x3, x1, x2

**ANOVA<sup>b</sup>**

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	2667,899	4	666,975	111,479	,000 <sup>a</sup>
	Residual	47,864	8	5,983		
	Total	2715,763	12			

a. Variables predictoras: (Constante), x4, x3, x1, x2

b. Variable dependiente: y

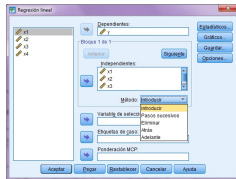
**Coefficientes<sup>b</sup>**

Modelo		Coefficients no estandarizados		Coefficients tipificados	t	Sig.	Estadísticos de colinealidad	
		B	Error tip.	Beta			Tolerancia	FIV
1	(Constante)	62,405	70,071		,891	,399		
	x1	1,551	,745	,607	2,083	,071	,026	38,496
	x2	,510	,724	,528	,705	,501	,004	254,423
	x3	,102	,755	,043	,135	,896	,021	46,868
	x4	-,144	,709	-,160	-,203	,844	,004	282,513

a. Variable dependiente: y

# Selección de variables en el modelo de regresión lineal múltiple

- ▶ **Eliminación hacia atrás (Backward Stepwise Regression)**. Se introducen todas las variables en la ecuación y después se van excluyendo una tras otra. En cada etapa se elimina la variable menos influyente según el contraste individual (de la  $t$  o de la  $F$ ).
- ▶ **Selección hacia adelante (Forward Stepwise Regression)**. Las variables se introducen secuencialmente en el modelo. La primera variable que se introduce es la de mayor correlación (+ o -) con la variable dependiente. Dicha variable se introducirá en la ecuación sólo si cumple el criterio de entrada. A continuación se considerará la variable independiente cuya correlación parcial sea la mayor y que no esté en la ecuación. El procedimiento termina cuando ya no quedan variables que cumplan el criterio de entrada.
- ▶ **Pasos sucesivos (Stepwise Regression)**. Este método es una combinación de los procedimientos anteriores. En cada paso se introduce la variable independiente que no se encuentre ya en la ecuación y que tenga la probabilidad para  $F$  más pequeña. Las variables ya introducidas en la ecuación de regresión pueden ser eliminadas del modelo. El método termina cuando ya no hay más variables candidatas a ser incluidas o eliminadas.



Puedes encontrar los contenidos de este tema desarrollados en el libro "Modelos Estadísticos Aplicados" de Juan M. Vilar Fernández (2006). Publicaciones de la UDC.

# Selección de variables en el modelo de regresión lineal múltiple

Eliminación hacia atrás (opción **Atrás** de SPSS)

**Variables introducidas/eliminadas<sup>b</sup>**

Modelo	Variables introducidas	Variables eliminadas	Método
1	x4, x3, x1, x2 <sup>a</sup>	-	Introducir
2	-	x3	Hacia atrás (criterio: Prob. de F para salir >= ,100).
3	-	x4	Hacia atrás (criterio: Prob. de F para salir >= ,100).

a. Todas las variables solicitadas introducidas.

b. Variable dependiente: y

**Resumen del modelo**

Modelo	R	R cuadrado	R cuadrado corregido	Error típ. de la estimación
1	,991 <sup>a</sup>	,982	,974	2,44601
2	,991 <sup>b</sup>	,982	,976	2,30874
3	,989 <sup>c</sup>	,979	,974	2,40634

a. Variables predictoras: (Constante), x4, x3, x1, x2

b. Variables predictoras: (Constante), x4, x1, x2

c. Variables predictoras: (Constante), x1, x2

**ANOVA<sup>d</sup>**

Modelo		Suma de cuadrados	gl	Medja cuadrática	F	Sig.
1	Regresión	2667,899	4	666,975	111,479	,000 <sup>a</sup>
	Residual	47,864	8	5,983		
	Total	2715,763	12			
2	Regresión	2667,790	3	889,263	166,832	,000 <sup>b</sup>
	Residual	47,973	9	5,330		
	Total	2715,763	12			
3	Regresión	2657,859	2	1328,929	229,504	,000 <sup>c</sup>
	Residual	57,904	10	5,790		
	Total	2715,763	12			

a. Variables predictoras: (Constante), x4, x3, x1, x2

b. Variables predictoras: (Constante), x4, x1, x2

c. Variables predictoras: (Constante), x1, x2

d. Variable dependiente: y

**Coefficientes<sup>a</sup>**

Modelo		Coeficientes no estandarizados		Coeficientes tipificados		t	Sig.	Estadísticos de colinealidad	
		B	Error típ.	Beta				Tolerancia	FIV
1	(Constante)	62,405	70,071			,891	,399		
	x1	1,551	,745	,607		2,083	,071	,026	38,496
	x2	,510	,724	,528		,705	,501	,004	254,423
	x3	,102	,755	,043		,135	,896	,021	46,868
	x4	-,144	,709	-,160		-,203	,844	,004	282,513
2	(Constante)	71,648	14,142			5,066	,001		
	x1	1,452	,117	,568		12,410	,000	,938	1,066
	x2	,416	,186	,430		2,242	,052	,053	18,780
	x4	-,237	,173	-,263		-1,365	,205	,053	18,940
3	(Constante)	52,577	2,286			22,998	,000		
	x1	1,468	,121	,574		12,105	,000	,948	1,055
	x2	,662	,046	,685		14,442	,000	,948	1,055

a. Variable dependiente: y

# Selección de variables en el modelo de regresión lineal múltiple

Selección hacia adelante (opción **Adelante** de SPSS)

**Variables introducidas/eliminadas<sup>a</sup>**

Modelo	Variables introducidas	Variables eliminadas	Método
1	x4	.	Hacia adelante (criterio: Prob. de F para entrar <= ,050)
2	x1	.	Hacia adelante (criterio: Prob. de F para entrar <= ,050)

a. Variable dependiente: y

**ANOVA<sup>c</sup>**

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	1831,896	1	1831,896	22,799	,001 <sup>a</sup>
	Residual	883,867	11	80,352		
	Total	2715,763	12			
2	Regresión	2641,001	2	1320,500	176,627	,000 <sup>b</sup>
	Residual	74,762	10	7,476		
	Total	2715,763	12			

a. Variables predictoras: (Constante), x4

b. Variables predictoras: (Constante), x4, x1

c. Variable dependiente: y

**Resumen del modelo**

Modelo	R	R cuadrado	R cuadrado corregido	Error tip. de la estimación
1	,821 <sup>a</sup>	,675	,645	8,96390
2	,986 <sup>b</sup>	,972	,967	2,73427

a. Variables predictoras: (Constante), x4

b. Variables predictoras: (Constante), x4, x1

**Coefficientes<sup>b</sup>**

Modelo		Coefficients no estandarizados		Coefficients tipificados	t	Sig.	Estadísticos de colinealidad	
		B	Error tip.	Beta			Tolerancia	FIV
1	(Constante)	117,568	5,262		22,342	,000	1,000	1,000
	x4	-,738	,155	-,821	-4,775	,001		
2	(Constante)	103,097	2,124		48,540	,000	,940	1,064
	x4	-,614	,049	-,683	-12,621	,000		
	x1	1,440	,138	,563	10,403	,000		

a. Variable dependiente: y

# Selección de variables en el modelo de regresión lineal múltiple

Pasos sucesivos (opción Pasos sucesivos de SPSS)

**Variables introducidas/eliminadas<sup>a</sup>**

Modelo	Variables introducidas	Variables eliminadas	Método
1	x4	.	Por pasos (criterio: Prob. de F para entrar <= ,050, Prob. de F para salir >= ,100).
2	x1	.	Por pasos (criterio: Prob. de F para entrar <= ,050, Prob. de F para salir >= ,100).

a. Variable dependiente: y

**ANOVA<sup>c</sup>**

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	1831,896	1	1831,896	22,799	,001 <sup>a</sup>
	Residual	883,867	11	80,352		
	Total	2715,763	12			
2	Regresión	2641,001	2	1320,500	176,627	,000 <sup>b</sup>
	Residual	74,762	10	7,476		
	Total	2715,763	12			

a. Variables predictoras: (Constante), x4

b. Variables predictoras: (Constante), x4, x1

c. Variable dependiente: y

**Resumen del modelo**

Modelo	R	R cuadrado	R cuadrado corregido	Error típ. de la estimación
1	,821 <sup>a</sup>	,675	,645	8,96390
2	,986 <sup>b</sup>	,972	,967	2,73427

a. Variables predictoras: (Constante), x4

b. Variables predictoras: (Constante), x4, x1

**Coefficientes<sup>a</sup>**

Modelo		Coeficientes no estandarizados		Coeficientes tipificados		t	Sig.	Estadísticos de colinealidad	
		B	Error típ.	Beta				Tolerancia	FIV
1	(Constante)	117,568	5,262			22,342	,000		
	x4	-,738	,155	-,821		-4,775	,001	1,000	1,000
2	(Constante)	103,097	2,124			48,540	,000		
	x4	-,614	,049	-,683		-12,621	,000	,940	1,064
	x1	1,440	,138	,563		10,403	,000	,940	1,064

a. Variable dependiente: y