EJEMPLO COMPLETO DE AJUSTE DE REGRESIÓN LINEAL

El experimento realizado ha tenido por objetivo el estudio de la dependencia de la resistencia eléctrica de un hilo de cobre con la temperatura; teóricamente esta relación responde a la fórmula **$R= R_o (1 + \alpha t)$**, donde R y R_o se miden en ohmios (Ω) y t en grados centígrados (^oC). Se ha obtenido la siguiente tabla de registros experimentales:

t (^oC)	R (Ω)				
30	2.242	2.247	2.245	2.248	2.243
40	2.331	2.330	2.332	2.329	2.328
60	2.482	2.480	2.483	2.484	2.481
75	2.599	2.597	2.595	2.597	2.597
95	2.765	2.763	2.767	2.765	2.765

La incertidumbre experimental en las medidas de la temperatura es de ± 0.5 ^oC y en el caso de las resistencias ± 0.001 Ω. El análisis estadístico de los datos tendrá por objetivo, en primer lugar, demostrar que efectivamente la dependencia de R con t es lineal, y una vez comprobada esta hipótesis, calcular los valores de los parámetros R_o y Ω con sus incertidumbres.

Como primer paso se calcularán los valores de la resistencia, y sus incertidumbres, correspondientes a cada temperatura, para lo cual se aplicarán las fórmulas pertinentes del *valor medio*, $< X > = \frac{\sum_1^5 X_i}{5}$, y de la *desviación media*, $s = \sqrt{\frac{\sum_1^5 (X_i - < X >)^2}{5 - 1}}$, y de la *desviación de la media*, $s_{< X >} = \frac{s}{\sqrt{5}}$; el valor del *intervalo de confianza* se calculará mediante el **criterio de Student**, **$\Delta = t\ s_{<X>}$**, donde el parámetro t para un *nivel de significación* de 0.05, o sea *nivel de confianza* del 95%, y 4 *grados de libertad* (5 medidas) vale 2.7764. Se obtiene la siguiente tabla de valores (ver la *Hoja de Cálculo* construida para el ejemplo):

t (^oC)	R (Ω)
30.0 ± 0.5	2.245 ± 0.003
40.0 ± 0.5	2.330 ± 0.003
60.0 ± 0.5	2.482 ± 0.003
75.0 ± 0.5	2.597 ± 0.003
95.0 ± 0.5	2.765 ± 0.003

Por razones de comodidad, consideramos que todos los valores de R están afectados por la misma incertidumbre, asignando el *intervalo de confianza* ± 0.003 Ω para todos los valores; aunque según se observa en la *Hoja* este valor corresponde al Δ de 30 ^oC, si consideramos que la incertidumbre instrumental es ± 0.001 Ω, podemos asignar el valor de 0.003 Ω para todos. Una posibilidad, que se utiliza con frecuencia, sería sumar la incertidumbre instrumental, en este caso 0.001 Ω, a los valores de Δ para obtener el intervalo de confianza definitivo; en este caso las incertidumbres de los valores de R no serían todas iguales, y, rigurosamente, habría que utilizar un análisis de regresión ponderado, complicando los cálculos sin que supusiese una diferencia sustancial .

A continuación sobre esta tabla de valores t ⟵⟶ R se calcularía el **coeficiente de correlación lineal** $r = \frac{\sum_1^5 (X_i - \langle X \rangle)(Y_i - \langle Y \rangle)}{\sqrt{\sum_1^5 (X_i - \langle X \rangle)^2 (Y - \langle Y \rangle)^2}}$. Se obtendría el valor de r= 0.9998 con una desviación , $s(r) = \sqrt{\frac{1 - r^2}{n-2}}$ donde n= 5, s(r)= 0.0118; de acuerdo con estos valores, puede asegurarse que la dependencia de R con t es lineal con más de un 99% de probabilidades, y por tanto es pertinente un ajuste de regresión lineal.

Finalmente se realizaría el ajuste por mínimos cuadrados, **R= m t + p**, donde los parámetros m y p se calcularían con las fórmulas:

$$m = \frac{\sum_1^5 (X_i - \langle X \rangle) Y_i}{\sum_1^5 (X_i - \langle X \rangle)^2} \quad ,,, \quad p = \langle Y \rangle - m \langle X \rangle$$

Se obtienen los valores m= 0.0079 y p= 2.0089; o sea, de acuerdo con la expresión:
$R = R_o (1 + \alpha t)$ resultarían los valores:

R_o= 2.0089 Ω ; αR_o= 0.0079 Ω $^oC^{-1}$ => α = 0.0039 $^oC^{-1}$

Para calcular las incertidumbres de estos valores, calculemos en primer lugar las desviaciones de los parámetros p y m, por medio de las expresiones:

$$s_m = s \sqrt{\frac{n}{n \sum_1^5 X_i^2 - (\sum_1^5 X_i)^2}} \quad ,,, \quad s_p = s \sqrt{\frac{\sum_1^5 X_i^2}{n \sum_1^5 X_i^2 - (\sum_1^5 X_i)^2}} \quad ,,, \quad s = \sqrt{\frac{\sum_1^5 [Y_i - (m X_i + p)]^2}{n-2}}$$

Se obtienen los valores **s(m)= 0.000094 y s(p)= 0.006031**.

Para obtener las *incertidumbres* de s(m) y s(p) multiplicamos a las desviaciones por el correspondiente *parámetro t de Student*, en este caso 3.1825 porque ahora hay *tres grados de libertad* (5-2) al considerar dos parámetros, y dividimos por √5, dado que 5 es el número de medidas de las muestras. Como resultado se obtienen las incertidumbres de m y p:

E(m)= 0.000133 ,,, E(p)= 0.00858

En primera instancia ya podemos afirmar, después de redondear a cuatro cifras significativas incluida la estimada, que el valor de R_o es:

R_o= (2.009 ± 0.009) Ω

Para calcular el valor de la incertidumbre del parámetro α debemos aplicar la fórmula de la *propagación cuadrática de errores* a la expresión **α=m/R_o**; si utilizamos las incertidumbres de m y p en vez de sus desviaciones obtendríamos directamente E(α):

$$E^2(\alpha) = \frac{1}{R_o^2} E^2(m) + \left(\frac{m}{R_o^2} \right)^2 E^2(R_o)$$

Resulta un valor de E(α)= 6.84x10^{-5} y por tanto, si redondeamos al número mínimo de cifras para poder considerar este último valor, podemos finalmente escribir que:

α = (0.0039 ± 0.0001) $^oC^{-1}$ = (39 ± 1)x 10^{-4} $^oC^{-1}$