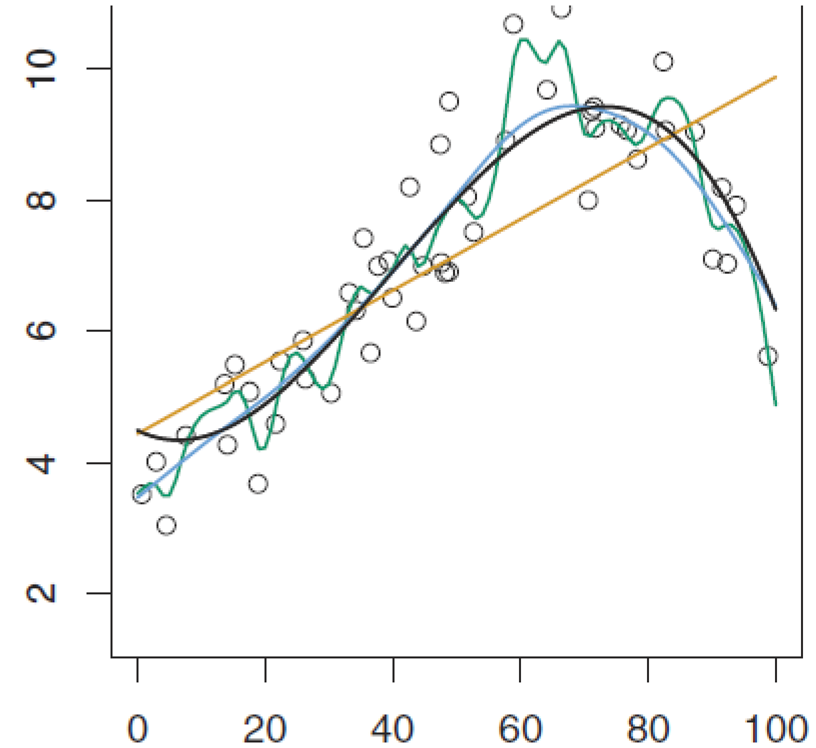
Evaluando la precisión del modelo

Debemos mensurar qué tan bien las respuestas que arroja el modelo predicen los valores observados.
En regresión se utiliza el error cuadrático medio

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2$$



El MSE se calcula utilizando los datos de entrenamiento. En situaciones de la vida real nos interesa la precisión de las predicciones que obtenemos cuando aplicamos nuestro modelo a datos **no vistos anteriormente.**

observaciones de entrenamiento $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$

Entonces podemos calcular $\hat{f}(x_1), \hat{f}(x_2), \ldots, \hat{f}(x_n)$

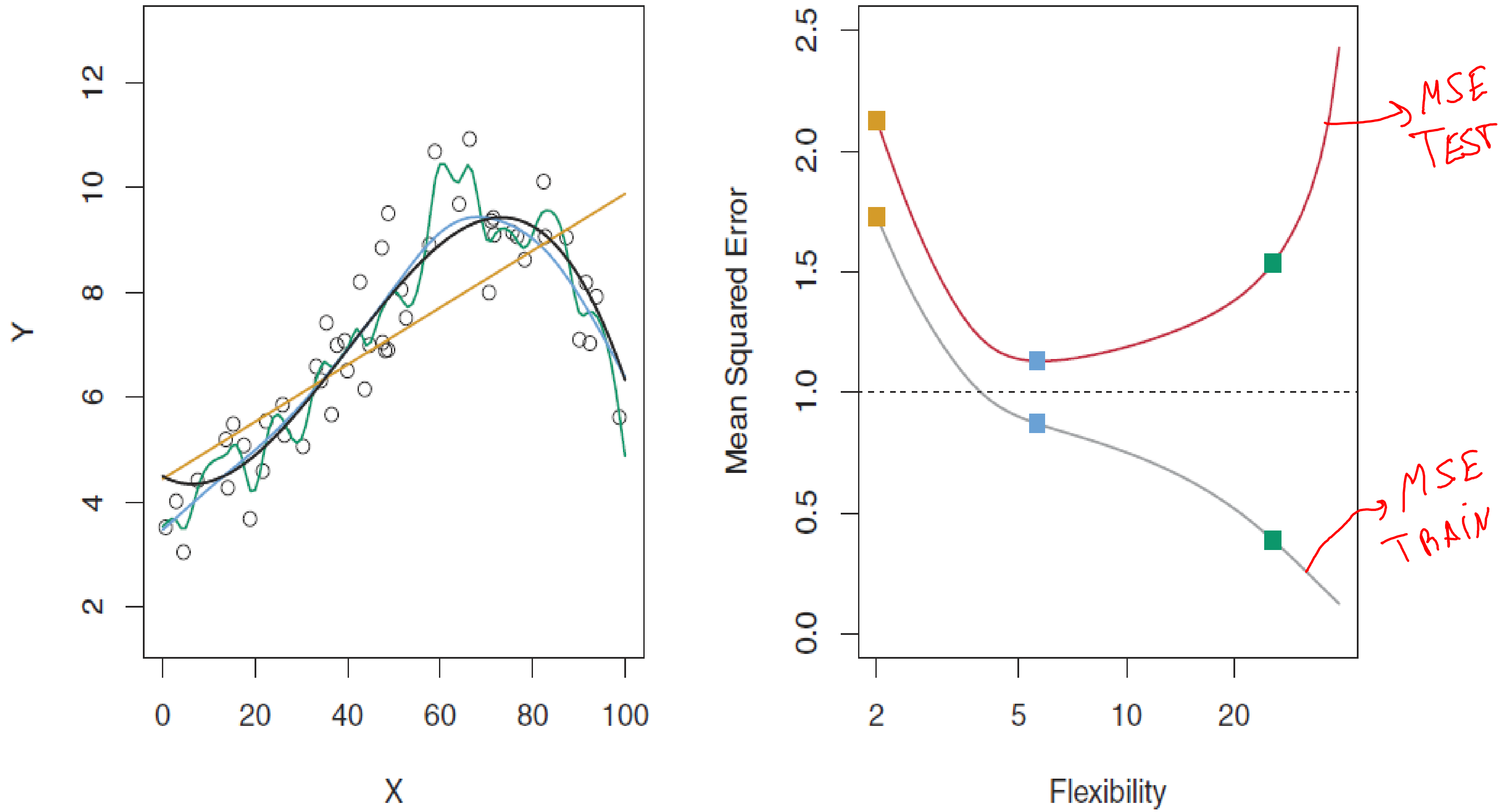
Si estos son aproximadamente iguales a $y_1, y_2, \ldots, y_n$

-→ **entonces sabiendo que el MSE de entrenamiento es pequeño**

queremos saber si $\hat{f}(x_0)$ es aproximadamente igual y_0 ,donde (x_0, y_0) es una observación de prueba **no vista previamente y no utilizada para entrenar** el modelo

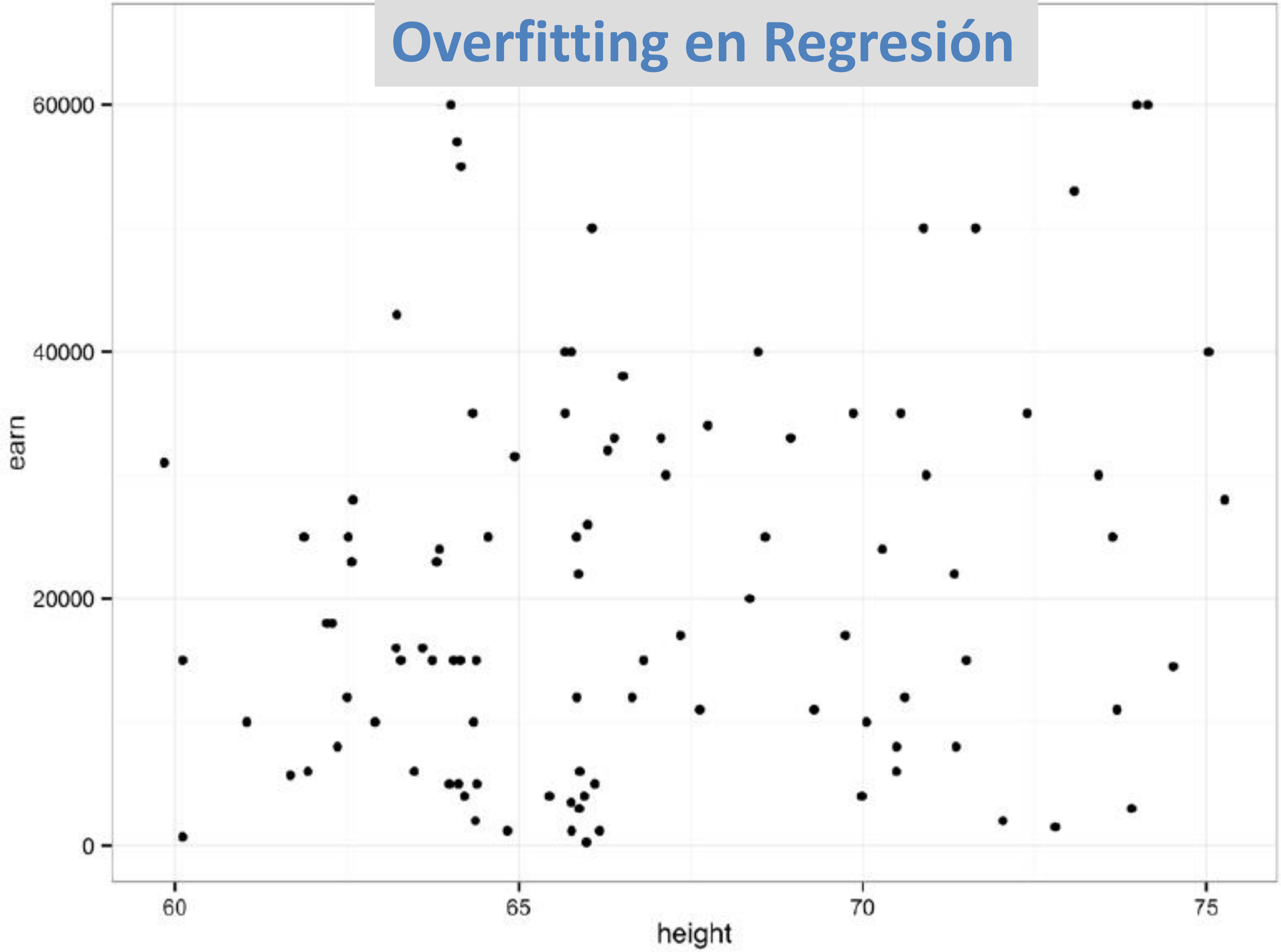
$$\text{test MSE} = \text{Ave}(y_0 - \hat{f}(x_0))^2$$

-→ **Este valor, el Error Cuadrático Medio del set de PRUEBA, es el que me interesa minimizar**

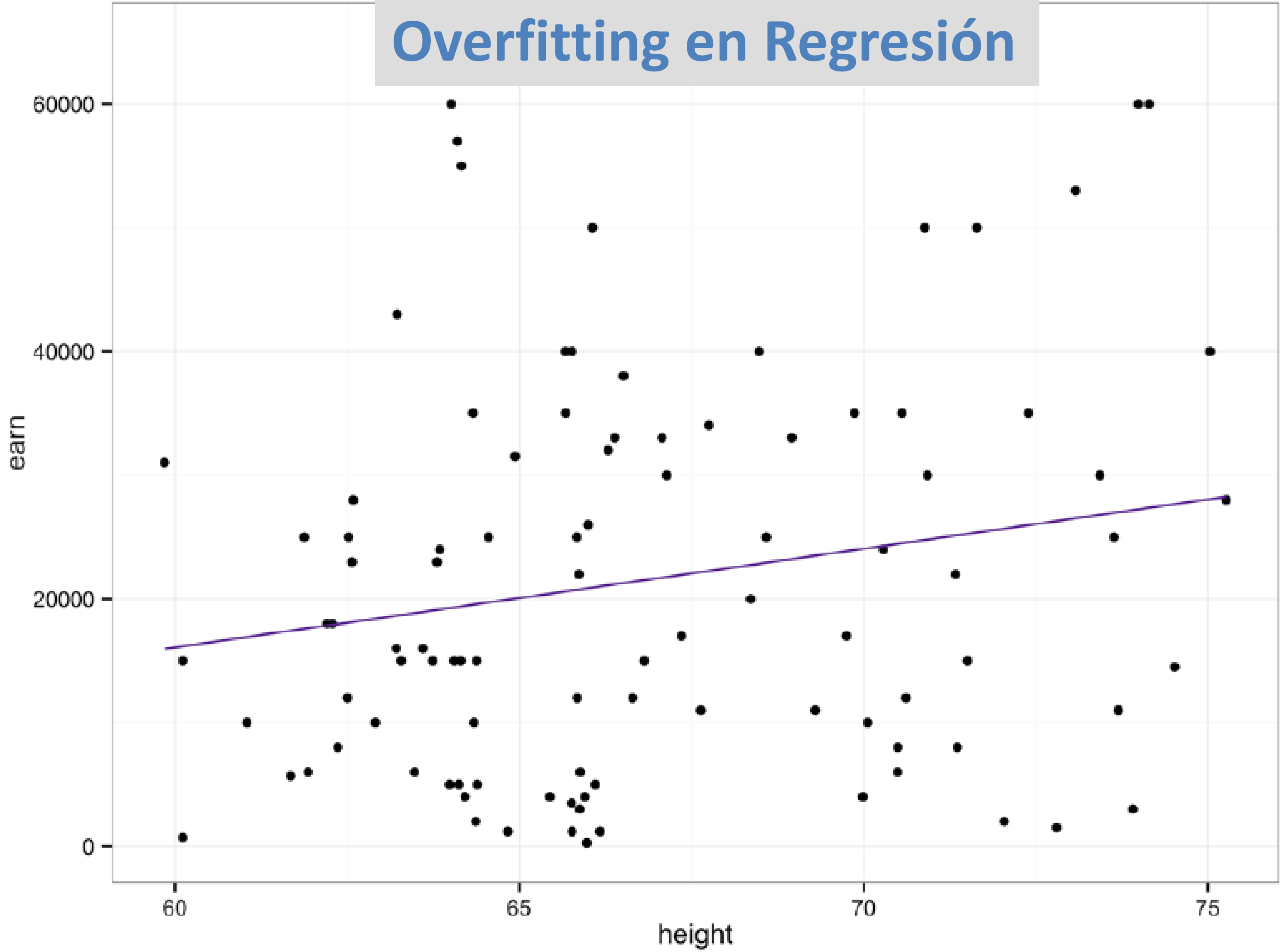


Las curvas naranja, azul y verde del gráfico de la izquierda muestra tres posibles estimaciones para f obtenidas usando métodos con niveles crecientes de flexibilidad. A la derecha vemos sus correspondientes MSE de entrenamiento y de prueba (abajo y arriba resp)

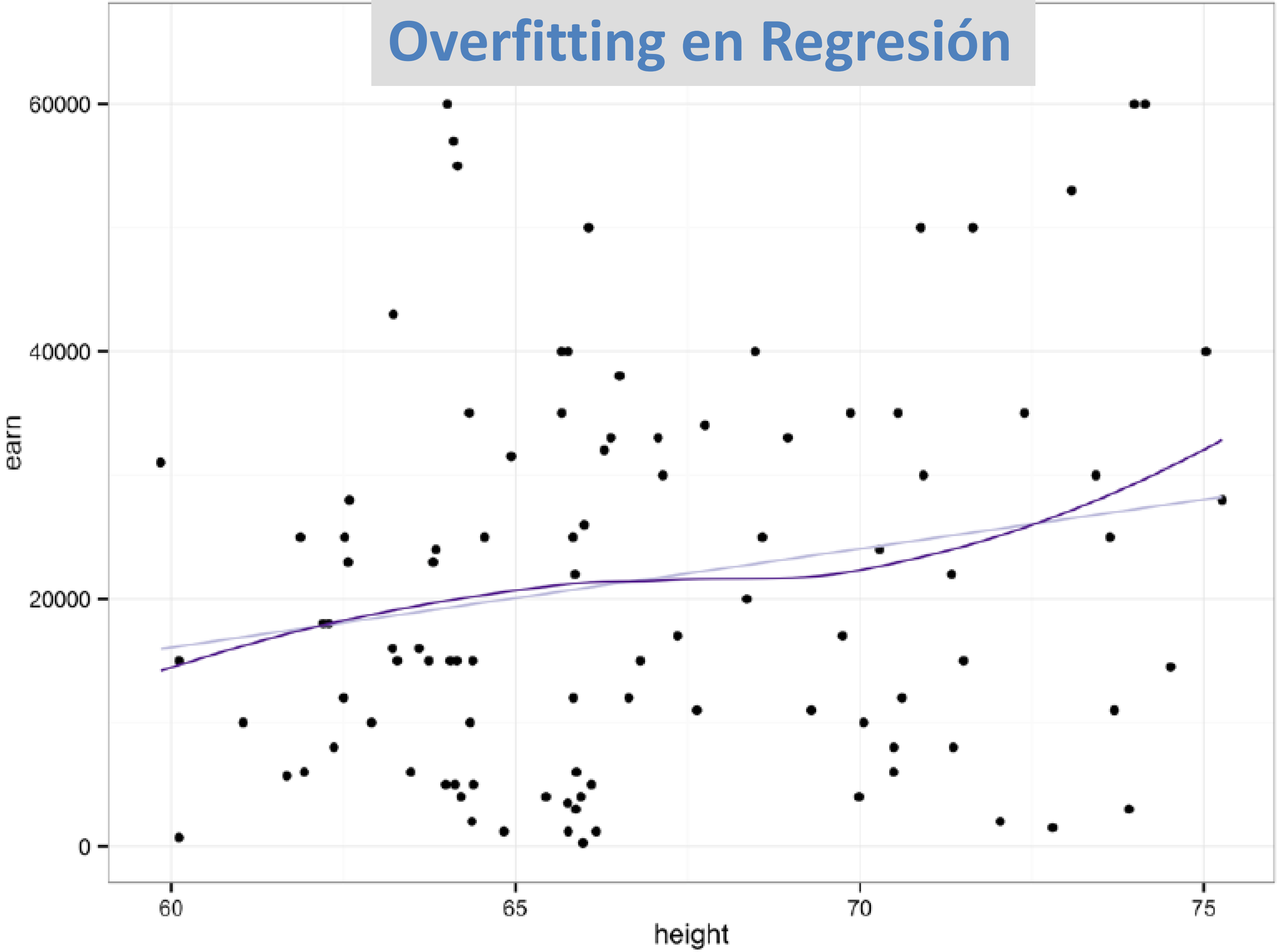
Overfitting en Regresión



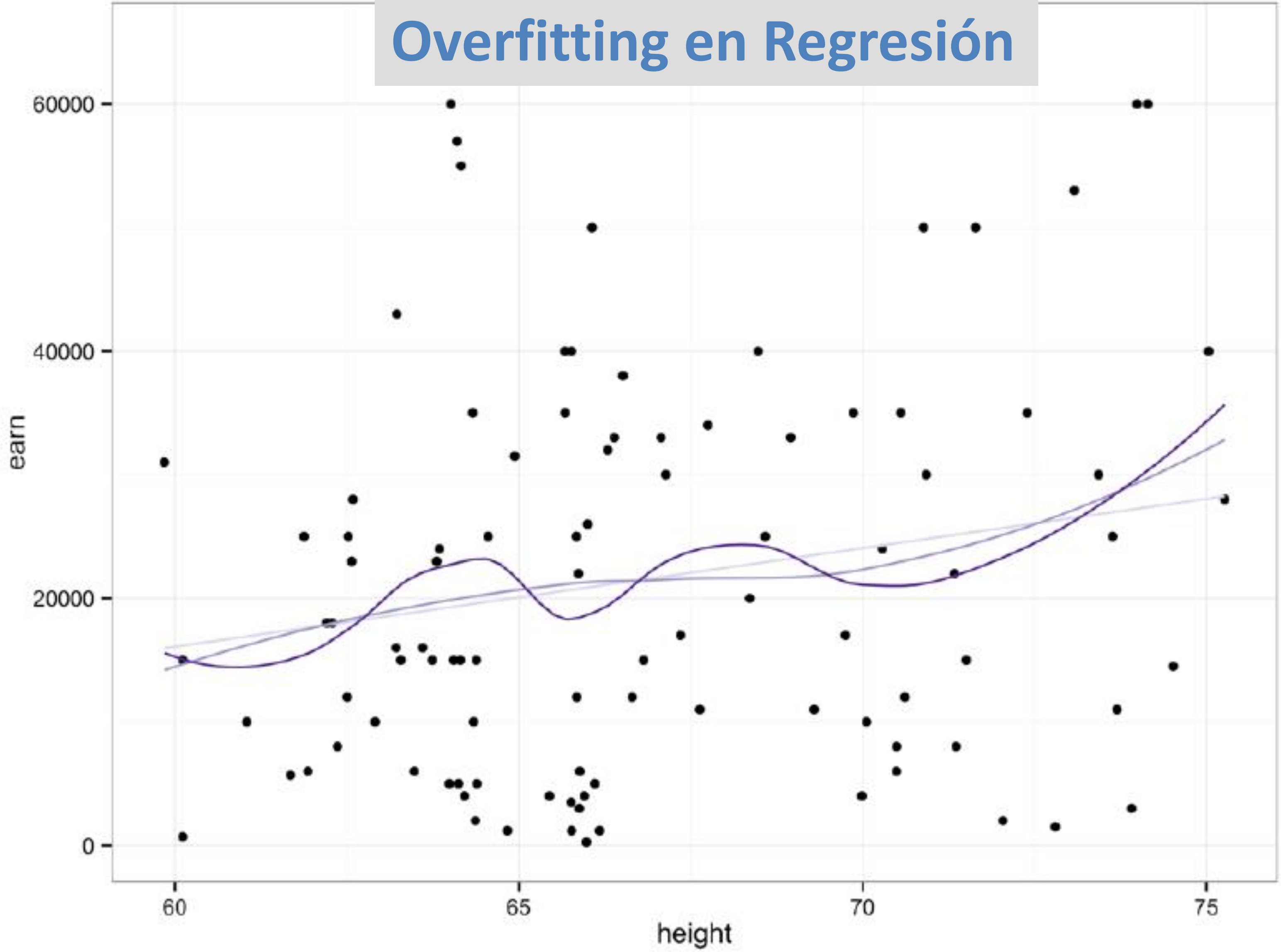
Overfitting en Regresión



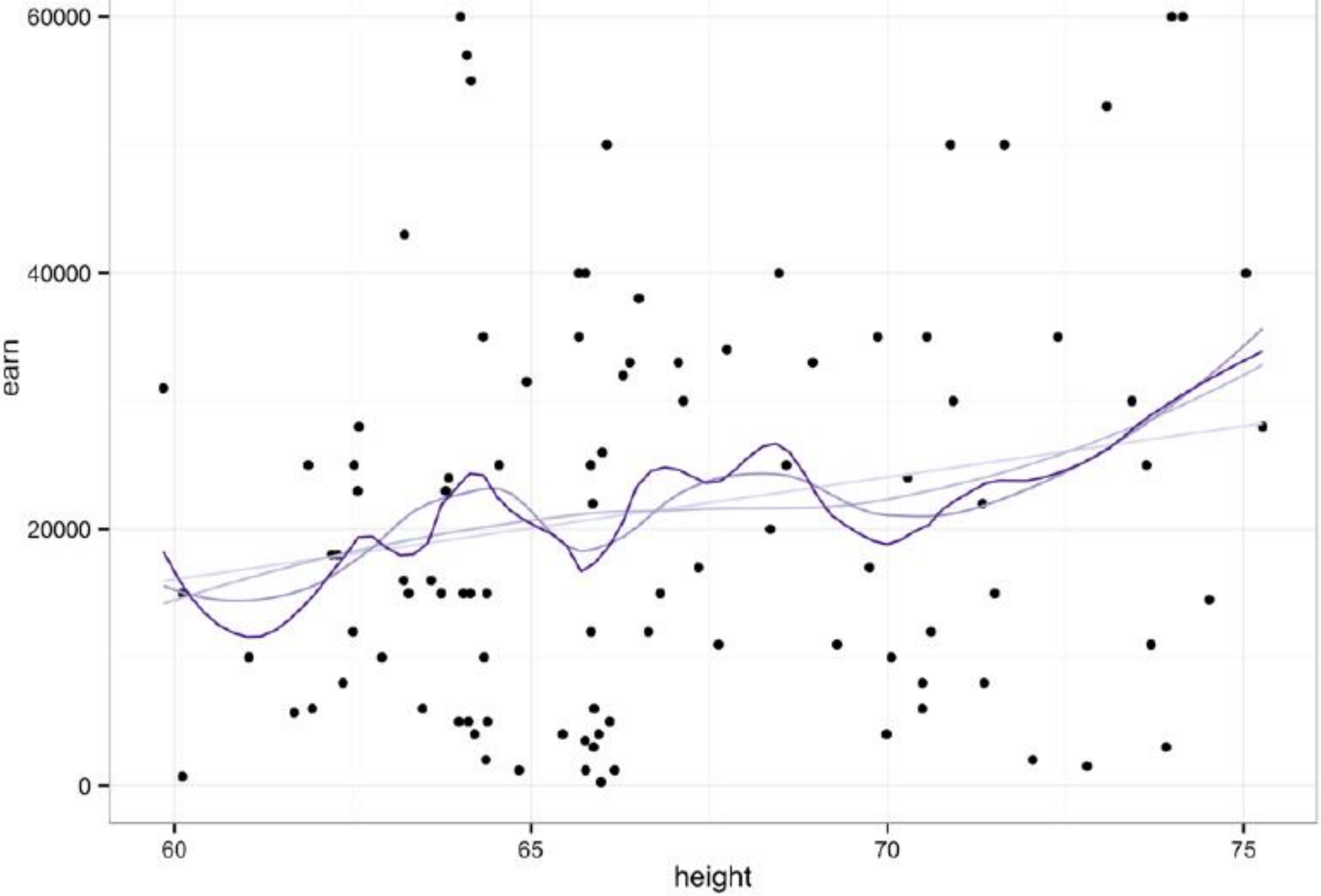
Overfitting en Regresión



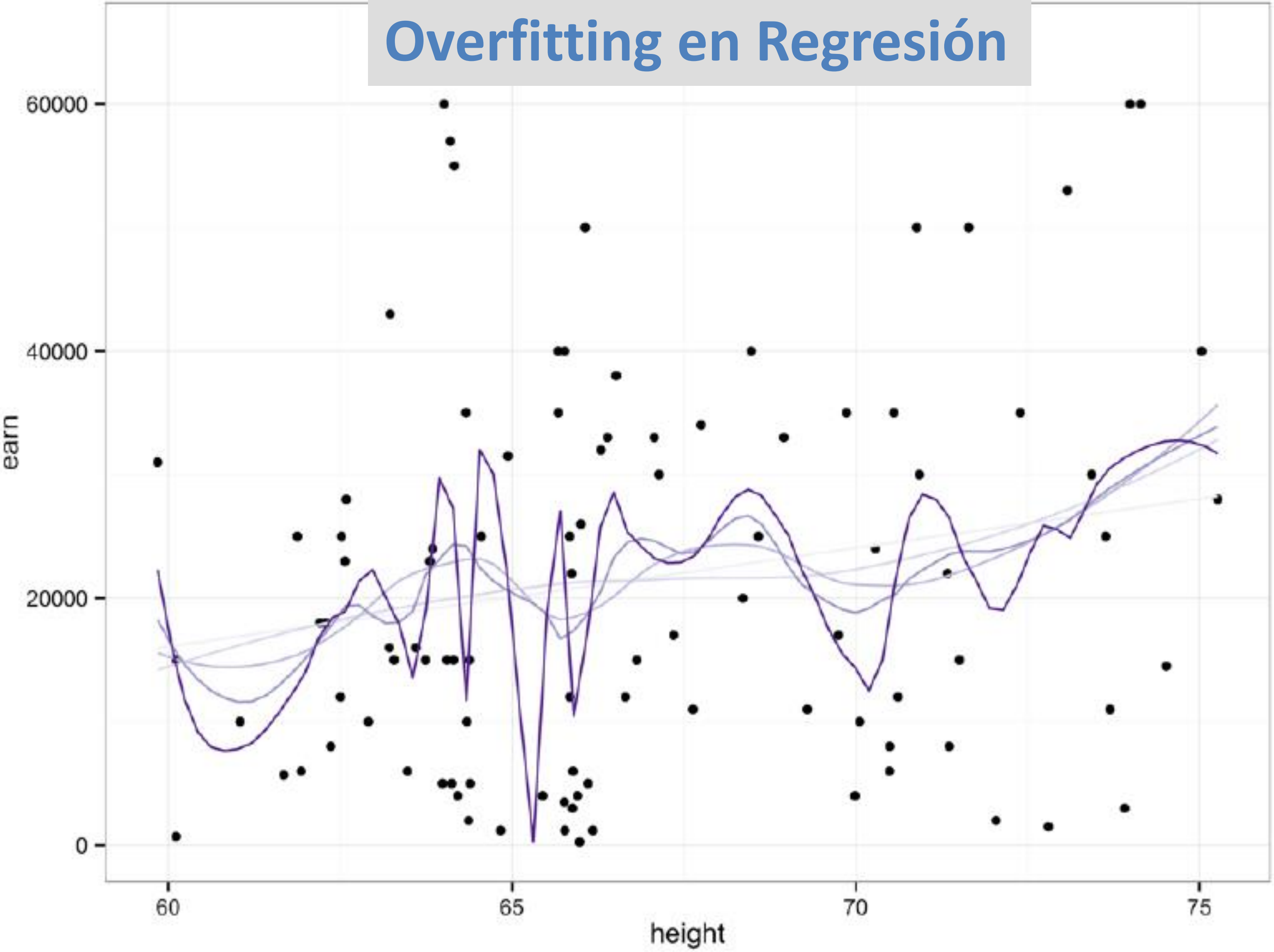
Overfitting en Regresión



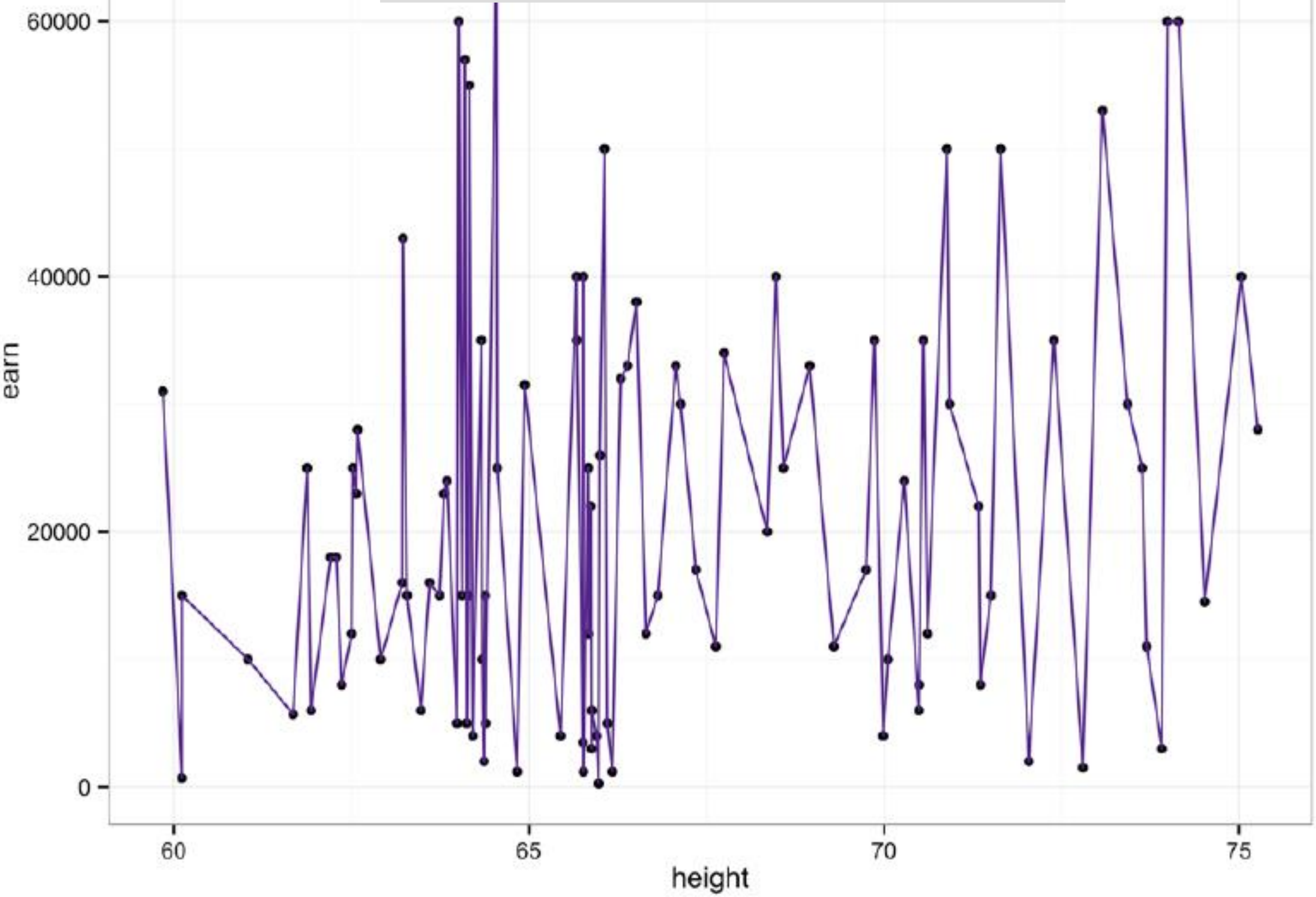
Overfitting en Regresión



Overfitting en Regresión



Overfitting en Regresión



Evaluando la precisión en modelos de clasificación

Supongamos que buscamos estimar ***f*** sobre la base de observaciones de entrenamiento $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, donde ahora $y_1, \ldots, y_n$ son **cualitativos.**

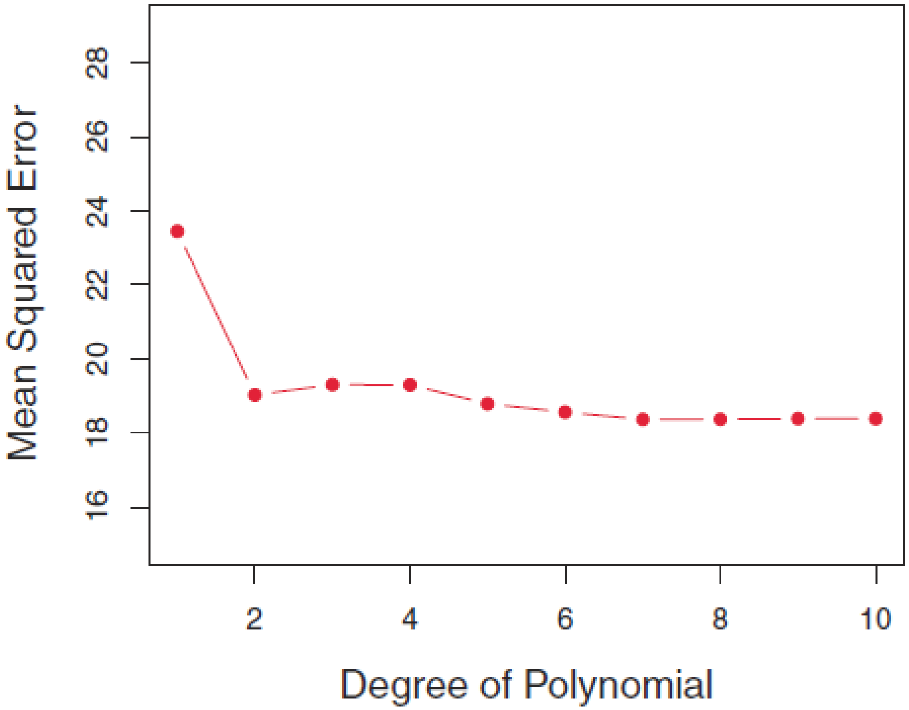
El enfoque más común para cuantificar la **precisión** de nuestra estimación f es la **tasa de error de entrenamiento**, que es la proporción de errores que se cometen si aplicamos nuestra estimación f para las observaciones de entrenamiento:

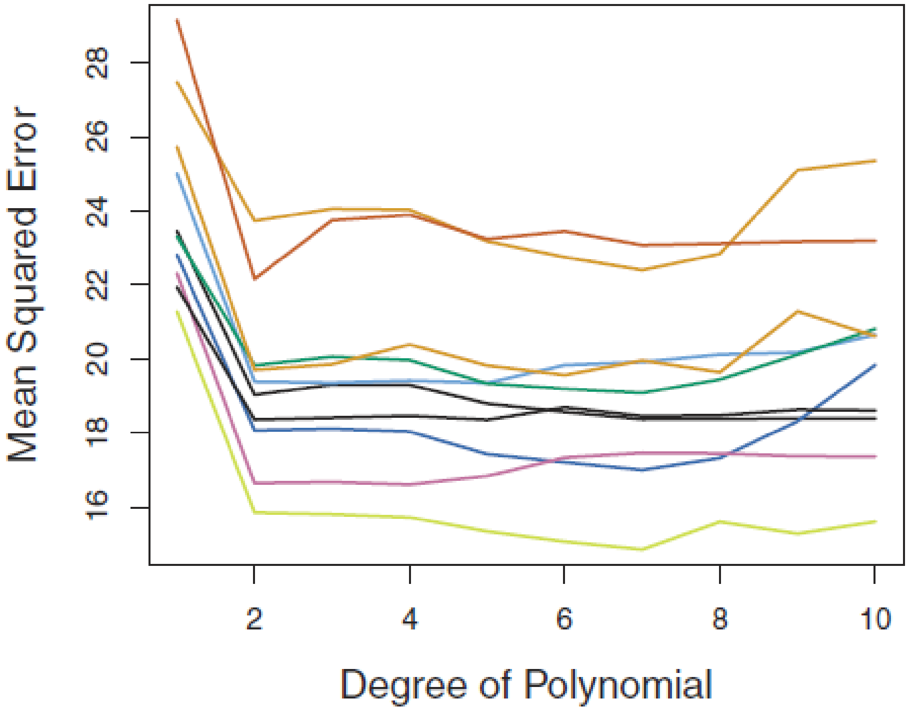
$$\frac{1}{n}\sum_{i=1}^{n} I(y_i \neq \hat{y}_i)$$

Al igual que en la configuración de regresión, estamos más interesados en las tasas de error que resultan de aplicar nuestro clasificador **a las observaciones de prueba que no se usaron en el entrenamiento.**

$$\mathrm{Ave}\left(I(y_0 \neq \hat{y}_0)\right)$$

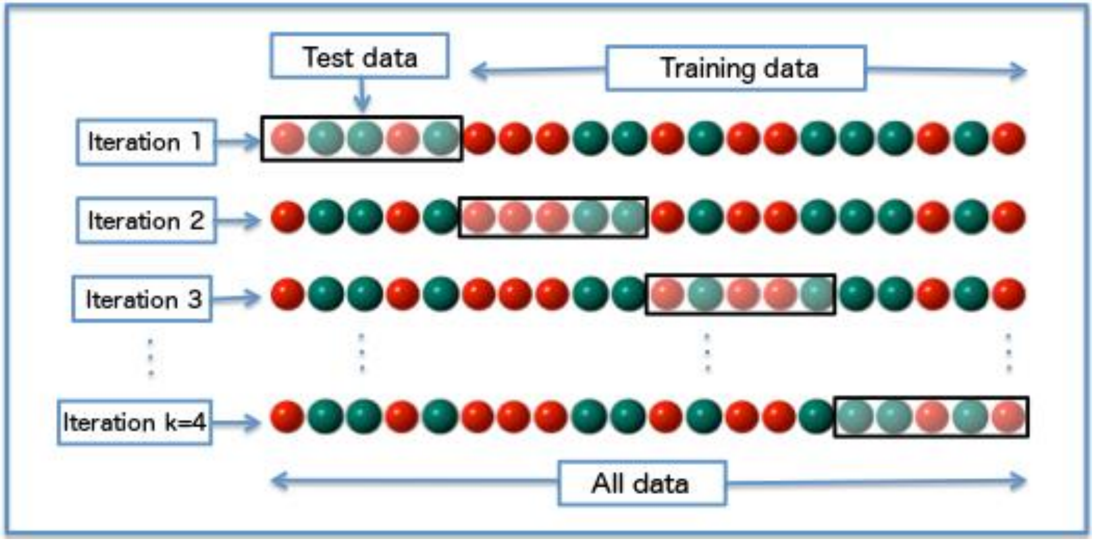
Diferentes set de validación





Validación Cruzada de k iteraciones

Separamos el dataset en k partes iguales y entrenamos k modelos eligiendo el set de validación de cada uno de la siguiente manera:



Para luego tener una mejor estimar del error de prueba!

$$\boxed{CV_{(k)} = \frac{1}{k}\sum_{i=1}^{k} MSE_i}$$

Validación Cruzada de k iteraciones en Clasificación

$$\boxed{\mathrm{CV}_{(k)} = \frac{1}{k}\sum_{i=1}^{k}\mathrm{Err}_i}$$

Siendo $\mathrm{Err}_i = I(y_i \neq \hat{y}_i)$

Entrenamiento del Modelo

K=3