



UNIVERSIDAD
DE SANTIAGO
DE CHILE

Trabajo ML Finanzas

Magister en Economía Financiera

Ignacio Sepúlveda

Profesor

Robinson Dettoni

Cliff Bahamondes

Fecha

14 de Sept, 2022

Respuesta 1

Para la **figura 2.25** vemos que el log ratio se mantiene mas o menos constante, hasta que llega al valor aproximado de 550 donde empieza a decrecer. La dispersión hasta los 600 parece ser aproximadamente constante después de esos valor existe mas amplitud en los valores. Evidentemente un ajuste lineal no parece ser el correcto.

En la **figura 2.26** cuando ajustamos modelos polinomiales vemos que mientras mas grados tenga mas flexibles serán. En general los 3 lo hacen bien y parecido al principio pero al final del rango presentan mayor dificultad para capturar el exceso de dispersión comparado al rango del principio. El libro además menciona los problemas de estabilidad de polinomios de alto grado, lo que eventualmente afectaría el resultado fuera de muestra.

Figure 2.25

Scatter plot

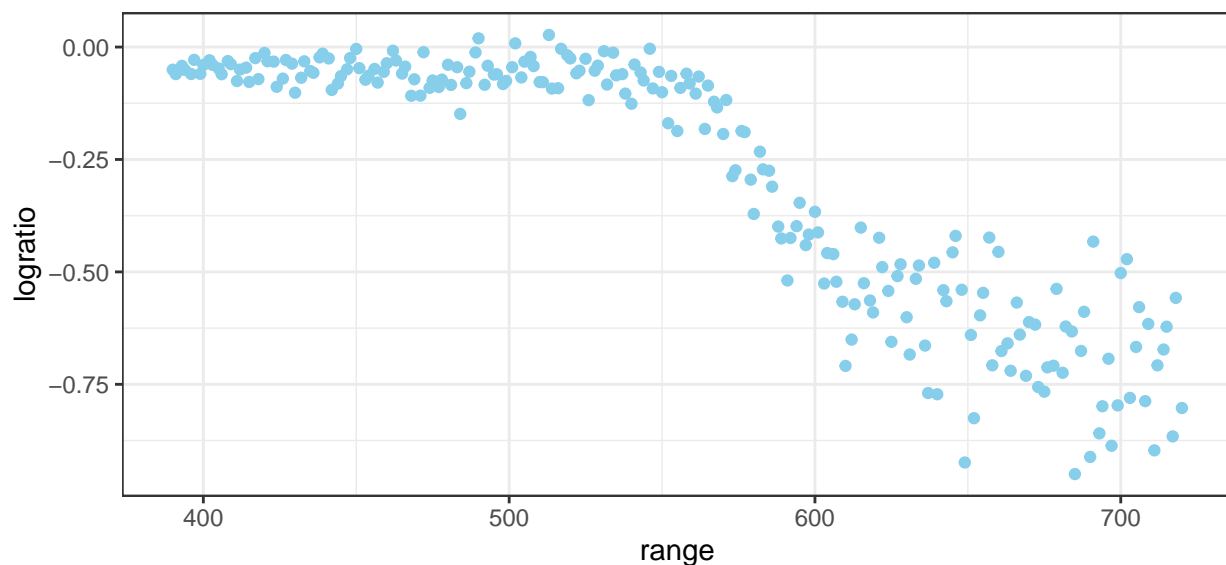
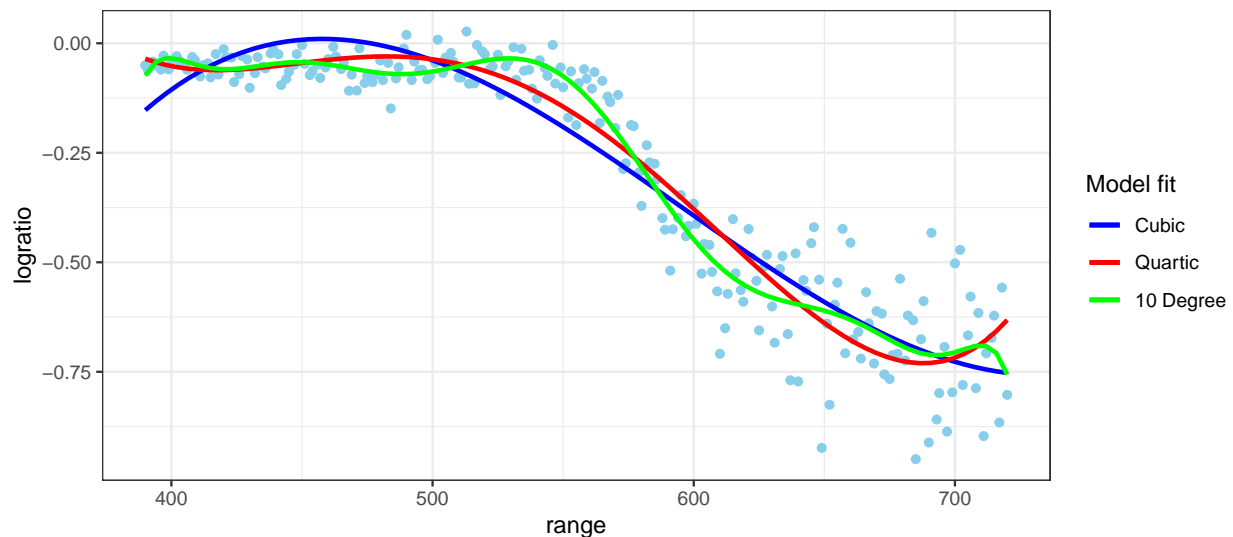


Figure 2.26

Polynomial regression Degree(3,4,10)



Respuesta 2

Vemos, en **figura 3.6**, que el ajuste con dos knots igual es bueno, presenta mas parsimonia que los modelos polinomiales. Pero igual tiene errores como que al principio tiene un tendencia decreciente cuando los puntos parecen estar mas planos y siguen siendo un problema la volatilidad de rangos altos.

La introducción de dos knots mas, en **figura 3.7**, mejora la calidad del ajuste tanto al inicio como al final. Pero sigue faltando mejoras sobretodo al final.

Para la **figura 3.8** ocurre algo similar a cuando aumentamos el grado del polinomio en cuando a la flexibilidad, mejora el ajuste al principio pero sobre todo al final . Pero es evidente que podríamos tener problemas de sobreajuste.

Finalmente en la **figura 3.9**, vemos que sigue siendo un buen ajuste, pero el sobreajuste en el rango final disminuyo.

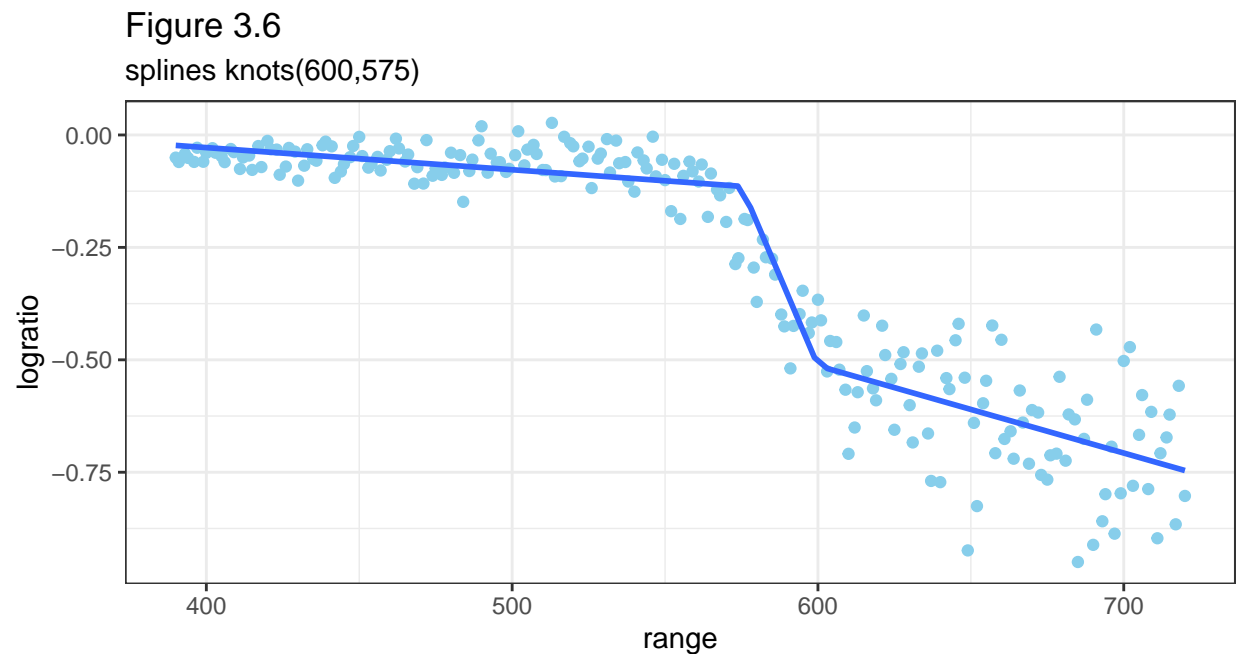


Figure 3.7

splines knots(500,550,600,650)

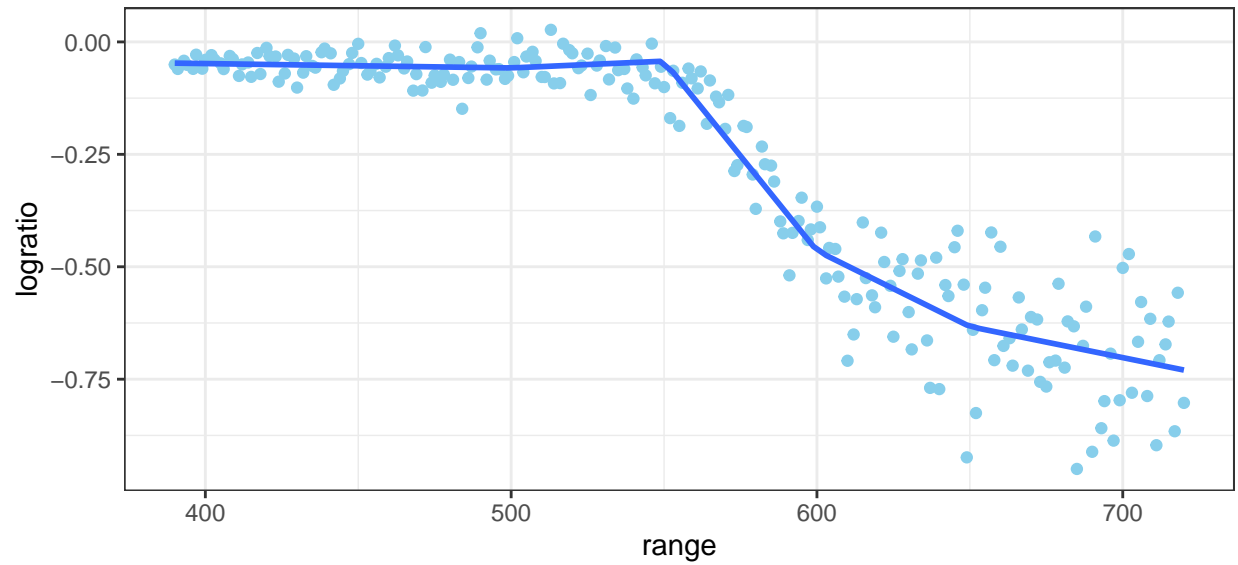


Figure 3.8

splines knots(400,412.5,...,700)

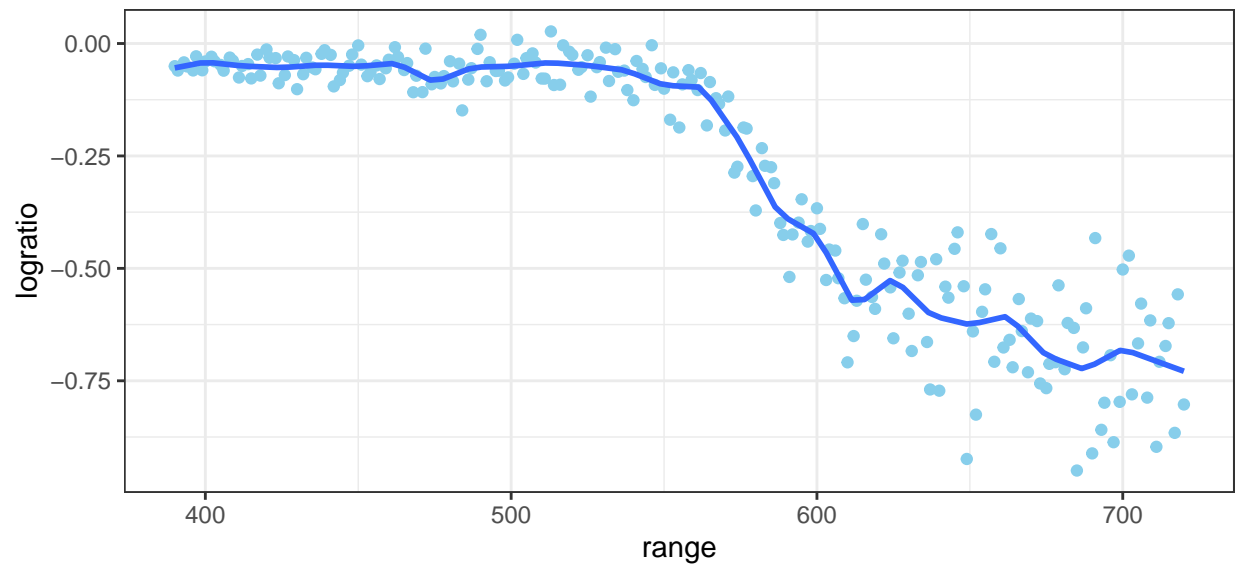
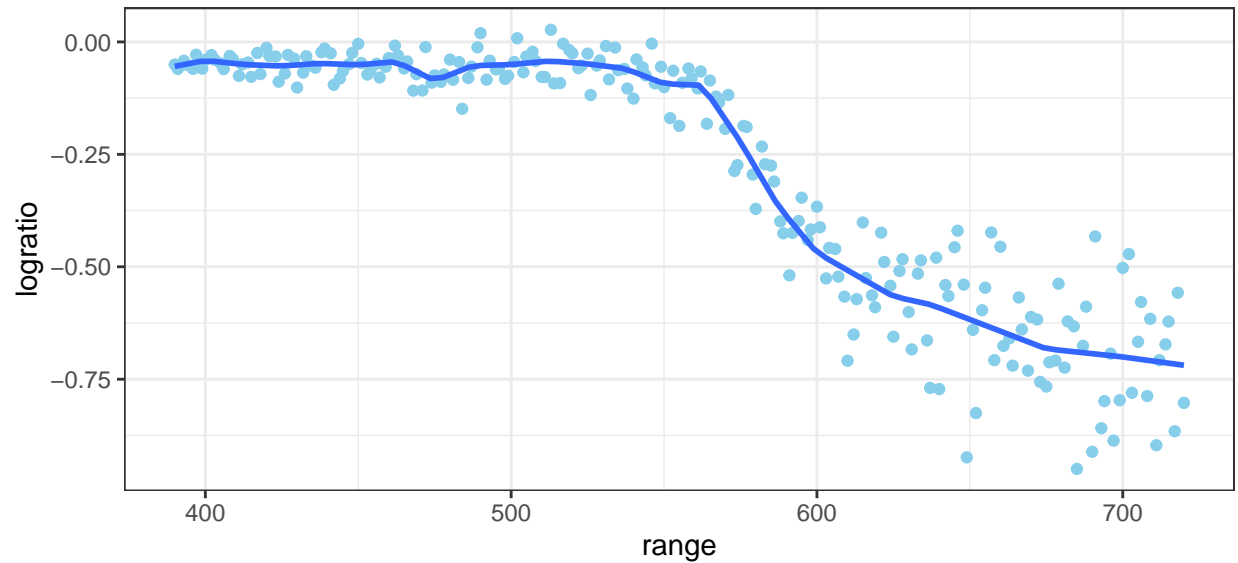


Figure 3.9

splines knots(400,412.5,...,700) sin (612.5,650,662.5,687.5)



Respuesta 3

En la **figura 3.10** el ajuste es mejor, logramos suavizar y mejoramos la captura de la tendencia en los datos.

En la **figura 3.11** notamos que a medida que aumenta el λ la curva cada vez es menos flexible, lo que podría permitir obtener mejores y más robustas predicciones fuera de muestra, si logramos optimizar este parámetro λ .

Figure 3.10

splines knots(400,412.5,...,700) con $\lambda=30$

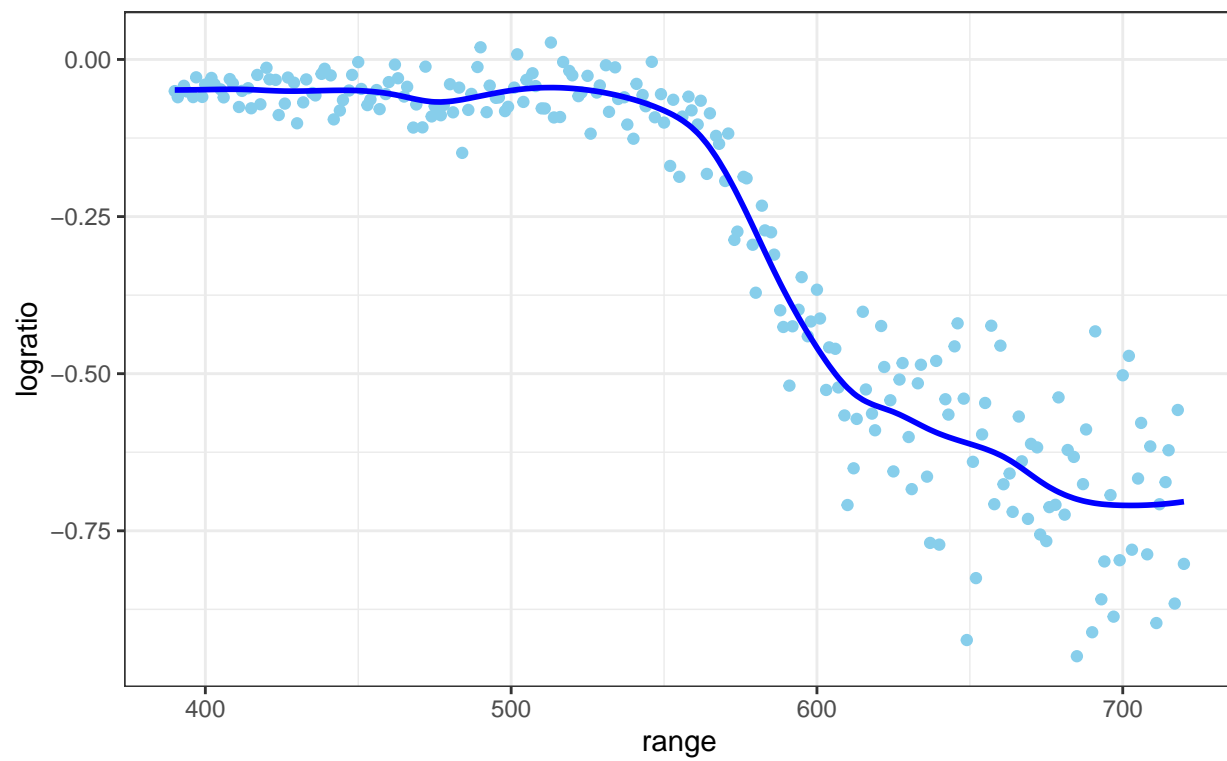
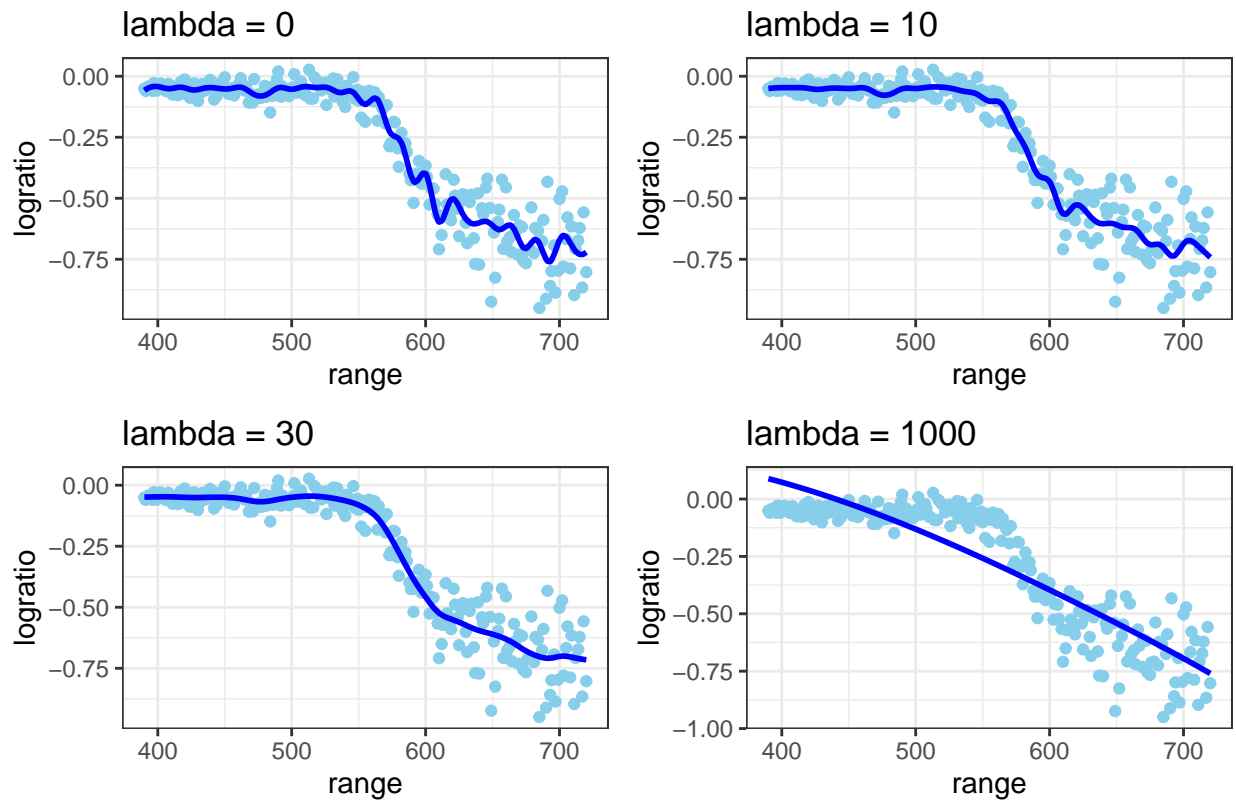


Figure 3.11



Respuesta 4

La **figura 3.20** nos muestra que el ajuste es casi idéntico para estos dos λ . La diferencia en grado de libertad es pequeña y eso se ve refleja en la poca diferencia de los ajustes.

La **figura 3.21** nos muestra la relación monotonica y decreciente de la $\lambda - \log(\lambda)$ vs los grados de libertad, lo que indica que para cada λ solo existe un correspondiente df.

Finalmente, la **figura 3.22** agrega otra linea a la figura anterior, vemos que la mayor diferencia entre la dos es al centro del gráfico. Dado que los extremos converge hacia una regresión paramétrica.

Figure 3.20

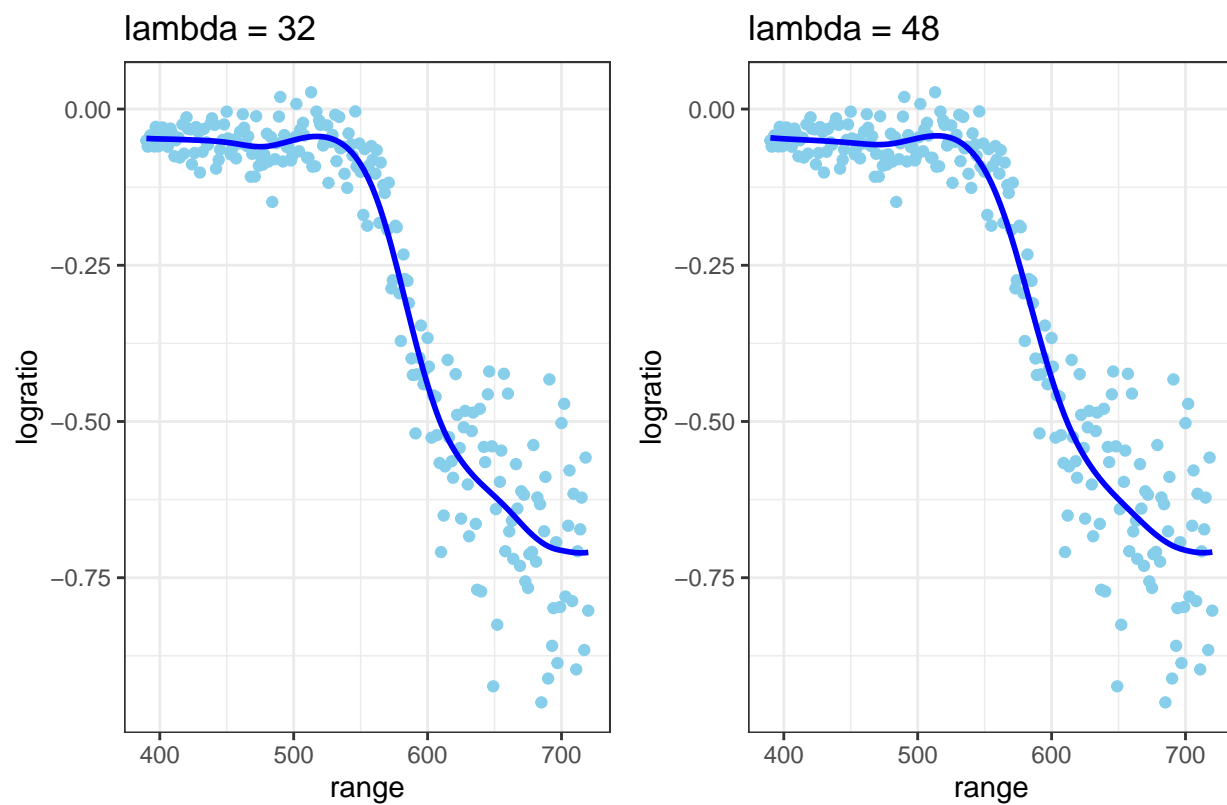


Figure 3.21

df(lambda) vs log(lambda)

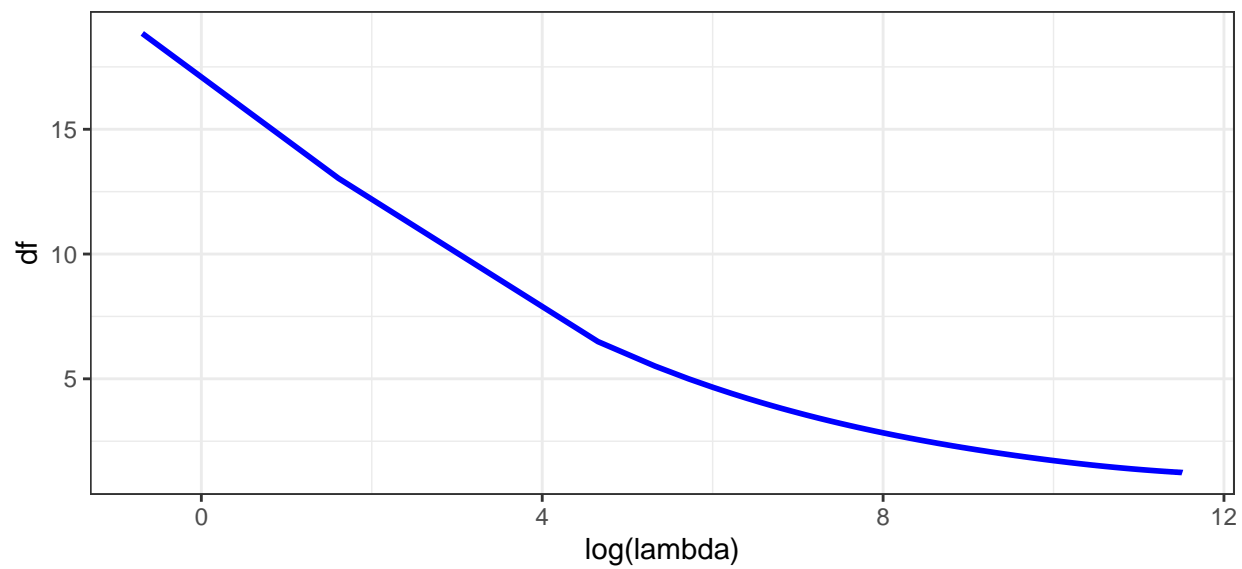
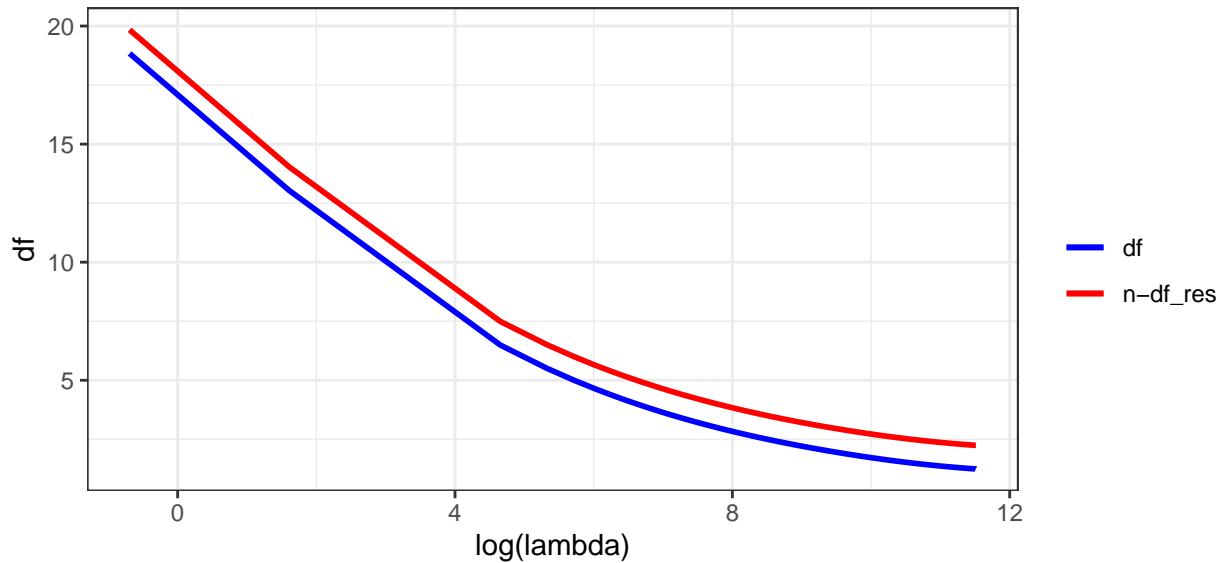


Figure 3.22

$df(\lambda)$ y $n-df_{res}$ vs $\log(\lambda)$



Respuesta 5

La **figura 5.2** nos muestra que los splines penalizados son extremadamente similares independiente del método de selección escogido, ML o REML.

La **figura 5.3** muestra que CV el λ minimiza en la parte donde los RSS empieza a converger de manera asintótica mientras CV crece lo que da estabilidad al mínimo.

Figure 5.2

Automatic penalized linear spline

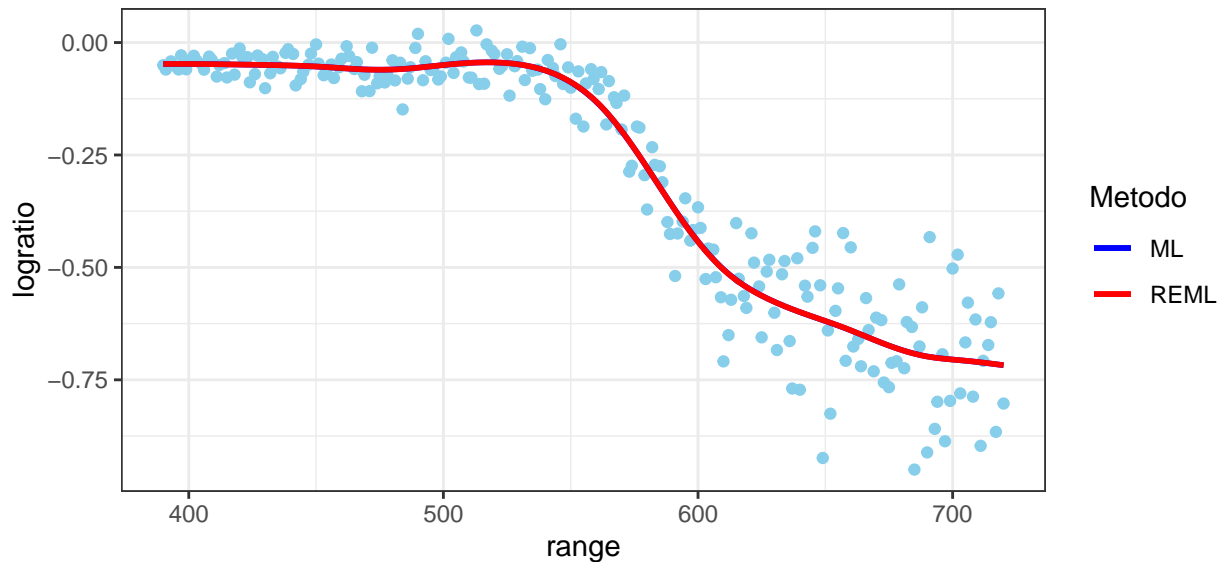
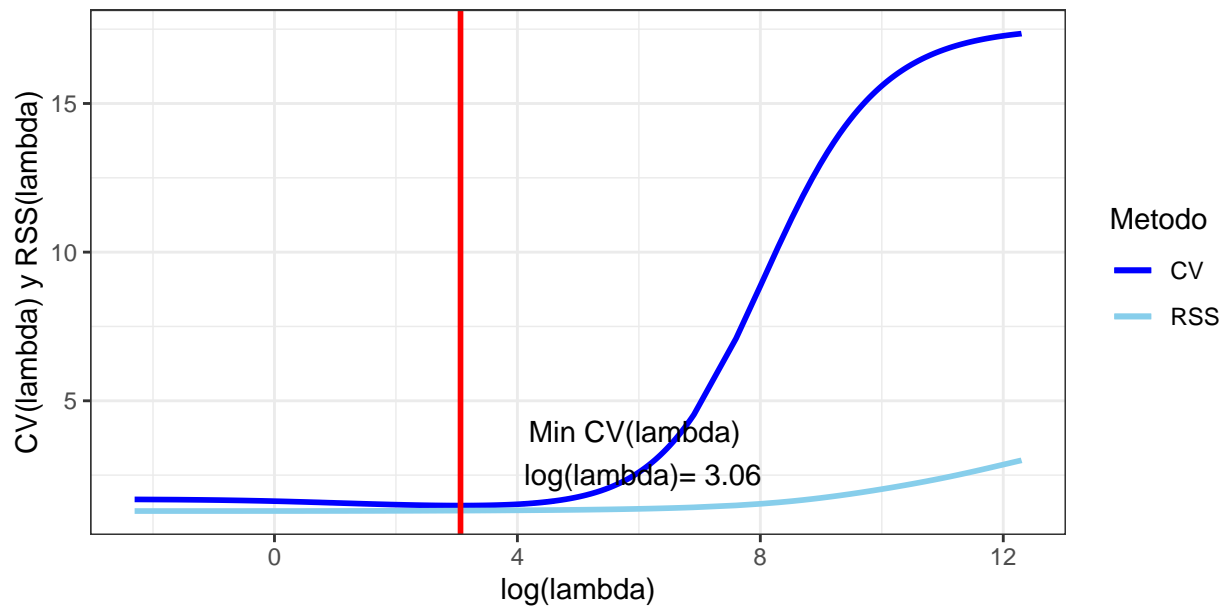


Figure 5.3



Respuesta 6

La **figura 5.9** nos muestra que una vez escogido el λ que penalizara nuestra funciones bases, y que a su vez minimiza el CV, el resultado es parecido independiente de la cantidad de knots que usemos.

Por su parte, la **figura 5.10** nos dice algo similar, pero de igual forma notamos que a medida que crece los knots la regresión se hace mas smooth.

Finalmente, la **figura 5.11** muestra que independiente del orden –el grado al cual se elevara la basis– de la basis escogida, el λ que optimiza es prácticamente el mismo.

Figure 5.9

CV vs df con diferentes knots, y como funcion de lambda

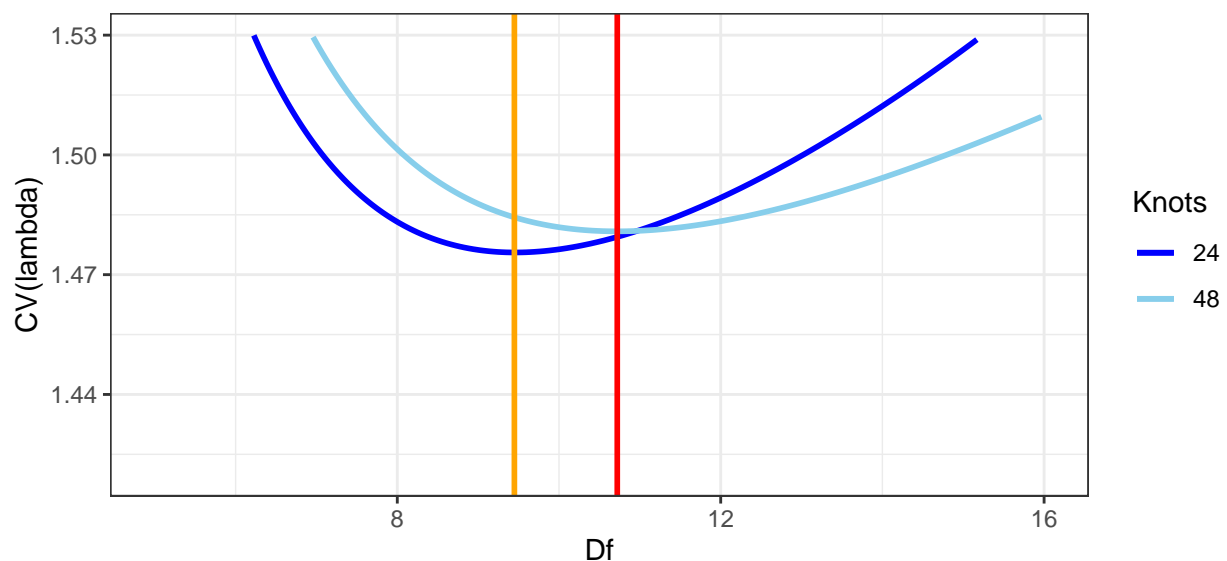


Figure 5.10

Minimo CV lineal con 6,12 y 24 knots.

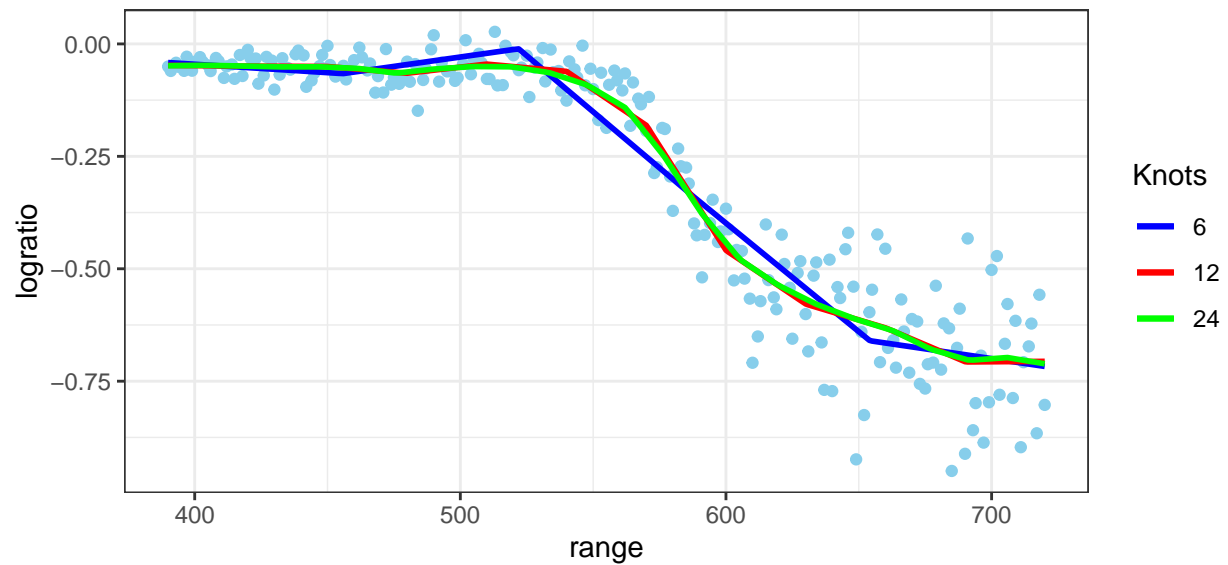
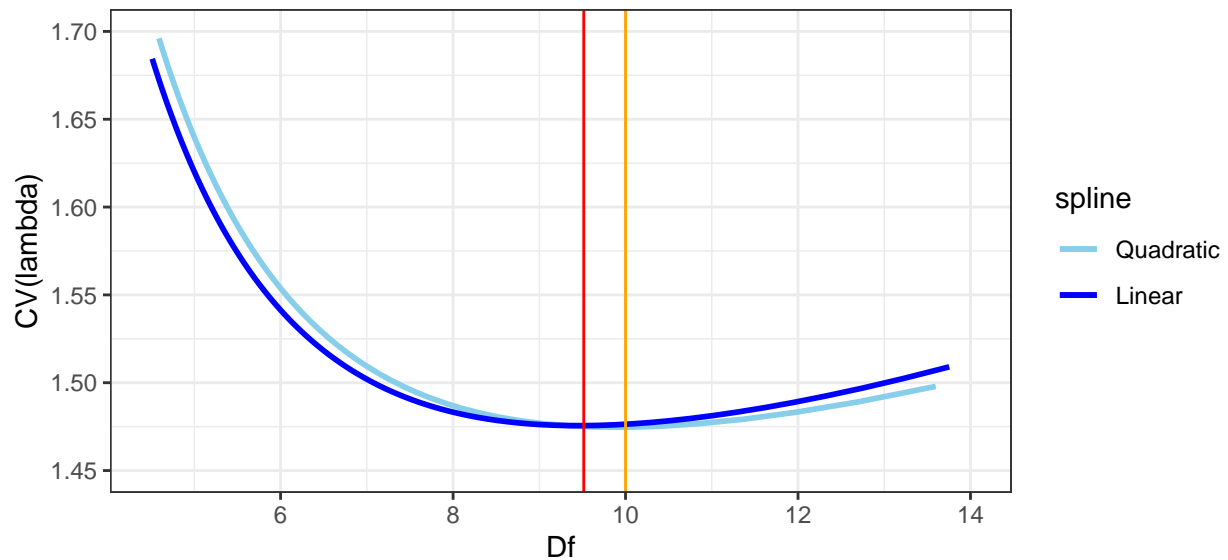


Figure 5.11

Linear vs Quadratic spline

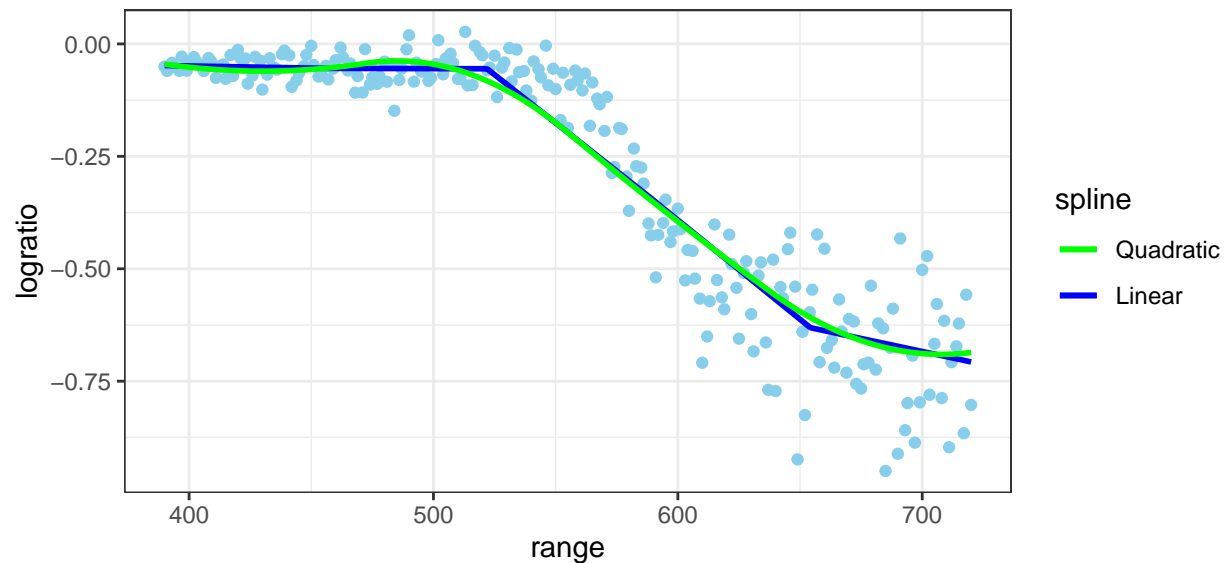


Pregunta 7

La **figura 5.13** hace el énfasis en que a mayor grado de la base mas smooth sera el ajuste. Por lo tanto, una lineal se ve que tiene un ajuste peor que una cuadrática.

Figure 5.13

Min CV 6 knot, linear and quadratic splines



Pregunta 8

Ocuparemos un splines lineal con 6 knots como ejemplo. La **figura 6.1** nos muestra el intervalo de confianza para un subconjunto del total de rangos muestreados, entre el 450 al 500. Se les gráfico su intervalo de confianza con un 95% de confianza.

La **Figura 6.2** muestra el intervalo de confianza para toda la serie, tiene una cobertura relativamente baja.

Por su parte, la **Figura 6.3** incluye los intervalos predictivos los cuales son mas amplios, y por tanto tiene una mejor cobertura. Dado que ahora es la desviación del ajuste mas el del residuo.

Figure 6.1

6 knot linear spline, con IC(95%) entre 450–500

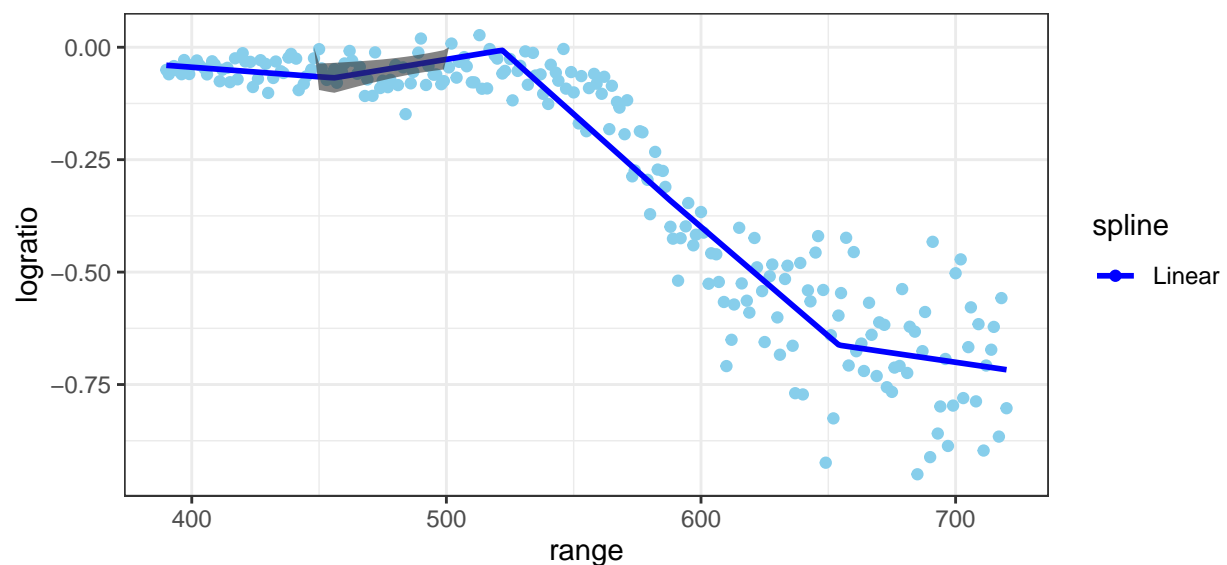


Figure 6.2

6 knot linear spline, con IC(95%)

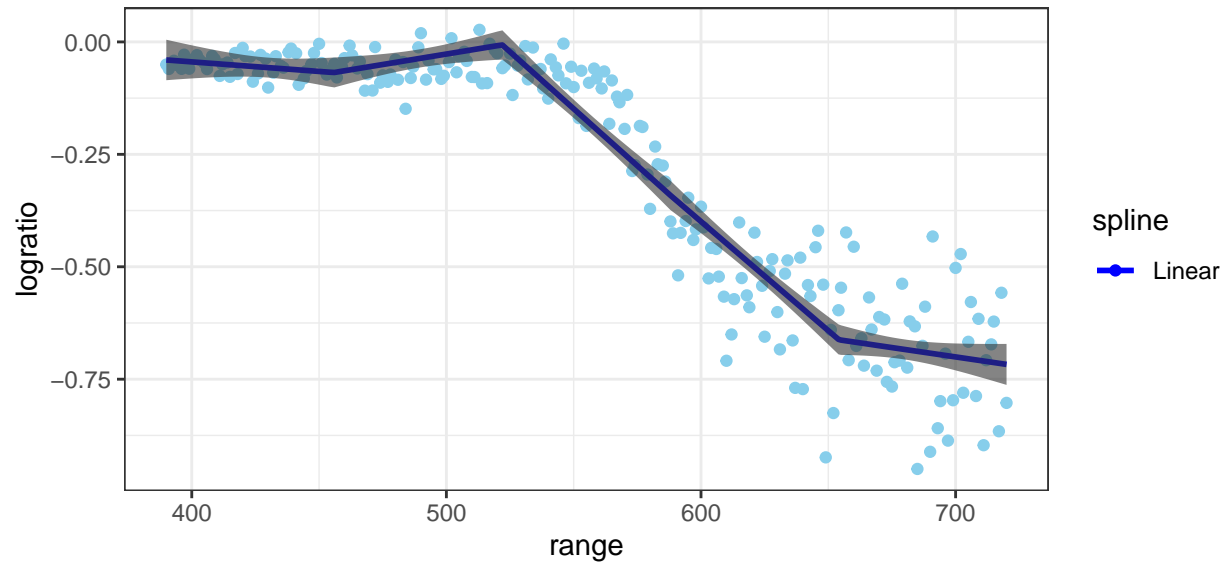


Figure 6.3

6 knot linear spline, con IC(95%) gris oscuro y PI(95%) gris claro

