

COLGATE VS. CREST: ANÁLISIS DE INTERVENCIÓN

Titulación: Máster en Data Science

Autor: Marta Divassón Carribero

Profesor: Ricardo Queralt

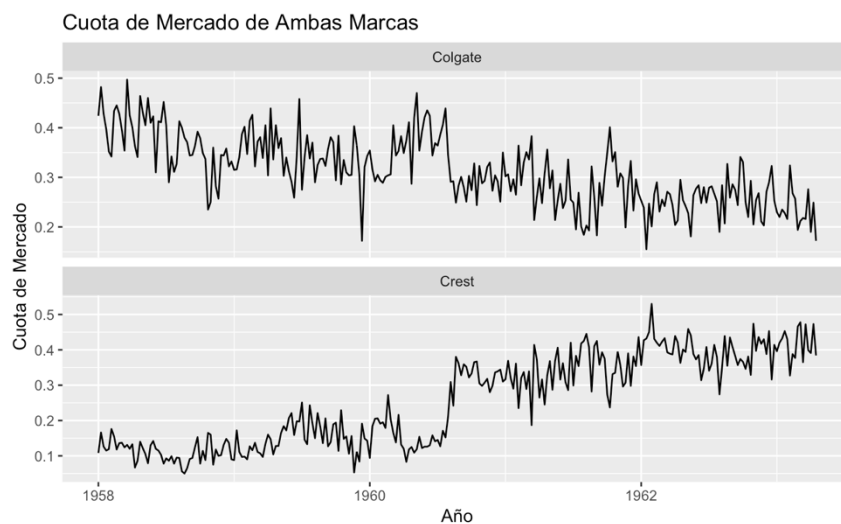
Fecha: enero 2019

Introducción

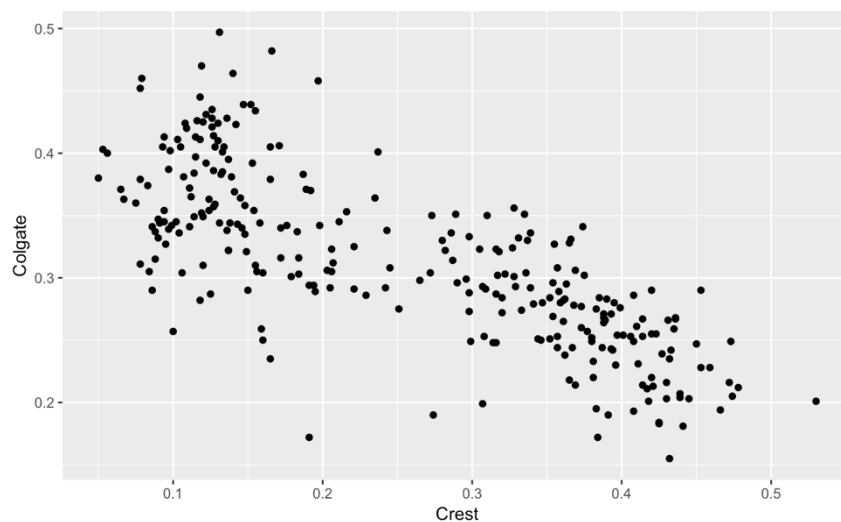
Con este informe se va a desarrollar un modelo de predicción de las cuotas de mercado de las marcas internacionales *Colgate Palmolive* y *Crest* de *Procter & Gamble*. Esto se hará sobre las 16 semanas del año 1963 para encontrar el modelo más preciso que evalué los gastos de publicidad invertidos por *Procter & Gamble* en su producto *Crest*, tras la intervención el 1 de agosto de 1961 de la ADA, que generó una subida en las ventas a mas consumidores. No obstante, también se estimará su dinamismo o permanencia en el tiempo.

Análisis

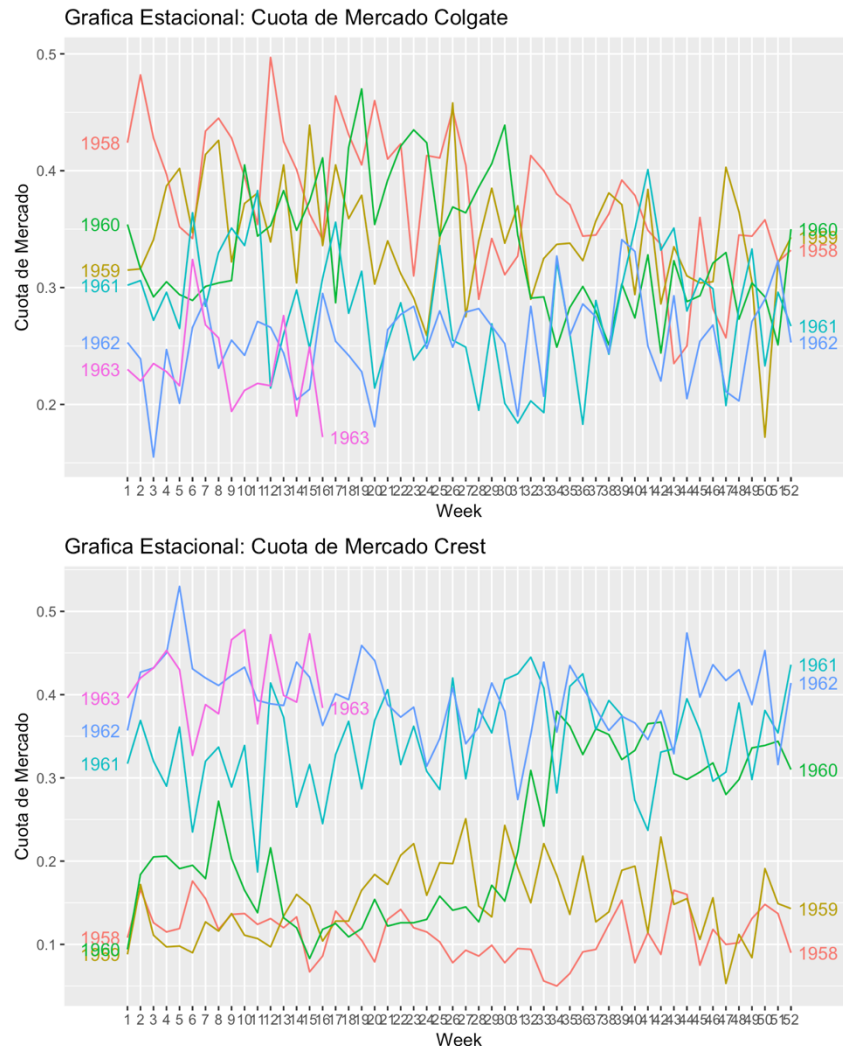
Al graficar las series temporales para ambos productos se pueden ver series *no estacionarias* con una tendencia alcista en el caso de *Crest* y una bajista menos marcada en cuestión de *Colgate*.



En cuanto a la relación entre los dos productos, se representa en el diagrama de dispersión la alta cuota de mercado de *Colgate* en momentos en los que *Crest* no alcanza cifras del 15%, pero como a medida que *Crest* va ganando proporción de los consumidores, *Colgate* disminuye sus ventas.



Es importante recalcar la estacionalidad, ya que, en estas graficas con datos superpuestos, se puede ver con mas claridad el patrón de ventas. En el caso de Colgate, el primer año, en el que la cuota de mercado era mas elevada, está en niveles de 0.45 y disminuye a final de año a niveles de casi 0.35 de la participación en el mercado. Mientras que para Crest, los años en el que hay mayor participación en el mercado son los últimos, llegando a niveles de mas de la mitad del mercado en 1962. Cabe destacar la existencia de autocorrelacion, es decir, como los valores mas cercanos en el tiempo tienen valores cercanos y para conseguir un modelo optimo de predicción se necesita trabajar con datos aleatorios sobre esto y así ajustar un modelo (sobre las 16 semanas de predicción).



Tras esto, se aplican una serie de modelos sencillos, como es el “*Naïve, Mean y Seasonal Naïve*” que se utilizan para evaluar los residuos, que son incluidos en nuestra predicción para que los valores ajustados sean cada vez mas precisos. Uno de los mas ajustados es el método Naïve para Crest con un error (MASE) de 0.4004030.

En el caso de Colgate, los residuos tienen una media constante, excepto por 2 puntos en los que se desvía de la trayectoria. Crest, por otra parte, sigue una media mas o menos constante, excepto para el punto en el año 1961, que va en sentido ascendente (contrario a su bajada en la cuota de mercado de Colgate) y da una pista de ser la posible intervención de la ADA.

Se busca la estacionariedad para que no haya autocorrelacion y así estemos ante lo conocido como “ruido blanco”. Esto se hará a partir de los modelos ARIMA. Miramos el comportamiento de los residuos, aunque tengamos dos modelos ARIMA para Colgate y Crest (**ARIMA (0,1,1)**) y con valores (AIC) muy representativos de que la predicción se está enfocando correctamente. Pero aun hay información que desajusta y por tanto no hay ruido blanco aún (esto implica que la grafica del ACF tenga todos los valores dentro de las bandas azules). Esto se puede deber a los “outliers” que se estudiarán posteriormente con un modelo, es decir, por datos como la intervención de del 1 de agosto de 1961. Cabe mencionar la distribución normal para ambos productos.

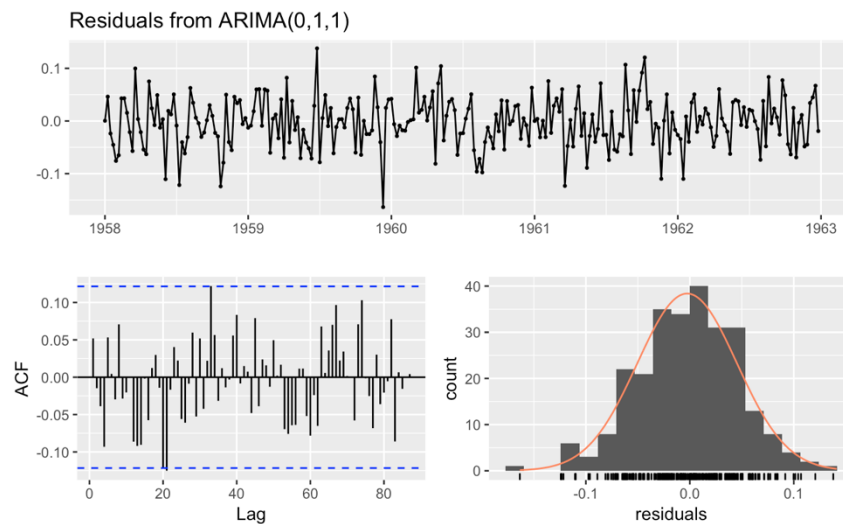


Figura Colgate

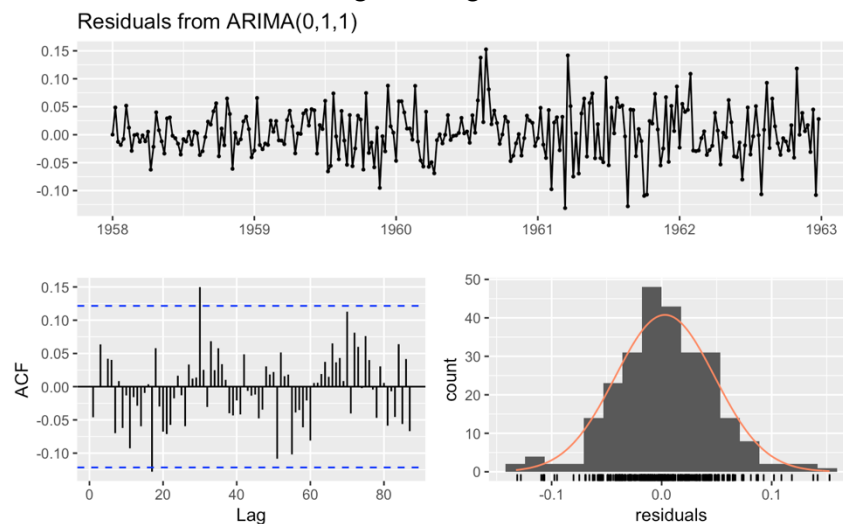
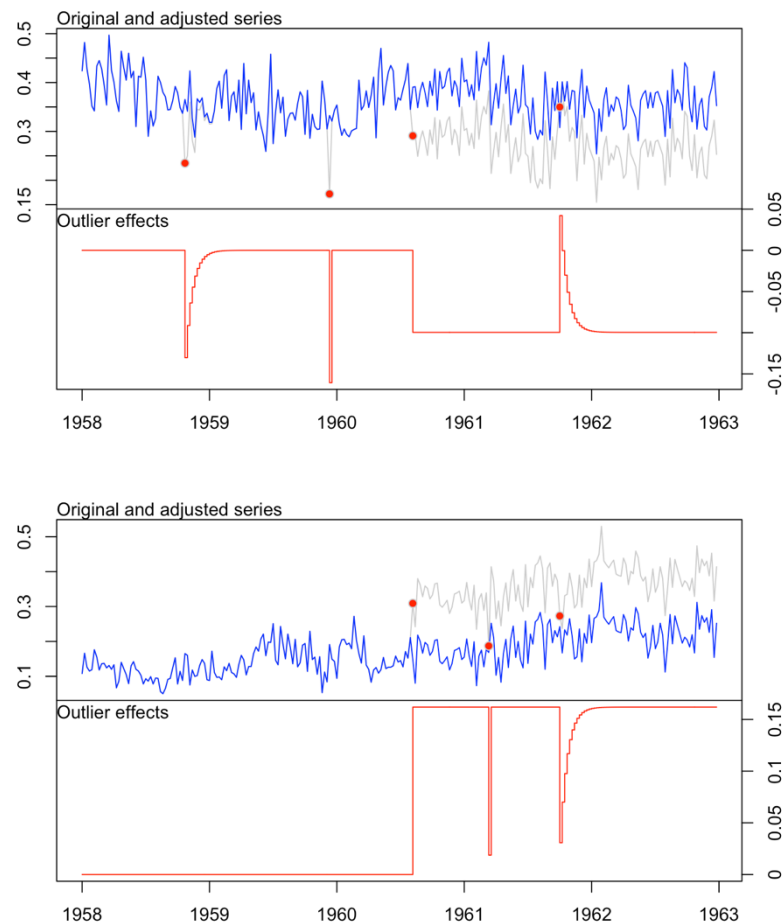


Figura Crest

Comportamiento de Atípicos

En las series temporales hay cambios a veces que cambian la dinámica y son cambios no sistemáticos que no pueden ser capturados por series temporales estándar. Esto se conoce como efectos exógenos o outliers. Estos son importantes en el análisis, ya que el posterior efecto de intervención puede haber sido causado por uno de estos. Bien, hay índices de outliers aditivos en 135, 136 y 138, esto es, se sitúan en las semanas en las que tuvo lugar la intervención de la ADA.

En los siguientes gráficos de Colgate y Crest, se observa respectivamente, los distintos cambios exógenos y en azul se aprecia el modelo ajustado en caso de que no existieran. El efecto de los outliers provoca que la cuota de mercado sea mas baja de lo que podría haber sido si no hubiesen tenido lugar estos efectos.



En el ajuste del modelo ARIMA final con estos comportamientos anteriores, finalmente nos deshacemos de la autocorrelacion con un ARIMA de $(1,0,1)$ en el caso de Colgate y ARIMA $(0,1,1)$ para el de Crest y "ruido blanco" en ambos, con valores de AIC muy indicativos, que muestran mejor comportamiento que el primer ARIMA introducido. Ahora nos encontramos ante series estacionarias.

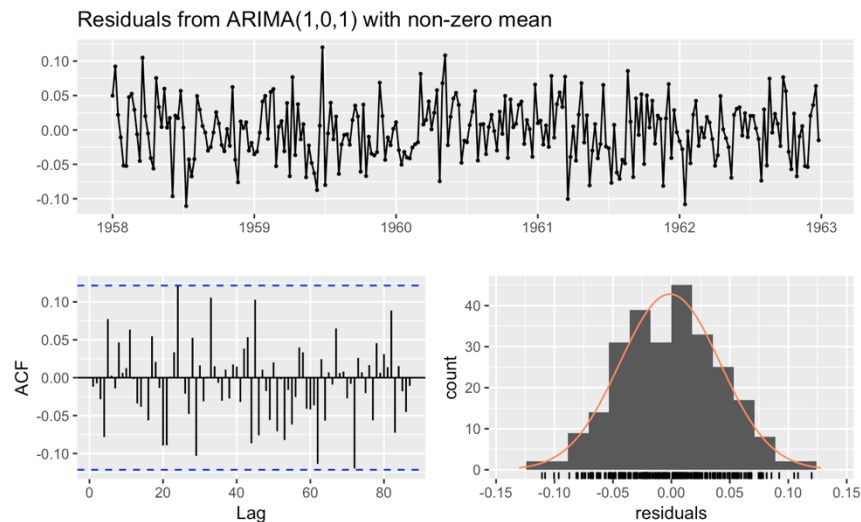


Figura Colgate

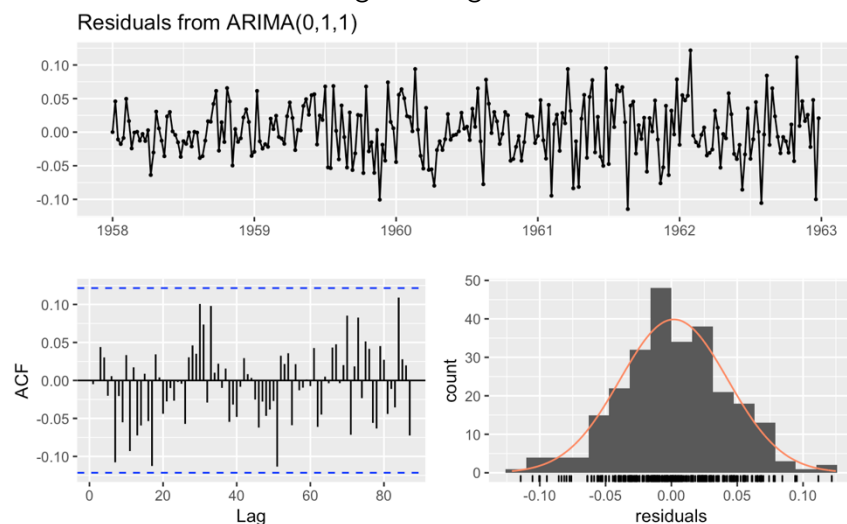
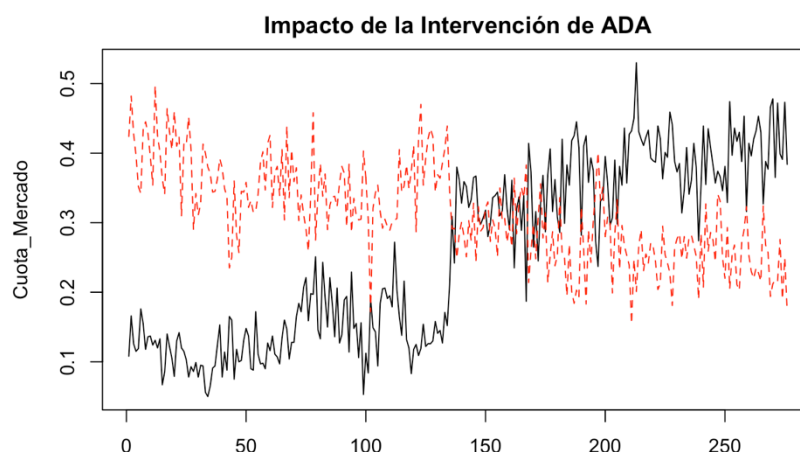


Figura Crest

Tras esto, es importante tener en cuenta el efecto de la publicidad sobre las dos cuotas de mercado de las diferentes marcas. Esto será estudiado mediante un análisis de intervención y la función de transferencia.

Análisis de intervención

Para evaluar el impacto de la intervención de los gastos de publicidad de Procter & Gamble, hay que centrarse en este ajuste del modelo. Esto se dividirá en dos etapas perfectamente segregables visualmente (pre-periodo y post periodo). El objetivo es detectar si con esta intervención se está recibiendo algún tipo de beneficio ya que se analiza como habría evolucionado si no hubiera habido ningún tipo de intervención y se comparan los modelos.



Durante el período posterior a la intervención, la variable de respuesta tuvo un valor promedio de aprox. 0.37. Por el contrario, en ausencia de una intervención, hubiéramos esperado una respuesta promedio de 0.16. El intervalo del 95% de esta predicción contrafactual es [0.15, 0.18]. Al restar esta predicción de la respuesta observada se obtiene una estimación del efecto causal que tuvo la intervención en la variable de respuesta. Este efecto es 0.20 con un intervalo del 95% de [0.19, 0.22].

Al resumir los puntos de datos individuales durante el período posterior a la intervención (que a veces solo se puede interpretar de manera significativa), la variable de respuesta tuvo un valor global de 52.14. Por el contrario, si la intervención no hubiera tenido lugar, habríamos esperado una suma de 23.22. El intervalo del 95% de esta predicción es [20.96, 25.45].

Los resultados anteriores se dan en términos de números absolutos. En términos relativos, la variable de respuesta mostró un incremento de + 125%. El intervalo del 95% de este porcentaje es [+ 115%, + 134%].

Esto significa que el efecto positivo observado durante el período de intervención es *estadísticamente significativo* y es poco probable que se deba a fluctuaciones aleatorias. Sin embargo, debe observarse que la pregunta de si este aumento también tiene un significado sustancial solo puede responderse comparando el efecto absoluto (0.20) con el objetivo original de la intervención subyacente.

La probabilidad de obtener este efecto por azar es muy pequeña (probabilidad Bayesiana unilateral del área de la cola $p = 0.001$). Esto significa que el efecto causal puede considerarse *estadísticamente significativo*.

Esto es en cuestión del efecto del impacto para ambos mercados juntos, visto en 2 gráficos para cada mercado abajo. Si se efectúan por separado sin la influencia del uno sobre el otro directamente, también se observa que el efecto causal es estadísticamente significativo para cada uno, por tanto, la intervención ha sido determinante en la posterior evolución de cada producto. En el caso de Crest ha sido positiva, ya que esta intervención, mas la campaña publicitaria de P&G ha contribuido en su crecimiento.

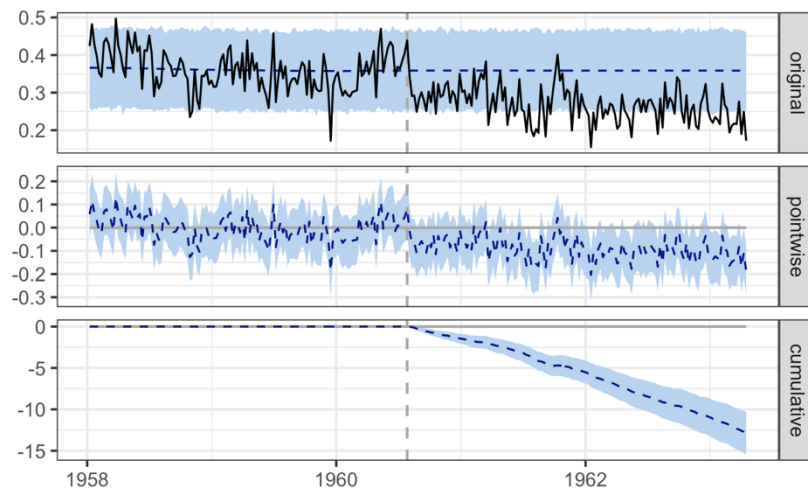


Figura Colgate

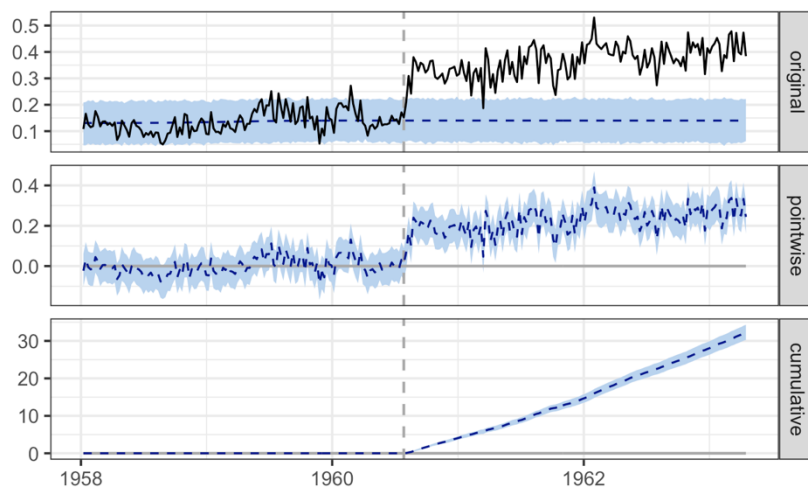


Figura Crest

Función de transferencia

En la ultima parte, tenemos en cuenta el modelo de función de transferencia que tiene efecto sobre la intervención. Hay dos formas de estudiarlo. O bien con la función de impulso o bien con la de salto. Estas dos afectan a la permanencia y al dinamismo ya que, cuando las condiciones cambian con el tiempo, los modelos se pueden ajustar aun mas introduciendo estos inputs que reflejarían los cambios en las condiciones del proceso.

Dicho esto, son como modelos de regresión, pero con respuesta independiente y en los que se valora su termino de error. Los coeficientes de ellos no dependen del tiempo y son estables.

Como se ha dicho, sirve para explicar el dinamismo que hay entre el gasto en publicidad y su impacto en la evolución de las cuotas de mercado. Es decir, se quiere medir si el impacto será dinámico o algo permanente. Cuando hay una intervención el efecto no se plasma en los datos hasta que pasa al menos un mes, es decir, siempre hay un retardo en el efecto.

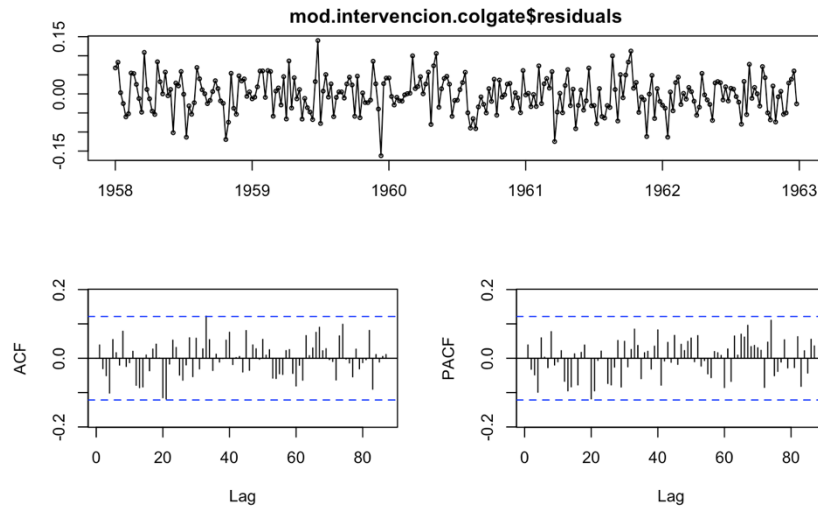


Figura Colgate

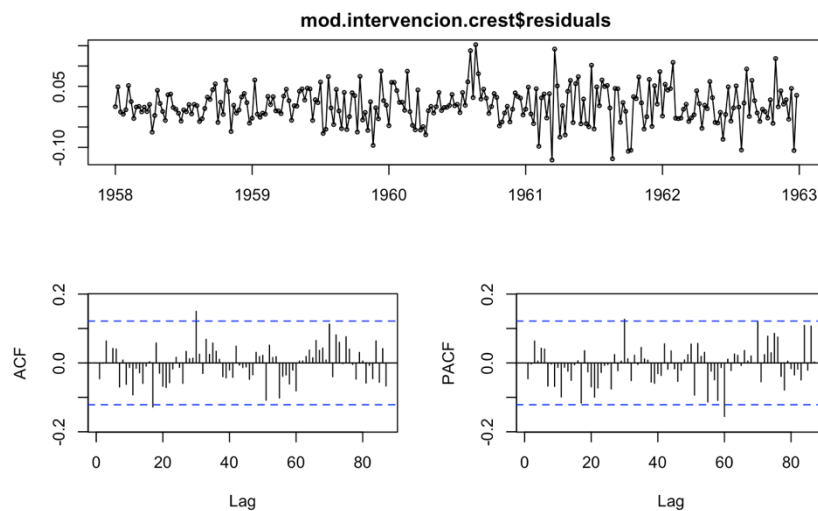


Figura Crest

Conclusión

El objetivo de este informe ha sido encontrar si existe una relación entre ambas cuotas tanto para los valores del pasado como los actuales y así predecir el comportamiento futuro.

Pero el curso normal de la serie temporal se ha interrumpido por las nuevas campañas publicitarias, de forma indirecta por la iniciativa de la ADA que declara a Crest como “marca fundamental para el programa de higiene dental”, como la posterior campaña publicitaria de P&G respaldada por ella.

Con el análisis e intervención (*Causal Impact*) se establece una relación causa y efecto en la que $\hat{\mu}$, en efecto, el paso de la ADA ha tenido un efecto significativo sobre las posteriores cuotas de mercado.

Por una parte, en los diferentes análisis de **Colgate**, se establece que con un solo impulso la intervención afecta (no hay necesidad de tener 2 impulsos), observado en los residuos del modelo tanto a priori como post periodo. Se concluye que el modelo se ajusta a los datos de forma muy ajustada, ya que la grafica de series temporales de los valores ajustados se sobre impone a los valores observados.

En el caso de **Crest**, también hay un ajuste razonablemente bueno, que se debe a que hay algunos residuos cercanos a la intervención, pero sin llegar a ajustarse al igual que el modelo de Colgate. Esto tiene sentido, teniendo en cuenta que este modelo tiene atípicos aditivos en las observaciones 135, 136 y 138.

En un análisis matemático, obtenemos que el efecto sobre Colgate en su descenso de cuota es de un 11.3% aproximadamente, mientras que Crest gana un 17.8% del mercado tras la intervención. Estos números sugieren que Crest no solo se hizo con parte de los consumidores de Colgate, sino que también obtuvo consumidores de otros consumidores menos competentes, en el que los efectos en el mercado no solo han sido momentáneos y dinámicos, sino que se han consolidado como permanentes.