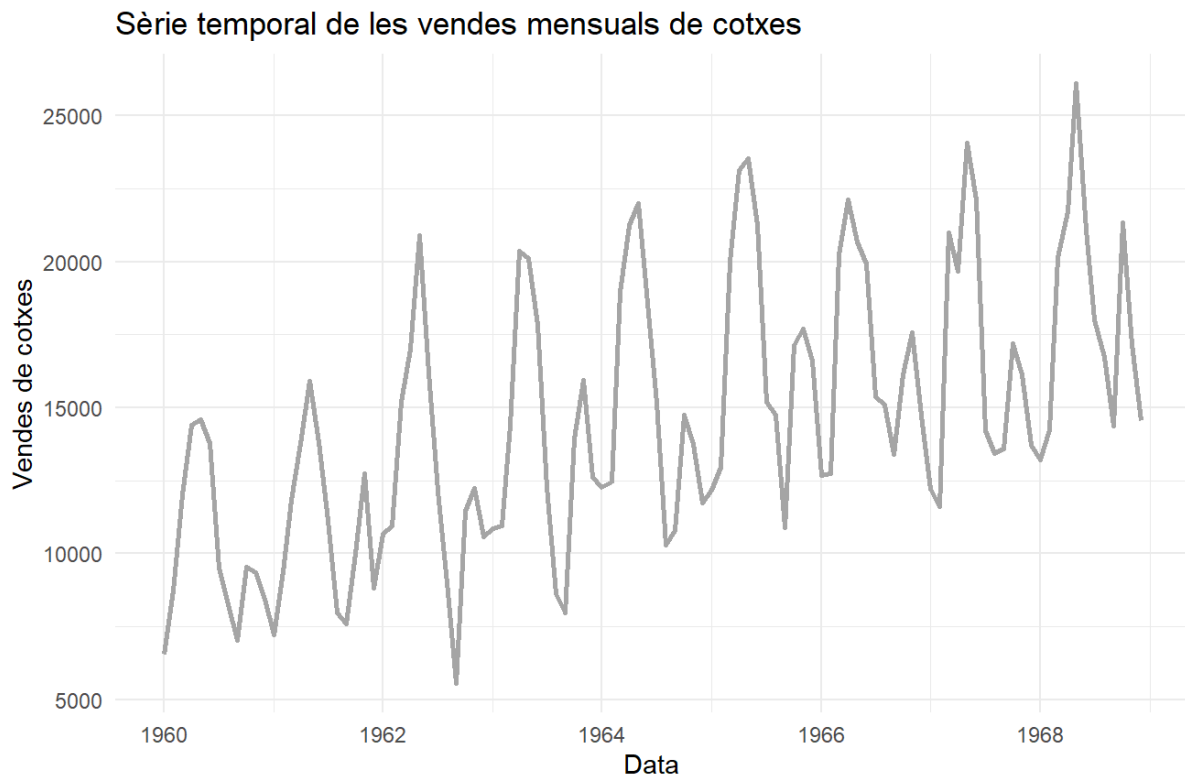


Sèrie temporal:

Aquest és el gràfic inicial de les vendes mensuals de cotxes. S'hi veuen clarament pujades i baixades amb una tendència a l'alça i patrons que es repeteixen cada any (estacionalitat). Aquest comportament ens indica que la mitjana i la variància canvien al llarg del temps, així que abans de fer una anàlisi estadística o prediccions caldrà transformar la sèrie: normalitzar la variància, eliminar la tendència i treure l'efecte estacional. Aquestes passes són bàsiques per poder aplicar models clàssics de sèries temporals i assegurar resultats fiables.



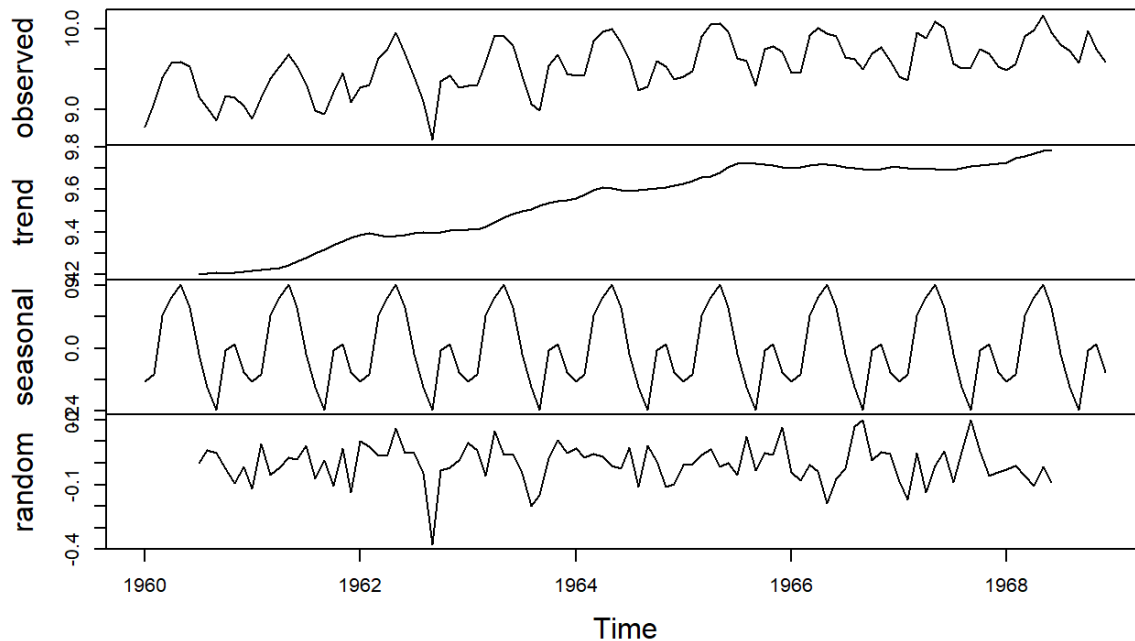
Descomposició:

La descomposició d'una sèrie temporal és fonamental perquè ens permet veure clarament de què està formada la sèrie: la tendència general, el patró estacional que es repeteix cada any i la part aleatòria o soroll.

En aquest gràfic es veu tot: la línia "trend" mostra com la mitjana de vendes creix al llarg dels anys (tendència), la línia "seasonal" reflecteix els cicles anuals molt marcats (estacionalitat) i la línia "random" ens ensenya les fluctuacions més imprevisibles.

Això ens indica que, per poder modelitzar i fer bones previsions amb aquestes dades, caldrà eliminar la tendència i l'estacionalitat perquè la part aleatòria sigui l'única protagonista: només així la sèrie serà estacionària i preparada per aplicar-hi models com ARIMA o predicció clàssica.

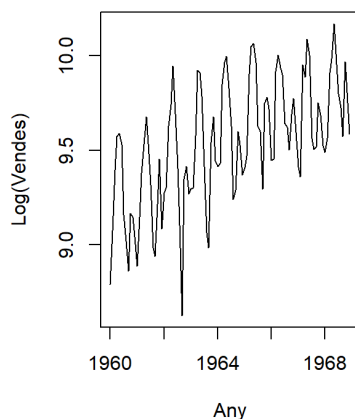
Decomposition of additive time series



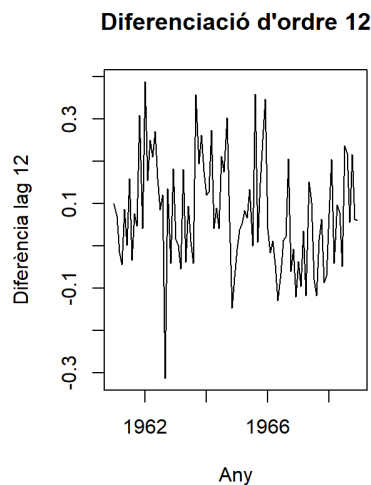
Serie estacionària?

Per determinar si la sèrie de vendes mensuals de cotxes és estacionària, comencem representant gràficament les dades. El primer gràfic mostra la sèrie log-transformada, on s'aprecien clarament un creixement sostingut en el nivell mig (tendència) i cicles estacionals repetits. Això indica, visualment, que la sèrie no és estacionària, ja que la seva mitjana i els patrons de variància canvien amb el temps.

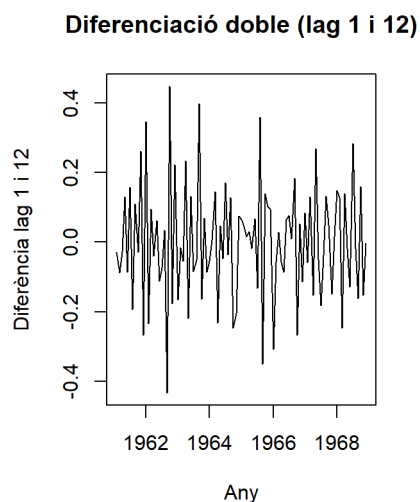
Log(Vendes mensuals de cotxes)



A continuació, utilitzem la diferenciació d'ordre 12, que elimina els patrons estacionals anuals. El segon gràfic mostra que la sèrie resultant ja no presenta estacionalitat visible, però encara pot persistir certa tendència.

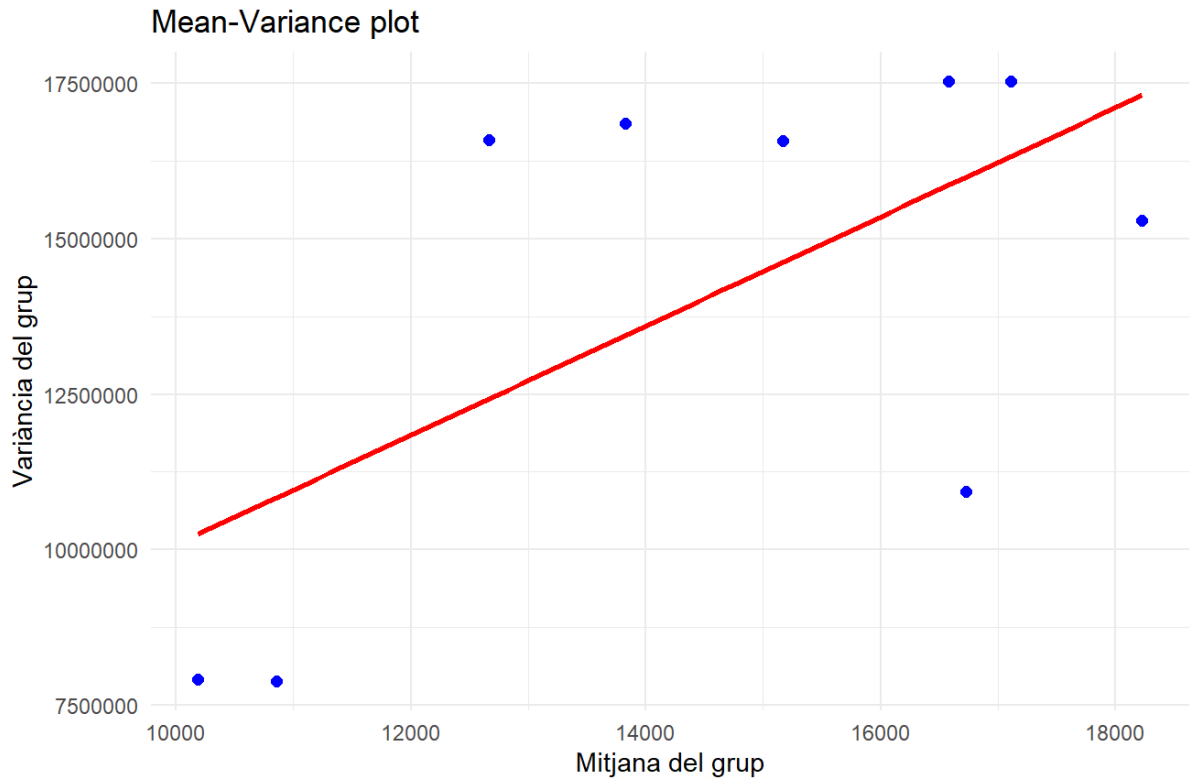


Finalment, fem una diferenciació addicional d'ordre 1 sobre la sèrie diferenciada d'ordre 12. El tercer gràfic resultant mostra una sèrie amb mitjana i variància estables i sense cap tendència o patró clar identificable. Això és indicatiu que la sèrie transformada és ara pràcticament estacionària: ha desaparegut tant la tendència com l'estacionalitat, i les fluctuacions semblen aleatòries al voltant del zero.

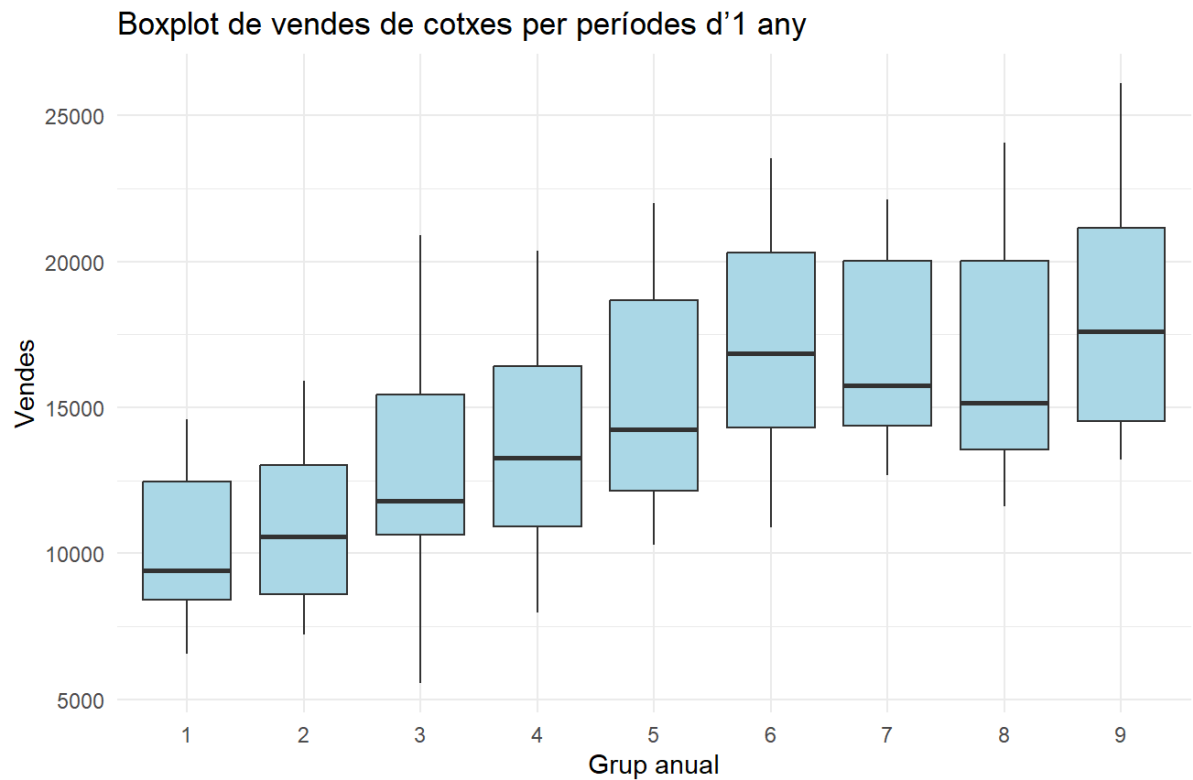


Transformacions: Canvi d'escala

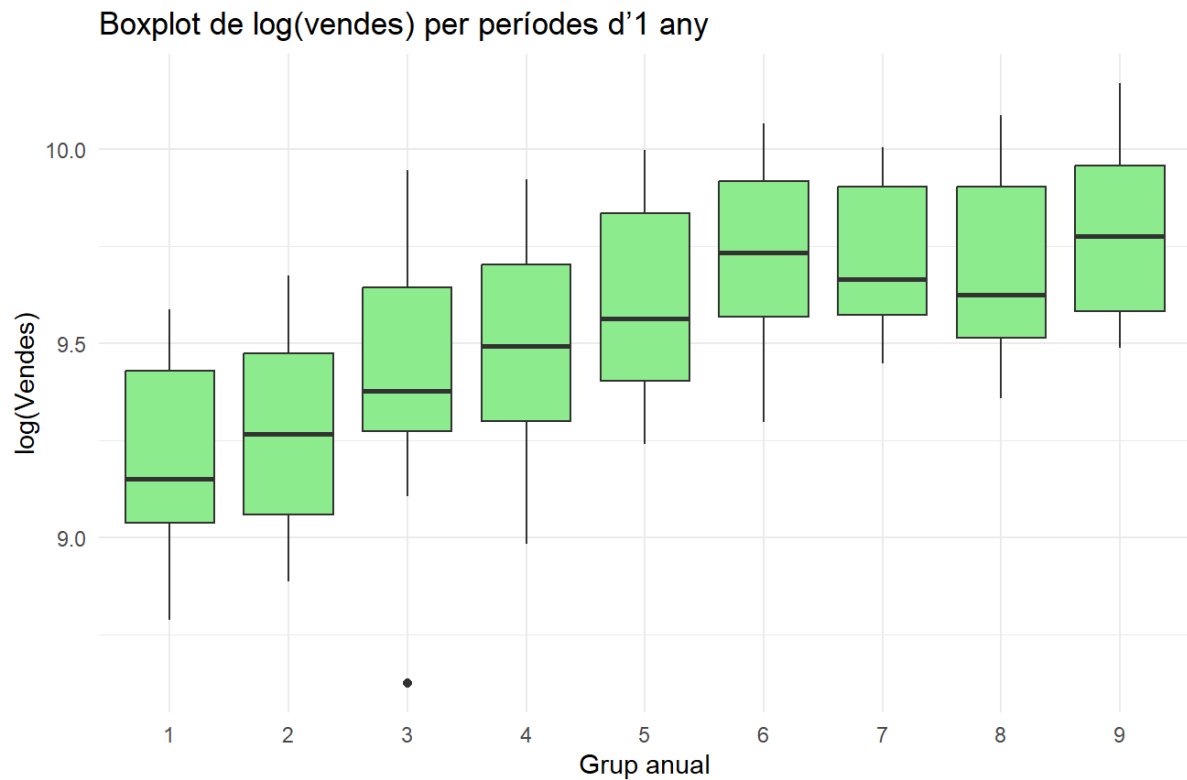
El gràfic Mean-Variance mostra que, a mesura que augmenta la mitjana de vendes dins de cada grup, la variància també creix. Aquesta relació positiva indica clarament que la variància no és constant al llarg del temps, sinó que depèn del nivell de la mitjana de la sèrie. Aquesta situació es coneix com a heteroscedasticitat i no és pròpia d'una sèrie estacionària. Quan s'observa aquest comportament, és aconsellable aplicar transformacions (com el logaritme) per aconseguir estabilitzar la variància abans de continuar analitzant o modelitzant la sèrie.



En dividir la sèrie de vendes mensuals de cotxes en grups d'un any i representar-ne el boxplot per cada període, s'observa que l'alçada de les caixes (IQR) augmenta progressivament a mesura que avança el temps. Aquesta major variabilitat en els grups amb valors mitjans més alts indica que la variància no és constant al llarg del temps. Per això, és necessari aplicar una transformació d'escala (normalment el logaritme) abans de treballar models estacionaris, ja que només així podem estabilitzar la variància i complir els requisits dels models clàssics de sèries temporals.

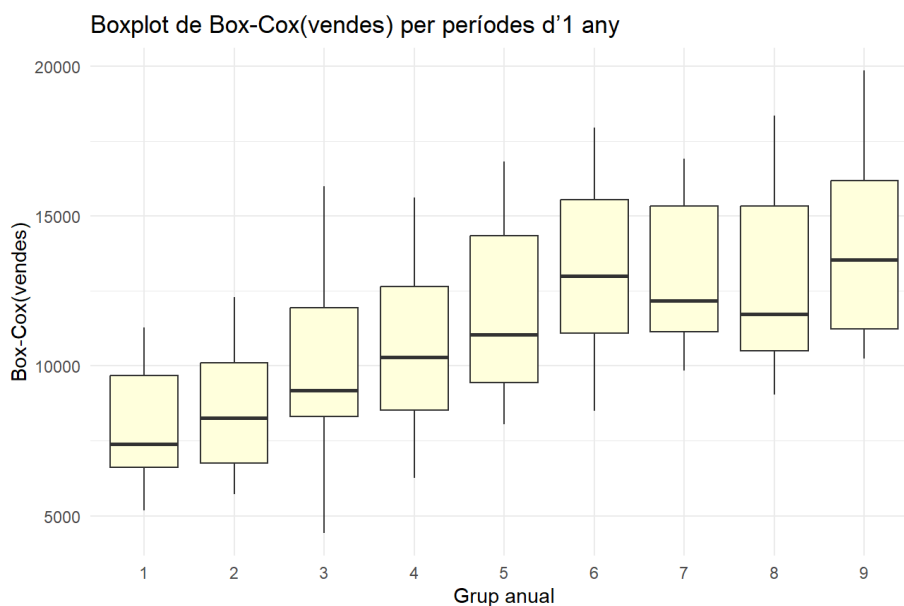


Amb la gràfica del boxplot de $\log(\text{vendes})$ per períodes d'un any, es pot concloure que la variància s'ha estabilitzat: les caixes (IQR) són molt més uniformes entre els diferents grups anuals, i la diferència d'alçada respecte als primers anys pràcticament ha desaparegut. Això vol dir que la transformació logarítmica ha estat efectiva per aconseguir variància aproximadament constant a la sèrie, fet que permet aplicar mètodes d'anàlisi estacionària amb més garanties de validesa estadística.



En el cas de la nostra base de dades de vendes mensuals de cotxes, la variància augmenta amb la mitjana: això es veu clarament en els boxplots anuals, on les caixes són més altes a mesura que les vendes creixen. Si només fem el logaritme, podem estabilitzar bastant la variància, però la transformació Box-Cox és millor perquè busca automàticament quin tipus de transformació (potència o logaritme) s'adapta millor als nostres valors.

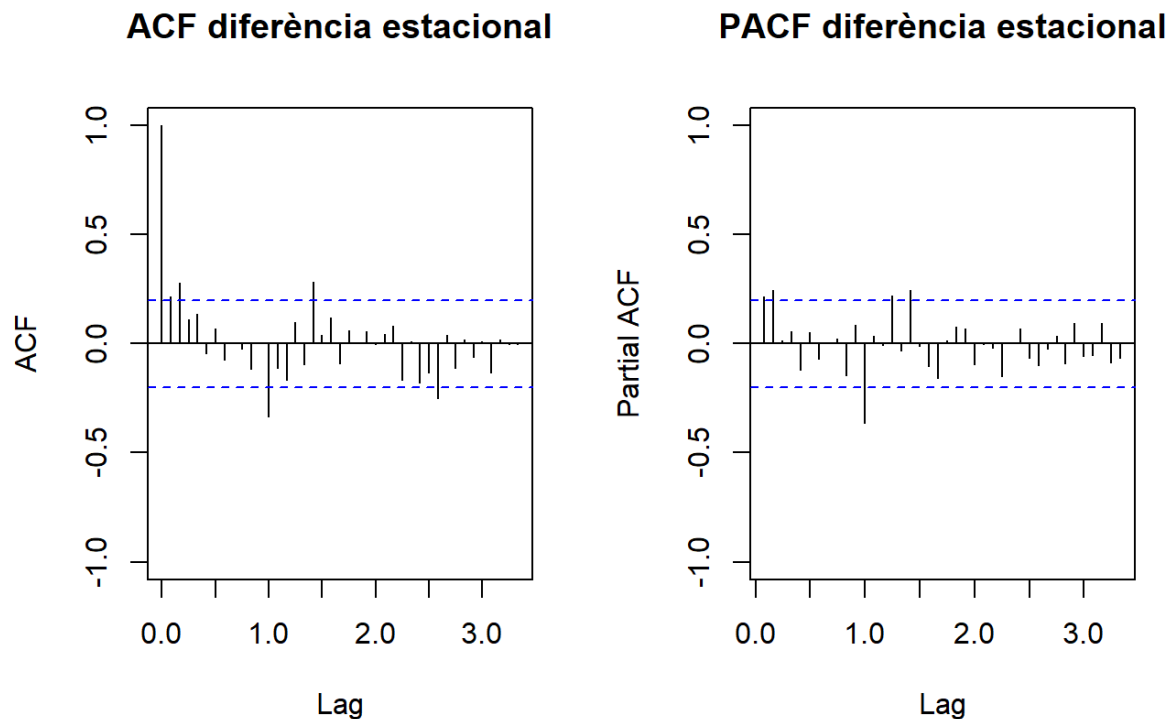
Això fa que la variància sigui encara més estable i la sèrie sigui més fàcil d'analitzar o de modelitzar després. Així, per les nostres dades, Box-Cox és la millor opció per aconseguir una sèrie més homogènia i adequada per a l'anàlisi de sèries temporals.



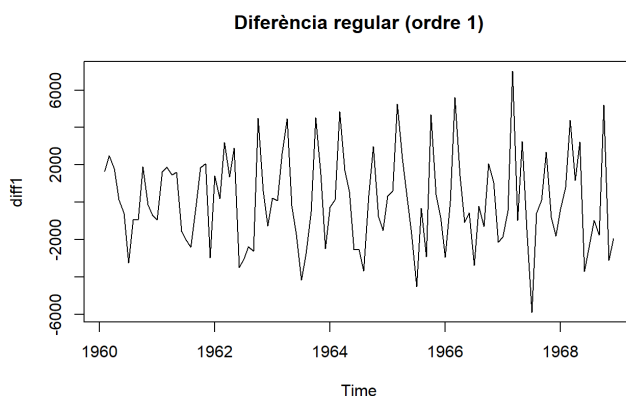
Transformacions: Diferència estacional

Després d'aplicar la diferenciació estacional d'ordre 12 i analitzar les funcions d'autocorrelació (ACF) i autocorrelació parcial (PACF), podem veure que gairebé totes les barres es situen dins dels límits de significació. Això indica que el patró estacional anual s'ha eliminat i la sèrie ja no presenta les dependències regulars que tenia inicialment.

Aquesta transformació ha convertit la sèrie de vendes mensuals de cotxes en una sèrie molt més homogènia i pròxima a l'estacionarietat, la qual cosa la fa més adequada per a l'anàlisi estadística i la modelització amb models ARIMA i similars.

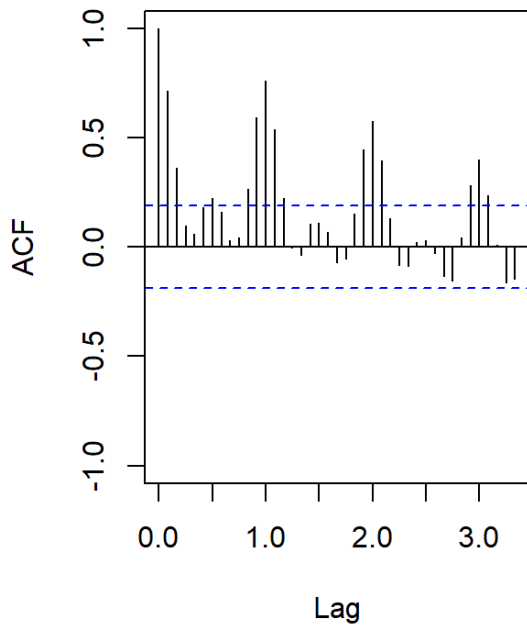


Després d'aplicar la diferenciació regular d'ordre 1, s'obté una sèrie que ja no mostra una tendència clara al llarg del temps. Ara, les variacions s'acumulen de forma irregular al voltant del zero i la mitjana de la sèrie es manté estable. Això vol dir que la diferenciació ha eliminat l'efecte de la tendència i la sèrie està més a prop de ser estacionària: és a dir, té una mitjana constant i pot ser utilitzada per a models com ARIMA on aquesta propietat és fonamental.

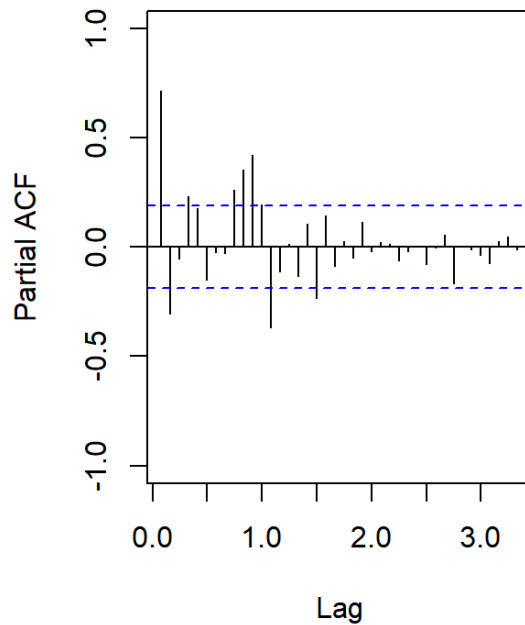


Després d'analitzar els gràfics d'ACF i PACF, es pot concloure que la sèrie de vendes mensuals, un cop diferenciada correctament, mostra un comportament pròxim a un AR(1): la PACF té un pic important al primer lag i després decau ràpidament, mentre que la ACF decau suaument i no talla en sec. Això indica que la dependència principal de la sèrie respecte a valors anteriors és autoregressiva d'ordre 1, i no de tipus mitjana mòbil. Aquest resultat serveix de base per escollir l'estructura del model ARIMA més adequat.

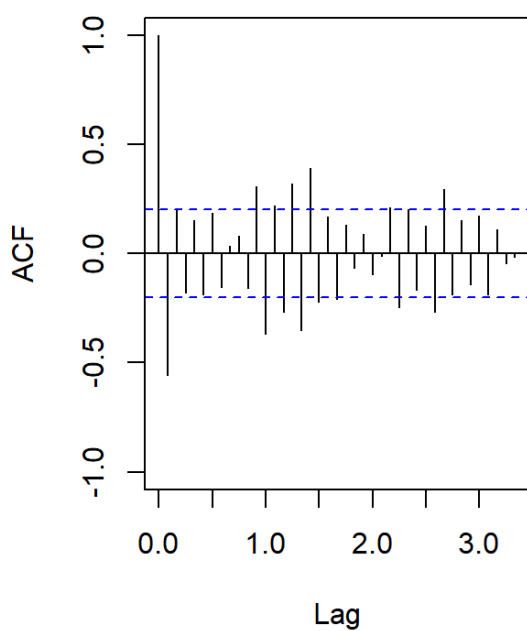
ACF mostra vendes cotxes



PACF mostra vendes cotxes



ACF dif (1,12)



PACF dif (1,12)

