CAPM aplicado.

Beta, costo de capital y alpha.

Dr. Martín Lozano https://mlozanoqf.github.io/

21 de February de 2026, 11:53 PM

	Fundamental	Intermedio	Especializado
Finanzas	×	×	✓
Estadística	×	✓	×
R	×	✓	×

Introducción.

- El CAPM se presenta como una herramienta operativa para decisiones financieras. El propósito no es contrastar el CAPM como teoría de equilibrio, sino obtener insumos cuantitativos con interpretación clara para análisis aplicado.
- Usamos rendimientos en exceso definidos como $r_{i,t} = R_{i,t} - R_{f,t}$ para el activo o fondo y $r_{M,t} = R_{M,t} - R_{f,t}$ para el mercado. Esta notación coincide con el video académico y mantiene consistencia con la teoría financiera.
- El benchmark de mercado se aproxima con el índice S&P 500, usando su serie de niveles (`^GSPC`) para construir rendimientos mensuales. En consecuencia, el rendimiento del mercado corresponde a un índice de precio. Esta elección se adopta por disponibilidad y transparencia, y se mantiene fija a lo largo del video.
- Se presentan tres aplicaciones basadas en la misma regresión de serie de tiempo en exceso,
$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t},$$
con objetivos distintos: (1) estimar $\widehat{\beta}$ para cuantificar exposición al mercado; (2) derivar un rendimiento requerido del equity $k_e = R_f + \widehat{\beta} \cdot ERP$; y (3) evaluar desempeño de un fondo con el alpha de Jensen $\widehat{\alpha}$.
- A diferencia del video académico, aquí no se realizan estimaciones en sección cruzada ni Fama–MacBeth, ya que esas técnicas se orientan a estimar el precio del riesgo y evaluar predicciones del modelo en un conjunto de activos. En este video el énfasis es el uso práctico de $\widehat{\beta}$, k_e y $\widehat{\alpha}$ como insumos para decisión, con supuestos explícitos.
- La muestra es mensual y utiliza una ventana de 120 observaciones hasta diciembre de 2025. Se enfatiza que β y α dependen de ventana, frecuencia y benchmark, por lo que su interpretación debe entenderse como condicional al diseño empírico adoptado.

Paquetes.

```r
library(dplyr)
library(tidyr)
library(tibble)
library(lubridate)
library(tidyquant)
library(broom)
library(lmtest)
library(sandwich)
library(readr)
library(ggplot2)
```

Descripción de datos.

- Descargamos precios desde Yahoo Finance con `tidyquant` para el índice S&P 500 (`^GSPC`) y para los instrumentos de las tres aplicaciones.
- Convertimos precios a rendimientos mensuales usando precios ajustados y rendimientos aritméticos.
- Para la tasa libre de riesgo usamos una serie mensual de `FRED` (`TB3MS`), que es una tasa anualizada; la transformamos a rendimiento mensual efectivo para construir $R_{f,t}$ en la misma frecuencia.
- La ventana se fija para obtener 120 rendimientos mensuales hasta diciembre de 2025.

Datos 1/3.

```r
# ----------------------------
# Parámetros de muestra
# ----------------------------
end_date   <- as.Date("2025-12-31")
start_date <- end_date %m-% months(120) %m-% months(3)

# Instrumentos
ticker_mkt    <- "^GSPC"
tickers_asset <- c("AAPL","KO","FDGRX")

# ----------------------------
# Helper robusto: asegurar clase Date
# ----------------------------
fix_date <- function(x) {
  if (inherits(x, "Date")) return(x)
  if (inherits(x, "POSIXct") || inherits(x, "POSIXt")) return(as.Date(x))
  if (is.numeric(x)) return(as.Date(x, origin = "1970-01-01"))
  as.Date(as.character(x))
}

# ----------------------------
# Rendimientos mensuales desde Yahoo (tidyquant)
# Fecha alineada al 1er día del mes
# ----------------------------
get_monthly_returns_yahoo <- function(tickers, start, end) {

  px <- tryCatch(
    tq_get(tickers, get = "stock.prices", from = start, to = end),
    error = function(e) NULL
  )

  if (is.null(px) || is.logical(px) || !is.data.frame(px) || nrow(px) == 0) {
    stop("No se pudo descargar precios desde Yahoo Finance con tq_get().", call. = FALSE)
  }

  out <- px |>
    group_by(symbol) |>
    tq_transmute(
      select     = adjusted,
      mutate_fun = periodReturn,
      period     = "monthly",
      type       = "arithmetic",
      col_rename = "ret"
    ) |>
    ungroup() |>
    mutate(
      date = as.Date(floor_date(fix_date(date), "month")),
      ret  = 100 * ret
    ) |>
    rename(ticker = symbol) |>
    distinct(ticker, date, .keep_all = TRUE) |>
    arrange(ticker, date)

  out
}
```

```r

# ----------------------------
# RF mensual (en %) con fallbacks, sin quantmod
# 1) FRED CSV (TB3MS) robusto a nombres
# 2) Yahoo ^IRX como proxy si FRED falla
# ----------------------------
get_rf_monthly <- function(start, end) {

  fred_url <- "https://fred.stlouisfed.org/graph/fredgraph.csv?id=TB3MS"

  rf1 <- tryCatch({
    raw <- readr::read_csv(fred_url, show_col_types = FALSE)
    if (!is.data.frame(raw) || nrow(raw) == 0) stop("FRED vacío")

    nms <- names(raw)

    date_col <- dplyr::case_when(
      "DATE" %in% nms ~ "DATE",
      "observation_date" %in% nms ~ "observation_date",
      "date" %in% nms ~ "date",
      TRUE ~ nms[1]
    )

    val_candidates <- c("TB3MS", "value", "VALUE")
    val_col <- val_candidates[val_candidates %in% nms][1]
    if (is.na(val_col)) val_col <- setdiff(nms, date_col)[1]

    raw |>
      transmute(
        date_raw = .data[[date_col]],
        tb3ms    = suppressWarnings(as.numeric(.data[[val_col]]))
      ) |>
      mutate(date = fix_date(date_raw)) |>
      filter(!is.na(date), !is.na(tb3ms), date >= start, date <= end) |>
      mutate(date = as.Date(floor_date(date, "month"))) |>
      group_by(date) |>
      summarise(tb3ms = last(tb3ms), .groups = "drop") |>
      transmute(
        date = date,
        rf   = 100 * ((1 + (tb3ms/100))^(1/12) - 1)
      )
  }, error = function(e) NULL)

  if (!is.null(rf1) && nrow(rf1) > 0) return(rf1)

  rf2 <- tryCatch({
    tq_get("^IRX", get = "stock.prices", from = start, to = end) |>
      transmute(
        date = as.Date(floor_date(fix_date(date), "month")),
        irx  = adjusted
      ) |>
      group_by(date) |>
      summarise(irx = last(irx), .groups = "drop") |>
      filter(!is.na(irx)) |>
      transmute(
        date = date,
        rf   = 100 * ((1 + (irx/100))^(1/12) - 1)
      )
  }, error = function(e) NULL)
```

```r

  if (!is.null(rf2) && nrow(rf2) > 0) {
    warning("RF: no se pudo obtener TB3MS desde FRED; usando ^IRX como proxy.")
    return(rf2)
  }

  stop("RF: falló FRED (TB3MS) y falló Yahoo (^IRX).", call. = FALSE)
}

# ----------------------------
# Descargas
# ----------------------------
mkt_mret <- get_monthly_returns_yahoo(ticker_mkt, start_date, end_date) |>
  transmute(date = fix_date(date), mkt_ret = ret) |>
  distinct(date, .keep_all = TRUE) |>
  arrange(date)

assets_mret <- get_monthly_returns_yahoo(tickers_asset, start_date, end_date) |>
  mutate(date = fix_date(date))

rf_m <- get_rf_monthly(start_date, end_date) |>
  mutate(date = fix_date(date))

# ----------------------------
# Normalización final: asegurar Date y calendario común real
# ----------------------------
mkt_mret2 <- mkt_mret |>
  mutate(date = as.Date(floor_date(fix_date(date), "month"))) |>
  distinct(date, .keep_all = TRUE) |>
  arrange(date)

rf_m2 <- rf_m |>
  mutate(date = as.Date(floor_date(fix_date(date), "month"))) |>
  group_by(date) |>
  summarise(rf = last(rf), .groups = "drop") |>
  filter(!is.na(rf)) |>
  arrange(date)

assets_mret2 <- assets_mret |>
  mutate(date = as.Date(floor_date(fix_date(date), "month"))) |>
  distinct(ticker, date, .keep_all = TRUE) |>
  arrange(ticker, date)

# Mercado + RF (meses donde existen ambos)
mkt_rf <- inner_join(mkt_mret2, rf_m2, by = "date") |>
  arrange(date)

# Calendario común (forzamos Date después de intersect por seguridad)
common_dates <- intersect(unique(assets_mret2$date), unique(mkt_rf$date))
common_dates <- fix_date(common_dates)
common_dates <- sort(common_dates)
```

Verificación de datos 1/2.

```r
# Diagnóstico: aquí common_dates debe imprimirse como fechas, no números
range(mkt_rf$date); nrow(mkt_rf)
```

```
## [1] "2015-10-01" "2025-12-01"

## [1] 123
```

```r
range(common_dates); length(common_dates)
```

```
## [1] "2015-10-01" "2025-12-01"

## [1] 123
```

```r
tail(common_dates, 5)
```

```
## [1] "2025-08-01" "2025-09-01" "2025-10-01" "2025-11-01" "2025-12-01"
```

```r
# Últimos 120 meses
last_120 <- tail(common_dates, 120)

# Dataset maestro
capm_practical_data <- assets_mret2 |>
  filter(date %in% last_120) |>
  inner_join(mkt_rf |> filter(date %in% last_120), by = "date") |>
  mutate(
    excess_ret = ret     - rf,
    mkt_excess = mkt_ret - rf
  ) |>
  select(date, ticker, ret, rf, mkt_ret, excess_ret, mkt_excess) |>
  arrange(ticker, date)
```

Verificación de datos 2/2.

```r
# Checks finales: debe dar 120 meses por ticker
capm_practical_data |>
  group_by(ticker) |>
  summarise(n_months = n_distinct(date),
            start    = min(date),
            end      = max(date),
            .groups = "drop")
```

```
## # A tibble: 3 x 4
##   ticker n_months start      end
##   <chr>     <int> <date>     <date>
## 1 AAPL        120 2016-01-01 2025-12-01
## 2 FDGRX       120 2016-01-01 2025-12-01
## 3 KO          120 2016-01-01 2025-12-01
```

```r
tail(capm_practical_data)
```

```
## # A tibble: 6 x 7
##   date       ticker   ret    rf mkt_ret excess_ret mkt_excess
##   <date>     <chr>  <dbl> <dbl>   <dbl>      <dbl>      <dbl>
## 1 2025-07-01 KO     -4.04 0.347   2.17       -4.39      1.82
## 2 2025-08-01 KO      1.62 0.337   1.91        1.28      1.57
## 3 2025-09-01 KO     -3.13 0.321   3.53       -3.45      3.21
## 4 2025-10-01 KO      3.89 0.313   2.27        3.58      1.96
## 5 2025-11-01 KO      6.12 0.310   0.130       5.82     -0.180
## 6 2025-12-01 KO     -3.50 0.294   0.688      -3.79      0.394
```

Aplicación 1. Contexto y objetivo: beta de `AAPL`.

- Objetivo: cuantificar la exposición de `AAPL` al riesgo de mercado usando el índice S&P 500 como benchmark.
- Trabajamos con rendimientos en exceso: $r_{i,t} = R_{i,t} - R_{f,t}$ y $r_{M,t} = R_{M,t} - R_{f,t}$, donde $R_{M,t}$ proviene de `^GSPC`.
- Estimamos la regresión de serie de tiempo en exceso:

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t},$$

  y reportamos $\widehat{\beta}$ como medida de sensibilidad del activo a variaciones del mercado netas de la tasa libre de riesgo.
- Interpretación operativa: $\widehat{\beta} > 1$ implica sensibilidad superior a la del mercado; $\widehat{\beta} < 1$ implica sensibilidad inferior. El parámetro es condicional a ventana, frecuencia y benchmark.

Aplicación 1. Estimación: beta de AAPL.

```r
# Estimación de $\beta$ para AAPL (OLS y opcional HAC).
# Filtrar AAPL
aapl_df <- capm_practical_data |>
  filter(ticker == "AAPL") |>
  select(date, excess_ret, mkt_excess) |>
  drop_na()

# Regresión CAPM en exceso
fit_aapl <- lm(excess_ret ~ mkt_excess, data = aapl_df)

# Resultados OLS
aapl_ols <- tidy(fit_aapl) |>
  mutate(term = recode(term, `(Intercept)` = "alpha", mkt_excess = "beta")) |>
  select(term, estimate, std.error, statistic, p.value)

aapl_ols
```

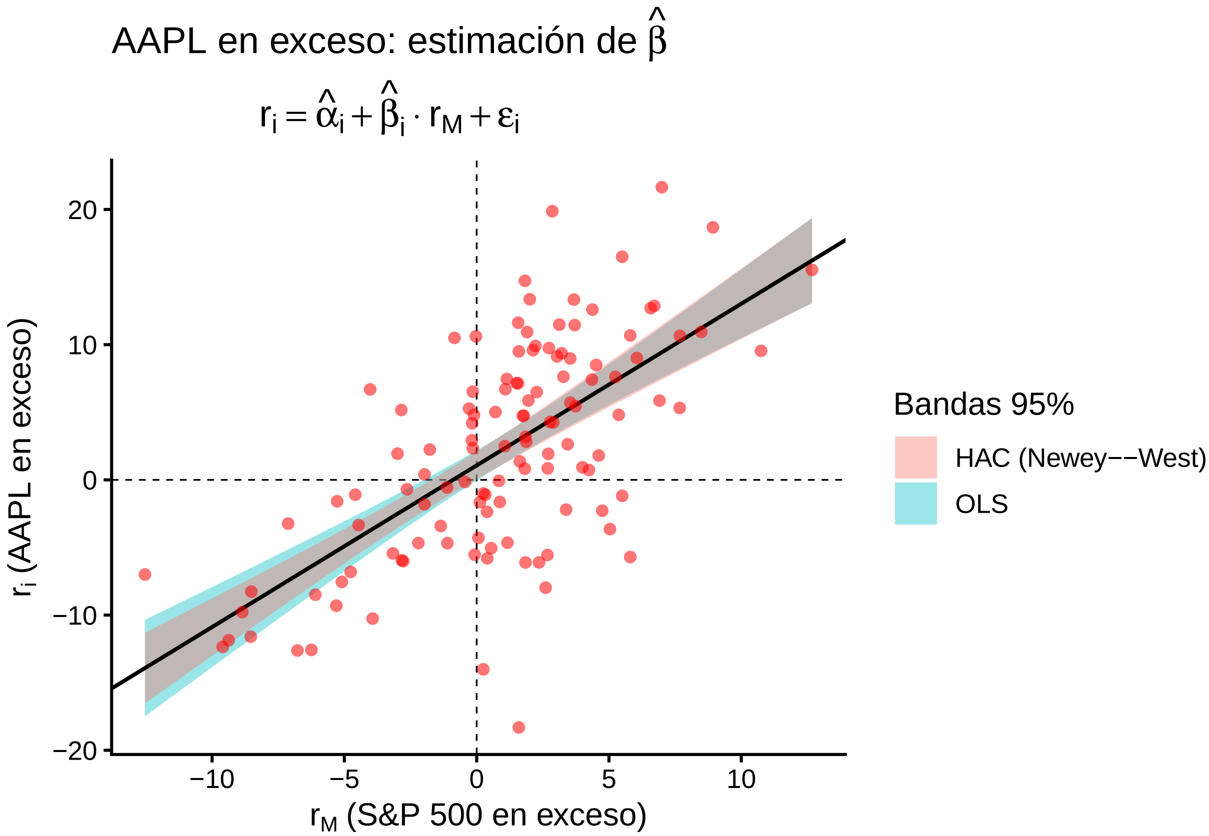
```
## # A tibble: 2 x 5
##   term  estimate std.error statistic  p.value
##   <chr>    <dbl>     <dbl>     <dbl>    <dbl>
## 1 alpha     1.07     0.564      1.89 6.05e- 2
## 2 beta      1.20     0.127      9.40 5.24e-16
```

```r
# (Opcional) Errores estándar HAC Newey--West para inferencia robusta
# Nota: el objetivo del video es operativo; HAC se reporta como referencia.
nw_L <- 3
aapl_hac <- coeftest(fit_aapl, vcov. = NeweyWest(fit_aapl, lag = nw_L, prewhite = FALSE)) |>
  tidy() |>
  mutate(term = recode(term, `(Intercept)` = "alpha", mkt_excess = "beta"),
         nw_lag = nw_L) |>
  select(term, estimate, std.error, statistic, p.value, nw_lag)

aapl_hac
```

```
## # A tibble: 2 x 6
##   term  estimate std.error statistic  p.value nw_lag
##   <chr>    <dbl>     <dbl>     <dbl>    <dbl>  <dbl>
## 1 alpha     1.07     0.523      2.04 4.33e- 2      3
## 2 beta      1.20     0.109     10.9  1.13e-19      3
```

Aplicación 1. Visualización: `AAPL`.

```r
# Aplicación 1. Visualización (AAPL): recta única y dos bandas (OLS vs HAC)
coef_aapl <- coef(fit_aapl)
alpha_hat <- unname(coef_aapl[1])
beta_hat  <- unname(coef_aapl[2])

x_grid <- tibble(mkt_excess = seq(min(aapl_df$mkt_excess, na.rm = TRUE),
                                  max(aapl_df$mkt_excess, na.rm = TRUE),
                                  length.out = 200))
# Banda OLS (media condicional)
pred_ols <- predict(fit_aapl, newdata = x_grid, interval = "confidence", level = 0.95)
pred_ols <- as_tibble(pred_ols) |>
  bind_cols(x_grid) |>
  rename(fit = fit, lwr = lwr, upr = upr)

# Banda HAC (media condicional)
V_nw <- NeweyWest(fit_aapl, lag = nw_L, prewhite = FALSE)

X_new   <- model.matrix(~ mkt_excess, data = x_grid)
fit_hat <- as.numeric(X_new %*% coef(fit_aapl))
se_nw   <- sqrt(diag(X_new %*% V_nw %*% t(X_new)))
crit    <- qnorm(0.975)

pred_nw <- tibble(mkt_excess = x_grid$mkt_excess, fit = fit_hat,
                  lwr = fit_hat - crit * se_nw, upr = fit_hat + crit * se_nw )

ggplot(aapl_df, aes(x = mkt_excess, y = excess_ret)) +
  geom_hline(yintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_vline(xintercept = 0, linetype = "dashed", linewidth = 0.3) +
  # Banda OLS (más transparente)
  geom_ribbon(data = pred_ols,
              aes(x = mkt_excess, ymin = lwr, ymax = upr, fill = "OLS"),
              alpha = 0.4, inherit.aes = FALSE) +
  # Banda HAC (menos transparente para distinguirla)
  geom_ribbon(data = pred_nw,
              aes(x = mkt_excess, ymin = lwr, ymax = upr, fill = "HAC (Newey--West)"),
              alpha = 0.40, inherit.aes = FALSE) +
  # Recta única (sin leyenda)
  geom_abline(intercept = alpha_hat, slope = beta_hat, linewidth = 0.7) +
  # Puntos rojos
  geom_point(color = "red", alpha = 0.55, size = 1.5) +
  labs(title = expression(atop("AAPL en exceso: estimación de"~hat(beta),
                          r[i,t] == hat(alpha)[i] + hat(beta)[i] %.% r[M,t] +
    epsilon[i,t])),
    x = expression(r[M,t]~"(S&P 500 en exceso)"),
    y = expression(r[i,t]~"(AAPL en exceso)"), fill = "Bandas 95%") +
  theme_classic(base_size = 12)
```



AAPL en exceso: estimación de $\hat{\beta}$
$r_i = \hat{\alpha}_i + \hat{\beta}_i \cdot r_M + \varepsilon_i$
20
10
0
−10
−20
−10
−5
0
5
10
r_M (S&P 500 en exceso)
r_i (AAPL en exceso)
Bandas 95%
HAC (Newey−−West)
OLS

Aplicación 1. Resultados: tres lecturas para interpretar el CAPM en exceso (AAPL).

A partir de la estimación en exceso

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t},$$

conviene distinguir tres lecturas que responden a preguntas distintas. La diferencia central es *qué objeto se quiere describir*: (i) la contribución sistemática asociada al mercado, (ii) la media condicional estimada por el modelo, o (iii) el rango plausible de una realización futura.

Aplicación 1. Lectura A: exposición sistemática y descomposición puntual.

Cuándo se usa. Como punto de partida, cuando el objetivo es construir intuición rápida y comparaciones: (i) cuantificar la exposición al mercado con $\hat{\beta}$ y (ii) evaluar escenarios simples (por ejemplo, un mes de mercado "bueno" o "malo") en términos de magnitudes puntuales.

- Interpretación de $\hat{\beta}$: la sensibilidad estimada es mayor que uno. En la muestra, `AAPL` reacciona más que proporcionalmente a variaciones del S&P 500 en exceso de la tasa libre de riesgo. En términos marginales, un cambio de 1% en $r_{M,t}$ se asocia con un cambio de 1.20% en $r_{i,t}$.
- Escenarios (solo exposición al mercado). Si el objetivo es aislar la parte atribuible *únicamente* a exposición sistemática, el componente asociado al mercado es
$$\hat{\beta}\, r_M.$$
Para un escenario $r_M = +5\%$, $\hat{\beta}\, r_M = 1.20 \times 5\% = 5.98\%$. Para $r_M = -5\%$, el componente es -5.98%.
- Escenarios (media condicional puntual). Si el objetivo es aproximar la media condicional en exceso bajo la especificación estimada,
$$\hat{r}_{i,t}(r_M) = \hat{\alpha} + \hat{\beta}\, r_M,$$
donde r_M se interpreta como un escenario para el rendimiento del mercado en exceso. Para $r_M = +5\%$, $\hat{r}_{i,t} = 1.07 + 1.20 \times 5\% = 7.05\%$. Para $r_M = -5\%$, $\hat{r}_{i,t} = -4.91\%$. En esta descomposición, $\hat{\beta}\, r_M$ captura exposición al mercado y $\hat{\alpha}$ representa un componente promedio residual en la muestra, condicionado al benchmark y al periodo.

Aplicación 1. Lectura B: media condicional con incertidumbre (IC 95% HAC).

Cuándo se usa. Cuando la pregunta no es solo "¿cuál es el valor puntual?", sino "*¿qué tan precisa es la media condicional que el modelo asigna?*". Esta lectura es útil para comunicar incertidumbre sobre $(\hat{\alpha}, \hat{\beta})$ sin confundirla con la variación idiosincrática mensual.

```r
# IC 95% HAC para la media condicional E[r_i | r_M = rM]
# Usa tus objetos: fit_aapl, nw_L

V_nw  <- NeweyWest(fit_aapl, lag = nw_L, prewhite = FALSE)
bhat  <- coef(fit_aapl)              # (alpha, beta)
crit  <- qnorm(0.975)                # 1.96 aprox.

ci_mean_hac <- function(rM, bhat, V, crit = qnorm(0.975)) {
  x   <- matrix(c(1, rM), ncol = 1)                 # (1, rM)'
  fit <- as.numeric(t(x) %*% bhat)                  # alpha + beta*rM
  se  <- sqrt(as.numeric(t(x) %*% V %*% x))         # HAC se for mean
  tibble(rM = rM, fit = fit, lwr = fit - crit*se, upr = fit + crit*se, se = se)
}

scenarios <- c(-5, 5)  # en % mensual
ci_mean_df <- bind_rows(
  ci_mean_hac(scenarios[1], bhat, V_nw, crit),
  ci_mean_hac(scenarios[2], bhat, V_nw, crit)
)

ci_mean_df
```

```
## # A tibble: 2 x 5
##      rM   fit   lwr   upr    se
##   <dbl> <dbl> <dbl> <dbl> <dbl>
## 1    -5 -4.91 -6.18 -3.63 0.651
## 2     5  7.05  5.38  8.71 0.849
```

Aplicación 1. Lectura B: interpretación.

- Los intervalos anteriores son rangos al 95% para la *media condicional* $E(r_{i,t} \mid r_{M,t} = r_M)$ bajo la especificación estimada. Estos rangos incorporan incertidumbre sobre $(\widehat{\alpha}, \widehat{\beta})$ mediante una matriz varianza–covarianza HAC (Newey–West).
- Interpretación: para un escenario r_M, el número central es

$$\widehat{r}_{i,t}(r_M) = \widehat{\alpha} + \widehat{\beta}\, r_M,$$

  y el intervalo $[\mathrm{lwr}, \mathrm{upr}]$ es un rango de incertidumbre para ese *promedio condicional*, no para una observación individual.
- Resultados en escenarios (HAC, 95%). Para $r_M = -5\%$, el modelo implica $\widehat{r}_{i,t} = -4.91\%$ y un intervalo al 95% de $[-6.18\%, -3.63\%]$, con EE HAC $= 0.65$. Para $r_M = +5\%$, $\widehat{r}_{i,t} = 7.05\%$ y el intervalo al 95% es $[5.38\%, 8.71\%]$, con EE HAC $= 0.85$.
- Lectura práctica: estos rangos responden a *"si el mercado en exceso fuera r_M, ¿qué tan imprecisa es la media que el modelo asigna a `AAPL`?"* dada la estimación de α y β y la corrección HAC.

Aplicación 1. Lectura C: intervalo de predicción (PI 95%).

Cuándo se usa. Cuando la pregunta es *prospectiva* y se refiere a una *realización mensual* posible, no a un promedio. Esta lectura es útil para planificación y gestión de riesgo porque incorpora la variación idiosincrática $\varepsilon_{i,t}$, que suele dominar la dispersión mensual.

```r
# PI 95% para una observación futura r_i | r_M = rM
# Aproximación: agrega la varianza residual a la incertidumbre del estimador

# Reutilizamos V_nw, bhat, crit del bloque anterior (si corres en secuencia).
# Si corres este chunk aislado, descomenta las tres líneas siguientes:
# V_nw <- NeweyWest(fit_aapl, lag = nw_L, prewhite = FALSE)
# bhat <- coef(fit_aapl)
# crit <- qnorm(0.975)

# Varianza residual (en %^2 mensual, consistente con excess_ret)
# Estimada por OLS como aproximación práctica
sigma2_hat <- sum(resid(fit_aapl)^2, na.rm = TRUE) / df.residual(fit_aapl)

pi_pred <- function(rM, bhat, V, sigma2, crit = qnorm(0.975)) {
  x   <- matrix(c(1, rM), ncol = 1)
  fit <- as.numeric(t(x) %*% bhat)
  se_mean <- sqrt(as.numeric(t(x) %*% V %*% x))
  se_pred <- sqrt(se_mean^2 + sigma2)
  tibble(
    rM = rM,
    fit = fit,
    pred_lwr = fit - crit*se_pred,
    pred_upr = fit + crit*se_pred,
    se_mean = se_mean,
    se_pred = se_pred
  )
}

scenarios <- c(-5, 5)
pi_df <- bind_rows(
  pi_pred(scenarios[1], bhat, V_nw, sigma2_hat, crit),
  pi_pred(scenarios[2], bhat, V_nw, sigma2_hat, crit)
)

pi_df
```

```
## # A tibble: 2 x 6
##      rM   fit pred_lwr pred_upr se_mean se_pred
##   <dbl> <dbl>    <dbl>    <dbl>   <dbl>   <dbl>
## 1    -5 -4.91   -16.8      7.00   0.651    6.08
## 2     5  7.05    -4.91    19.0    0.849    6.10
```

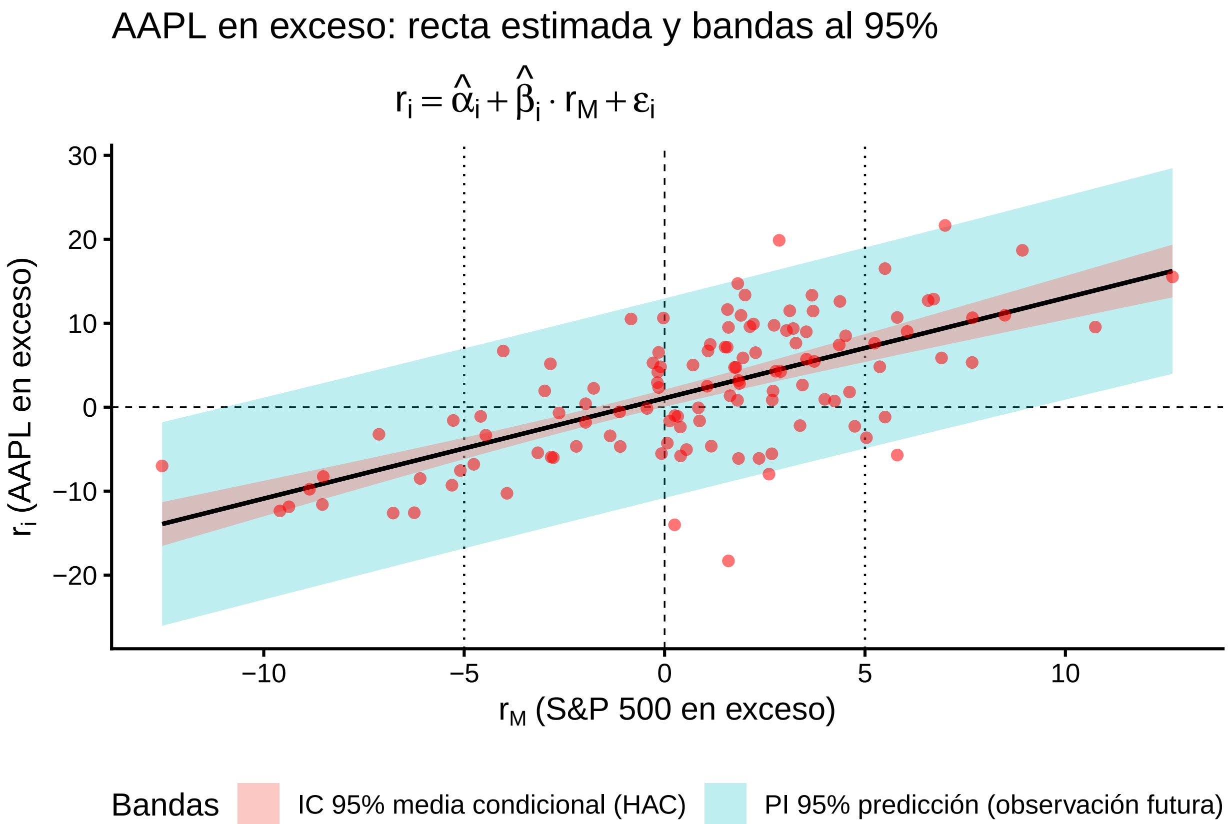
Aplicación 1. Lectura C: interpretación.

- Estos intervalos son de *predicción* al 95% para una observación futura $r_{i,t}$ condicionada en $r_{M,t} = r_M$. A diferencia del caso anterior, aquí se incorpora la dispersión no explicada por el mercado, $\varepsilon_{i,t}$, a través de la varianza residual $\hat{\sigma}^2$ (estimada por OLS como aproximación práctica en esta aplicación).
- Por construcción, el intervalo de predicción es sustancialmente más amplio que el intervalo para la media condicional: además de la incertidumbre sobre $(\hat{\alpha}, \hat{\beta})$, incorpora variación idiosincrática mensual de `AAPL` que el mercado no captura. En los escenarios considerados, el error estándar de la media condicional es relativamente pequeño, con $\widehat{se}_{mean} = 0.65$ para $r_M = -5\%$ y $\widehat{se}_{mean} = 0.85$ para $r_M = +5\%$, mientras que el error estándar de predicción es $\widehat{se}_{pred} = 6.08$ y $\widehat{se}_{pred} = 6.10$, reflejando que la componente residual domina la dispersión mensual.
- Resultados en escenarios (predicción, 95%). Para $r_M = -5\%$, el valor central del modelo es $\hat{r}_{i,t} = -4.91\%$ y el intervalo de predicción al 95% es $[-16.82\%, 7.00\%]$. Para $r_M = +5\%$, $\hat{r}_{i,t} = 7.05\%$ y el intervalo de predicción es $[-4.91\%, 19.01\%]$.
- Lectura práctica: estos rangos responden a *"si el mercado en exceso fuera r_M, ¿qué rango de retornos en exceso mensuales podría observar en `AAPL`?"* dentro del marco lineal estimado. La amplitud del intervalo subraya que, a frecuencia mensual, una fracción importante de la variación en retornos proviene de componentes idiosincráticos, aun cuando la exposición al mercado (capturada por $\hat{\beta}$) sea estimada con precisión.

Aplicación 1. Figura resumen: recta y bandas al 95% (AAPL).

```r
# Requiere objetos ya creados en tu flujo:
# aapl_df, fit_aapl, nw_L
# (Opcional, si ya corriste los bloques previos: V_nw, bhat, crit, sigma2_hat)

# ----------------------------
# Preparación: grid de escenarios r_M
# ----------------------------
x_grid <- tibble(
  rM = seq(min(aapl_df$mkt_excess, na.rm = TRUE),
           max(aapl_df$mkt_excess, na.rm = TRUE),
           length.out = 250)
)

# ----------------------------
# Matriz var-cov HAC y coeficientes
# ----------------------------
V_nw <- NeweyWest(fit_aapl, lag = nw_L, prewhite = FALSE)
bhat <- coef(fit_aapl)          # (alpha, beta)
crit <- qnorm(0.975)

# ----------------------------
# Media condicional y su IC 95% (HAC)
# ----------------------------
X_new   <- model.matrix(~ rM, data = x_grid)        # columnas: (Intercept), rM
fit_hat <- as.numeric(X_new %*% bhat)

se_mean <- sqrt(diag(X_new %*% V_nw %*% t(X_new)))

band_mean <- x_grid |>
  mutate(
    fit = fit_hat,
    lwr = fit_hat - crit * se_mean,
    upr = fit_hat + crit * se_mean,
    band = "IC 95% media condicional (HAC)"
  )

# ----------------------------
# Predicción: agrega varianza residual (aprox. práctica)
# ----------------------------
sigma2_hat <- sum(resid(fit_aapl)^2, na.rm = TRUE) / df.residual(fit_aapl)
se_pred <- sqrt(se_mean^2 + sigma2_hat)

band_pred <- x_grid |>
  mutate(
    fit = fit_hat,
    lwr = fit_hat - crit * se_pred,
    upr = fit_hat + crit * se_pred,
    band = "PI 95% predicción (observación futura)"
  )

# ----------------------------
# Gráfico: puntos + recta única + dos bandas
# Nota: la recta es la misma en ambos casos; solo cambiamos las bandas.
# ----------------------------
ggplot(aapl_df, aes(x = mkt_excess, y = excess_ret)) +
```

```
  geom_hline(yintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_vline(xintercept = 0, linetype = "dashed", linewidth = 0.3) +
  # Marcas de escenarios (opcional): r_M = ±5
  geom_vline(xintercept = c(-5, 5), linetype = "dotted", linewidth = 0.4) +
  # Banda de predicción (más amplia)
  geom_ribbon(
    data = band_pred,
    aes(x = rM, ymin = lwr, ymax = upr, fill = band),
    alpha = 0.25, inherit.aes = FALSE
  ) +
  # Banda de media condicional (más estrecha)
  geom_ribbon(
    data = band_mean,
    aes(x = rM, ymin = lwr, ymax = upr, fill = band),
    alpha = 0.40, inherit.aes = FALSE
  ) +
  # Recta estimada (sin leyenda)
  geom_line(
    data = band_mean,
    aes(x = rM, y = fit),
    linewidth = 0.8, inherit.aes = FALSE
  ) +
  # Puntos (rojo)
  geom_point(color = "red", alpha = 0.55, size = 1.5) +
  labs(
    title = expression(atop("AAPL en exceso: recta estimada y bandas al 95%",
                            r[i,t] == hat(alpha)[i] + hat(beta)[i] %.% r[M,t] +
    epsilon[i,t])),
    x = expression(r[M,t]~"(S&P 500 en exceso)"),
    y = expression(r[i,t]~"(AAPL en exceso)"),
    fill = "Bandas"
  ) +
  theme_classic(base_size = 12) +
  theme(legend.position = "bottom")
```



- La recta representa la media condicional estimada $\hat{r}_{i,t}(r_M) = \hat{\alpha} + \hat{\beta}\, r_M$.
- La banda de *media condicional (HAC)* refleja incertidumbre sobre $(\hat{\alpha}, \hat{\beta})$; típicamente es relativamente estrecha.
- La banda de *predicción* agrega variación idiosincrática $\varepsilon_{i,t}$ (vía $\hat{\sigma}^2$), por lo que es sustancialmente más amplia.

Aplicación 2. Contexto y criterio de uso (k_e vs WACC).

En esta aplicación distinguimos dos tasas que se usan según *a quién* pertenece el flujo que se está valuando. El costo de equity, k_e, es el rendimiento requerido por los accionistas y es la tasa adecuada para descontar flujos *al capital propio*. En términos prácticos, esto incluye dividendos, recompras y el *free cash flow to equity* (FCFE), entendido como el efectivo que queda *después* de cubrir gastos operativos, inversión y pagos a la deuda (intereses y amortizaciones), y que por lo tanto puede distribuirse a los accionistas sin comprometer la operación.

En cambio, el WACC se usa para descontar flujos *a la firma* o *a los activos*, antes de decidir cómo se reparte el valor entre acreedores y accionistas. Ese flujo es el *free cash flow to firm* (FCFF), que es el efectivo generado por las operaciones *antes* de pagos netos a la deuda y, por construcción, es un flujo disponible para *todos* los proveedores de capital.

Esta distinción también ayuda a entender la relación entre tasas. En el caso sin deuda ($D = 0$), el WACC coincide con el costo de equity, $WACC = k_e$, porque toda la financiación es capital propio. En situaciones típicas con deuda a un costo menor que el equity y con escudo fiscal, se espera $WACC < k_e$, ya que el WACC promedia una tasa de deuda usualmente más baja (después de impuestos) con la tasa de equity. Por ello, k_e no “compite” con el WACC: se usa con flujos distintos y, además, k_e es un insumo necesario para construir el propio WACC.

Aplicación 2. Definición operativa y pasos (CAPM para k_e).

- Objetivo: usar el CAPM para obtener un rendimiento requerido para el equity, k_e, a partir de (i) una estimación de exposición al mercado $\hat{\beta}$ y (ii) un supuesto explícito sobre la prima de riesgo del mercado.
- Paso 1 (estimación de β en exceso): mantenemos la notación del bloque anterior. Definimos rendimientos en exceso

  $$r_{i,t} = R_{i,t} - R_{f,t}, \qquad r_{M,t} = R_{M,t} - R_{f,t},$$

  y estimamos para `KO`

  $$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t}.$$

  El insumo empírico de esta aplicación es $\hat{\beta}_{KO}$.
- Paso 2 (costo de capital en niveles): el costo de equity se reporta como un rendimiento *total*. Bajo el CAPM,

  $$k_e = R_f + \hat{\beta}_{KO} \cdot ERP,$$

  donde ERP es la prima de riesgo del mercado definida como el exceso esperado del mercado,

  $$ERP \equiv E(r_M) = E(R_M) - R_f.$$

  En la regresión aparece $r_{M,t}$ como exceso *realizado* mes a mes; en costo de capital aparece ERP como exceso *esperado*, es decir, un supuesto de largo plazo.
- Punto clave: $\hat{\beta}_{KO}$ se estima con datos. En cambio, $ERP = E(r_M)$ es un *supuesto* (entrada externa) y suele dominar el nivel de k_e. Por ello conviene reportar un valor base y una sensibilidad de k_e ante distintos supuestos de ERP.
- Alcance: aquí solo construimos k_e (costo de equity). No construimos WACC ni incorporamos ajustes por apalancamiento, primas por país o por tamaño; esas extensiones requieren supuestos adicionales y se dejan fuera para mantener transparencia.

Aplicación 2. Estimación OLS de $\widehat{\beta}$ para `KO`.

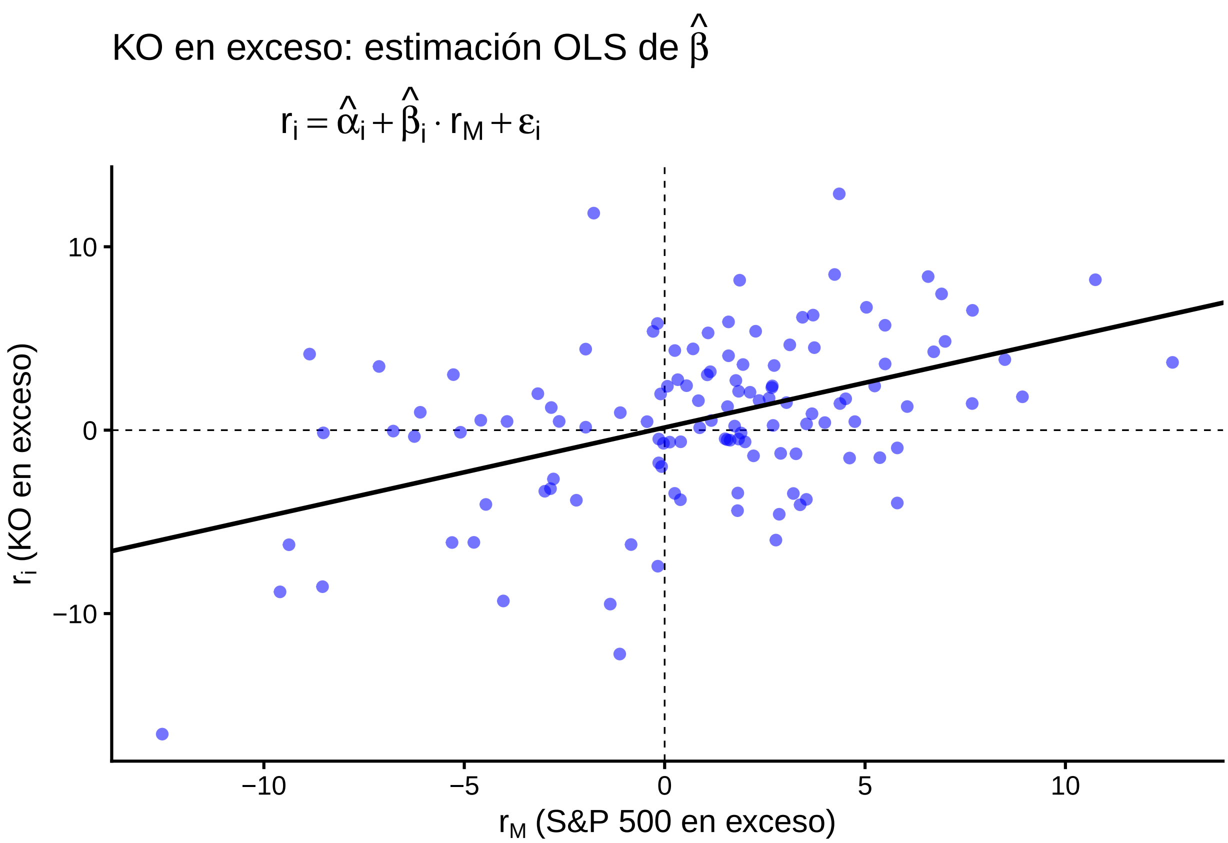
- Objetivo: estimar $\widehat{\beta}_{KO}$ como insumo para construir el costo de equity k_e bajo CAPM.
- Mantenemos exactamente la misma especificación en exceso que en la Aplicación 1:
  $$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t},$$
  donde $r_{i,t}$ corresponde a `KO` y $r_{M,t}$ al S&P 500 (`^GSPC`), ambos en exceso de $R_{f,t}$.
- En esta aplicación reportamos OLS únicamente; la corrección HAC se omitirá aquí porque ya fue discutida y visualizada en el bloque anterior.

```r
# Filtrar KO
ko_df <- capm_practical_data |>
  filter(ticker == "KO") |>
  select(date, excess_ret, mkt_excess) |>
  drop_na()

# Regresión CAPM en exceso (OLS)
fit_ko <- lm(excess_ret ~ mkt_excess, data = ko_df)

# Tabla OLS (alpha y beta)
ko_ols <- tidy(fit_ko) |>
  mutate(term = recode(term, `(Intercept)` = "alpha", mkt_excess = "beta")) |>
  select(term, estimate, std.error, statistic, p.value)

ko_ols
```

```
## # A tibble: 2 x 5
##   term  estimate std.error statistic     p.value
##   <chr>    <dbl>     <dbl>     <dbl>       <dbl>
## 1 alpha    0.144    0.384      0.375 0.709
## 2 beta     0.488    0.0865     5.64  0.000000116
```

Aplicación 2. Visualización (`KO`): dispersión $r_{i,t}$ vs $r_{M,t}$ y recta OLS.

```r
coef_ko <- coef(fit_ko)
alpha_hat_ko <- unname(coef_ko[1])
beta_hat_ko  <- unname(coef_ko[2])

ggplot(ko_df, aes(x = mkt_excess, y = excess_ret)) +
  geom_hline(yintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_vline(xintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_point(alpha = 0.55, size = 1.5, col = "blue") +
  geom_abline(intercept = alpha_hat_ko, slope = beta_hat_ko, linewidth = 0.8) +
  labs(
    title = expression(atop("KO en exceso: estimación OLS de"~hat(beta),
                            r[i,t] == hat(alpha)[i] + hat(beta)[i] %.% r[M,t] +
    epsilon[i,t])),
    x = expression(r[M,t]~"(S&P 500 en exceso)"),
    y = expression(r[i,t]~"(KO en exceso)")
  ) +
  theme_classic(base_size = 12)
```



Aplicación 2. Resultados (`KO`): insumo para costo de capital.

- Estimación (OLS):
  $\hat{\alpha}_{KO} = 0.14, \quad SE = 0.38,\ t = 0.37,\ p = 0.709,$
  $\hat{\beta}_{KO} = 0.49, \quad SE = 0.09,\ t = 5.64,\ p =< 0.001.$
- Lectura operativa: $\hat{\beta}_{KO} = 0.49$ implica que, en esta ventana y con el S&P 500 como benchmark, `KO` exhibe una exposición al riesgo de mercado inferior a la unidad. Este es el insumo central para convertir prima de mercado esperada en un rendimiento requerido para el equity.
- Nota sobre $\hat{\alpha}_{KO}$: el intercepto no es estadísticamente distinto de cero; en esta aplicación no se usa para construir k_e. El costo de equity bajo CAPM se apoya en R_f y en $ERP = E(r_M)$, escalados por $\hat{\beta}_{KO}$.
- Transición: en el siguiente bloque calculamos $k_e = R_f + \hat{\beta}_{KO} \cdot ERP$, reportando un caso base y sensibilidad ante distintos supuestos de ERP.

Aplicación 2. Costo de equity k_e con CAPM: caso base y sensibilidad a $ERP = E(r_M)$.

- Objetivo: construir k_e a partir de $\hat{\beta}_{KO}$ y supuestos explícitos para R_f y $ERP = E(r_M)$.
- Coherencia de frecuencia: $\hat{\beta}_{KO}$ se estimó con datos mensuales; para reportar un costo de capital anual, anualizamos R_f mensual y usamos un ERP anual (supuesto de largo plazo).
- Presentamos (i) un caso base con un ERP anual elegido y (ii) una tabla de sensibilidad de k_e ante varios valores plausibles de ERP.

```r
# --- Insumos ---
beta_ko_hat <- ko_ols |> filter(term == "beta") |> pull(estimate) |> as.numeric()

# RF anual: a partir del promedio mensual de rf en tu ventana (capm_practical_data)
# rf está en % mensual efectivo en tus datos
rf_mean_m <- capm_practical_data |>
  distinct(date, rf) |>
  summarise(rf_m = mean(rf, na.rm = TRUE)) |>
  pull(rf_m) |> as.numeric()

# Anualización compuesta desde % mensual -> % anual
rf_ann <- 100 * ((1 + rf_mean_m/100)^12 - 1)

# --- Supuestos para ERP anual = E(r_M) ---
# Puedes ajustar esta parrilla; por defecto muestro 4% a 8%.
erp_grid <- c(4, 5, 6, 7, 8)  # % anual

# --- Costo de equity (anual, %) ---
ke_table <- tibble(
  beta_hat = beta_ko_hat,
  rf_ann   = rf_ann,
  ERP_ann  = erp_grid
) |>
  mutate(ke_ann = rf_ann + beta_hat * ERP_ann)

ke_table
```

```
## # A tibble: 5 x 4
##   beta_hat rf_ann ERP_ann ke_ann
##      <dbl>  <dbl>   <dbl>  <dbl>
## 1    0.488   2.16       4   4.11
## 2    0.488   2.16       5   4.60
## 3    0.488   2.16       6   5.09
## 4    0.488   2.16       7   5.58
## 5    0.488   2.16       8   6.07
```

Aplicación 2. Interpretación: qué determina k_e y cómo leer la sensibilidad.

- Regla operativa (anual): con $ERP = E(r_M)$ en porcentaje anual,

  $$k_e = R_f + \hat{\beta}_{KO} \cdot ERP.$$

  Aquí R_f se anualiza a partir del promedio mensual observado en la ventana. En tu muestra, el promedio anualizado es $R_f = 2.16\%$.

- Resultados (caso base y sensibilidad): con $\hat{\beta}_{KO} = 0.49$, el costo de equity queda:

  $$k_e = 2.16\% + 0.49 \times ERP.$$

  En particular, si $ERP = 4\%$, $k_e = 4.11\%$; si $ERP = 5\%$, $k_e = 4.60\%$; si $ERP = 6\%$, $k_e = 5.09\%$; si $ERP = 7\%$, $k_e = 5.58\%$; y si $ERP = 8\%$, $k_e = 6.07\%$.

- Lectura de sensibilidad: un cambio de 1 punto porcentual en el ERP anual cambia k_e en $\hat{\beta}_{KO} = 0.49$ puntos porcentuales. Esta elasticidad es el vínculo operativo entre exposición al mercado y rendimiento requerido del equity.

- Mensaje central: $\hat{\beta}_{KO}$ determina *cuánto* del supuesto sobre ERP se traslada al costo de capital; el nivel de k_e depende de manera importante del valor adoptado para ERP. Por ello, en aplicaciones prácticas es preferible reportar un rango de ERP plausible en lugar de un único número.

Cierre (uso práctico, con supuestos explícitos): el propósito de estimar k_e no es "adivinar" el rendimiento de `KO`, sino fijar un *rendimiento requerido* para decisiones en las que el flujo relevante es del equity y el riesgo sistemático puede aproximarse por $\hat{\beta}_{KO}$ frente al S&P 500. Con $\hat{\beta}_{KO} = 0.49$ y $R_f = 2.16\%$ anual, el CAPM asigna un costo de equity entre 4.11% y 6.07% cuando se asume $ERP = E(r_M)$ entre 4% y 8%. Este rango se usa como referencia operativa en tres situaciones típicas: (i) *valuación por flujos al accionista* (FCFE/dividendos), donde se descuenta con k_e y la lectura es que el valor es sensible al supuesto de ERP; (ii) *evaluación de un proyecto desde el equity*, donde el criterio es comparar el rendimiento esperado del flujo al accionista con k_e y exigir un margen de seguridad si el proyecto tiene incertidumbre adicional no capturada por el mercado; y (iii) *insumo para WACC*, donde k_e entra como componente del capital propio y la comparación relevante se hace a nivel de WACC solo si el flujo es FCFF. En todos los casos, la condición crítica es de "comparabilidad de riesgo": si el flujo que se está evaluando no tiene un riesgo sistemático parecido al del equity de `KO` (por ejemplo, por apalancamiento distinto o por exposición sectorial distinta), entonces el ajuste correcto es revisar la β que representa ese riesgo (y, si es necesario, el benchmark), en lugar de interpretar k_e como una tasa universal aplicable a cualquier flujo.

Aplicación 2. Ejemplo hipotético: uso práctico de k_e.

Suponga que un analista quiere valorar el equity de `KO` con un flujo anual al accionista (FCFE) relativamente estable. Tiene una proyección simple: el flujo del próximo año será 100 unidades y espera que crezca 2% anual en el largo plazo. En este punto, el problema no es "si el flujo es correcto", sino *qué rendimiento exigir* dado el riesgo. Con $\widehat{\beta}_{KO} = 0.49$, el CAPM sugiere que el costo de equity es moderado y que depende del supuesto de prima de mercado: si el analista adopta un escenario conservador (ERP bajo), el rendimiento requerido para el equity cae cerca de la parte baja del rango (4.11%); si adopta un escenario más exigente (ERP alto), sube hacia la parte alta (6.07%).

La consecuencia económica es inmediata: con un k_e más bajo, el mismo flujo "vale más" hoy; con un k_e más alto, "vale menos". Por eso el resultado se usa como *regla de consistencia*: el analista no reporta un único valor como verdad, sino un intervalo de valuación coherente con supuestos explícitos sobre ERP y con una β estimada. Además, el analista pregunta si el flujo realmente tiene el riesgo del equity de `KO`. Si el flujo proviene de una expansión en un país con alta incertidumbre regulatoria, o de una línea de negocio con exposición distinta al ciclo económico, entonces la "comparabilidad de riesgo" falla: el ajuste correcto no es insistir en el mismo k_e, sino justificar por qué la exposición sistemática del proyecto debería ser mayor o menor que la de `KO`.

Aplicación 3. Contexto y objetivo: alpha de Jensen (`FDGRX`).

- Cuándo es útil: esta aplicación es relevante cuando se quiere evaluar el desempeño de un fondo o estrategia más allá de su rendimiento promedio, distinguiendo cuánto proviene de exposición al mercado y cuánto queda como excedente residual. El objetivo no es “predecir” rendimientos, sino evaluar si el desempeño observado es consistente con el riesgo sistemático asumido.

- Planteamiento: mantenemos la notación en exceso. Para el fondo (aquí `FDGRX`) definimos $r_{\text{fondo},t} = R_{\text{fondo},t} - R_{f,t}$ y para el mercado $r_{M,t} = R_{M,t} - R_{f,t}$, con el S&P 500 como benchmark.

- Modelo: estimamos la regresión CAPM en exceso

$$r_{\text{fondo},t} = \alpha + \beta\, r_{M,t} + \varepsilon_t.$$

  En esta aplicación, $\hat{\alpha}$ es el objeto central: mide el rendimiento en exceso promedio del fondo que no queda explicado por su exposición al mercado dada por $\hat{\beta}$.

- Interpretación operativa: $\hat{\beta}$ resume “cuánto mercado” hay en el fondo. $\hat{\alpha}$ resume “qué queda” después de descontar ese componente de mercado. Un $\hat{\alpha} > 0$ sugiere desempeño en exceso no atribuible a la exposición al benchmark en la muestra; un $\hat{\alpha} < 0$ sugiere lo contrario.

- Alcance y cautelas: $\hat{\alpha}$ es ex post y es condicional a (i) la ventana de estimación, (ii) la frecuencia mensual, y (iii) el benchmark elegido. Por ello, se interpreta como evidencia descriptiva y comparativa en la muestra, no como garantía de desempeño futuro.

Aplicación 3. Estimación OLS: $\hat{\alpha}$ y $\hat{\beta}$ de `FDGRX`.

- Objetivo: obtener una primera lectura empírica del desempeño ajustado por mercado de `FDGRX` en la muestra.
- Conservamos la misma especificación en exceso usada en las aplicaciones anteriores:
$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \varepsilon_{i,t},$$
donde ahora $r_{i,t}$ corresponde al fondo `FDGRX`.
- En esta página reportamos OLS como línea base; en la siguiente página añadiremos inferencia robusta (HAC) centrada en $\hat{\alpha}$.

```r
# Filtrar FDGRX
fdgrx_df <- capm_practical_data |>
  filter(ticker == "FDGRX") |>
  select(date, excess_ret, mkt_excess) |>
  drop_na()

# Regresión CAPM en exceso (OLS)
fit_fdgrx <- lm(excess_ret ~ mkt_excess, data = fdgrx_df)

# Tabla OLS (alpha y beta)
fdgrx_ols <- tidy(fit_fdgrx) |>
  mutate(term = recode(term, `(Intercept)` = "alpha", mkt_excess = "beta")) |>
  select(term, estimate, std.error, statistic, p.value)

fdgrx_ols
```

```
## # A tibble: 2 x 5
##   term  estimate std.error statistic  p.value
##   <chr>    <dbl>     <dbl>     <dbl>    <dbl>
## 1 alpha     1.05    0.278       3.78 2.47e- 4
## 2 beta      1.18    0.0627     18.9  2.12e-37
```

Aplicación 3. Inferencia robusta (HAC): $\hat{\alpha}$ en `FDGRX`.

- Objetivo: evaluar la evidencia estadística de $\hat{\alpha}$ usando errores estándar robustos a heterocedasticidad y autocorrelación (Newey–West).
- Mantenemos la misma regresión estimada en la página anterior y comparamos OLS vs HAC para transparencia inferencial.
- Foco de decisión: en esta aplicación el parámetro central es $\hat{\alpha}$; $\hat{\beta}$ se reporta como control de exposición al mercado.

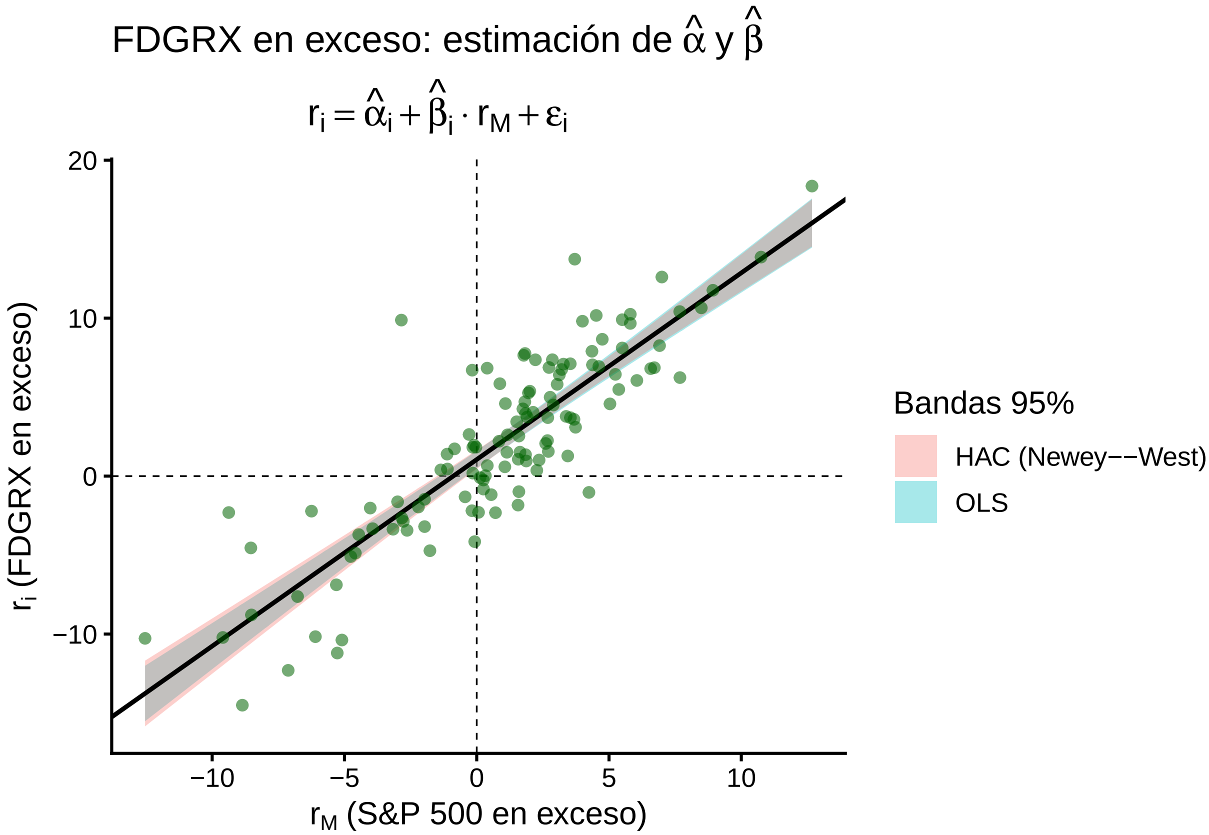
```r
# Errores estándar HAC Newey--West (misma lógica de Aplicación 1)
nw_L_fdgrx <- 3

fdgrx_hac <- coeftest(
  fit_fdgrx,
  vcov. = NeweyWest(fit_fdgrx, lag = nw_L_fdgrx, prewhite = FALSE)
) |>
  tidy() |>
  mutate(
    term = recode(term, `(Intercept)` = "alpha", mkt_excess = "beta"),
    method = "HAC (Newey-West)",
    nw_lag = nw_L_fdgrx
  ) |>
  select(method, term, estimate, std.error, statistic, p.value, nw_lag)

fdgrx_ols_cmp <- fdgrx_ols |>
  mutate(method = "OLS", nw_lag = NA_integer_) |>
  select(method, term, estimate, std.error, statistic, p.value, nw_lag)

fdgrx_inference <- bind_rows(fdgrx_ols_cmp, fdgrx_hac) |>
  arrange(term, method)

fdgrx_inference
```

```
## # A tibble: 4 x 7
##   method           term  estimate std.error statistic  p.value nw_lag
##   <chr>            <chr>    <dbl>     <dbl>     <dbl>    <dbl>  <dbl>
## 1 HAC (Newey-West) alpha     1.05    0.305       3.45 7.72e- 4      3
## 2 OLS              alpha     1.05    0.278       3.78 2.47e- 4     NA
## 3 HAC (Newey-West) beta      1.18    0.0694     17.0  1.36e-33      3
## 4 OLS              beta      1.18    0.0627     18.9  2.12e-37     NA
```

Aplicación 3. Visualización: recta y bandas de media condicional (FDGRX).

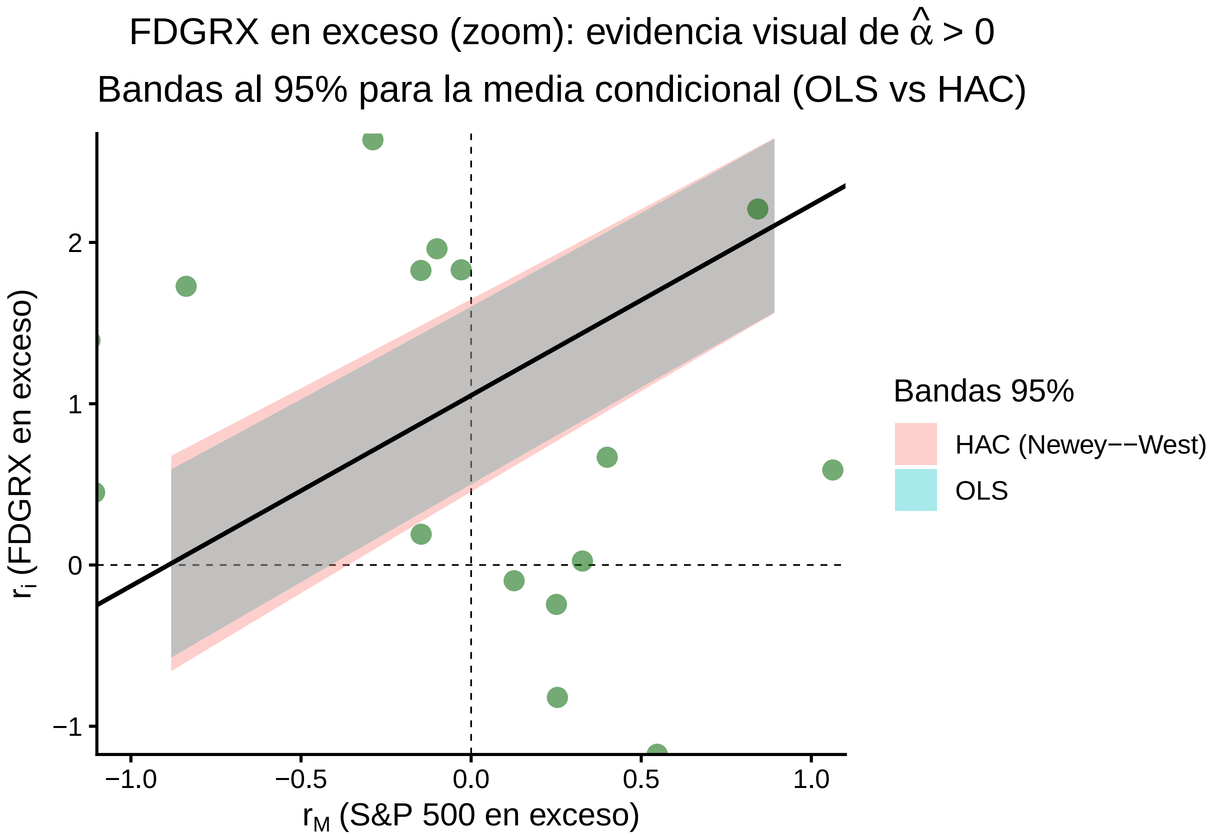
```r
# Coeficientes estimados
coef_fdgrx <- coef(fit_fdgrx)
alpha_hat_fdgrx <- unname(coef_fdgrx[1])
beta_hat_fdgrx  <- unname(coef_fdgrx[2])

# Grid para dibujar bandas
x_grid_fdgrx <- tibble(
  mkt_excess = seq(
    min(fdgrx_df$mkt_excess, na.rm = TRUE),
    max(fdgrx_df$mkt_excess, na.rm = TRUE),
    length.out = 200
  )
)

# Banda OLS (media condicional)
pred_ols_fdgrx <- predict(fit_fdgrx, newdata = x_grid_fdgrx, interval = "confidence", level =
↪ 0.95)
pred_ols_fdgrx <- as_tibble(pred_ols_fdgrx) |>
  bind_cols(x_grid_fdgrx) |>
  rename(fit = fit, lwr = lwr, upr = upr)

# Banda HAC (media condicional)
V_nw_fdgrx <- NeweyWest(fit_fdgrx, lag = nw_L_fdgrx, prewhite = FALSE)
X_new_fdgrx <- model.matrix(~ mkt_excess, data = x_grid_fdgrx)
fit_hat_fdgrx <- as.numeric(X_new_fdgrx %*% coef(fit_fdgrx))
se_nw_fdgrx <- sqrt(diag(X_new_fdgrx %*% V_nw_fdgrx %*% t(X_new_fdgrx)))
crit_fdgrx <- qnorm(0.975)

pred_nw_fdgrx <- tibble(
  mkt_excess = x_grid_fdgrx$mkt_excess,
  fit = fit_hat_fdgrx,
  lwr = fit_hat_fdgrx - crit_fdgrx * se_nw_fdgrx,
  upr = fit_hat_fdgrx + crit_fdgrx * se_nw_fdgrx
)

ggplot(fdgrx_df, aes(x = mkt_excess, y = excess_ret)) +
  geom_hline(yintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_vline(xintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_ribbon(
    data = pred_ols_fdgrx,
    aes(x = mkt_excess, ymin = lwr, ymax = upr, fill = "OLS"),
    alpha = 0.35, inherit.aes = FALSE
  ) +
  geom_ribbon(
    data = pred_nw_fdgrx,
    aes(x = mkt_excess, ymin = lwr, ymax = upr, fill = "HAC (Newey--West)"),
    alpha = 0.35, inherit.aes = FALSE
  ) +
  geom_abline(intercept = alpha_hat_fdgrx, slope = beta_hat_fdgrx, linewidth = 0.8) +
  geom_point(color = "darkgreen", alpha = 0.55, size = 1.5) +
  labs(
    title = expression(atop("FDGRX en exceso: estimación de"~hat(alpha)~"y"~hat(beta),
                            r[i,t] == hat(alpha)[i] + hat(beta)[i] %.% r[M,t] +
↪ epsilon[i,t])),
```

```r
    x = expression(r[M,t]~"(S&P 500 en exceso)"),
    y = expression(r[i,t]~"(FDGRX en exceso)"),
    fill = "Bandas 95%"
  ) +
  theme_classic(base_size = 12)
```



Aplicación 3. Visualización (zoom): foco en $\hat{\alpha}$ y bandas al 95% (`FDGRX`).

```r
# Zoom fijo alrededor de r_M = 0 para visualizar mejor el intercepto (alpha)
pred_ols_zoom <- pred_ols_fdgrx |>
  filter(mkt_excess >= -1, mkt_excess <= 1)

pred_nw_zoom <- pred_nw_fdgrx |>
  filter(mkt_excess >= -1, mkt_excess <= 1)

ggplot(fdgrx_df, aes(x = mkt_excess, y = excess_ret)) +
  geom_hline(yintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_vline(xintercept = 0, linetype = "dashed", linewidth = 0.3) +
  geom_ribbon(
    data = pred_ols_zoom,
    aes(x = mkt_excess, ymin = lwr, ymax = upr, fill = "OLS"),
    alpha = 0.35, inherit.aes = FALSE
  ) +
  geom_ribbon(
    data = pred_nw_zoom,
    aes(x = mkt_excess, ymin = lwr, ymax = upr, fill = "HAC (Newey--West)"),
    alpha = 0.35, inherit.aes = FALSE
  ) +
  geom_abline(intercept = alpha_hat_fdgrx, slope = beta_hat_fdgrx, linewidth = 0.8) +
  geom_point(color = "darkgreen", alpha = 0.55, size = 3) +
  coord_cartesian(xlim = c(-1, 1), ylim = c(-1, 2.5)) +
  labs(
    title = expression(atop("FDGRX en exceso (zoom): evidencia visual de"~hat(alpha)~"> 0",
                            "Bandas al 95% para la media condicional (OLS vs HAC)")),
    x = expression(r[M,t]~"(S&P 500 en exceso)"),
    y = expression(r[i,t]~"(FDGRX en exceso)"),
    fill = "Bandas 95%"
  ) +
  theme_classic(base_size = 12)
```



FDGRX en exceso (zoom): evidencia visual de α̂ > 0
Bandas al 95% para la media condicional (OLS vs HAC)
r_i (FDGRX en exceso)
r_M (S&P 500 en exceso)
2
1
0
−1
−1.0
−0.5
0.0
0.5
1.0
Bandas 95%
HAC (Newey−−West)
OLS

Aplicación 3. Resultados e interpretación: desempeño ajustado por mercado (`FDGRX`).

- Estimaciones principales (mensuales, en exceso):
  $\hat{\alpha}_{FDGRX} = 1.05, \quad SE_{OLS} = 0.28,\ t = 3.78,\ p =< 0.001,$
  $SE_{HAC} = 0.30,\ t_{HAC} = 3.45,\ p_{HAC} =< 0.001.$
  $\hat{\beta}_{FDGRX} = 1.18.$
- Lectura de $\hat{\alpha}$ (objeto central): el estimado puntual es positivo y estadísticamente significativo tanto con OLS como con HAC. En esta muestra hay evidencia de un desempeño residual positivo después de controlar por exposición al mercado.
- Lectura de $\hat{\beta}$: el fondo exhibe una exposición sistemática superior a la del mercado. Con $\hat{\beta} = 1.18$, un cambio de 1% en el exceso del mercado se asocia, en promedio condicional, con un cambio de 1.18% en el exceso de `FDGRX`.
- Mensaje operativo conjunto: en esta ventana, `FDGRX` combina una exposición sistemática elevada con un componente residual promedio positivo (alpha de Jensen) estadísticamente distinto de cero.
- Cautela de uso: esta evidencia es ex post y condicional a la ventana de 120 meses, frecuencia mensual y benchmark S&P 500. Se interpreta como diagnóstico de muestra y de gestión en ese periodo, no como garantía de desempeño futuro.

Aplicación 3. Cierre operativo: cómo usar $\hat{\alpha}_{FDGRX}$ en decisión.

- Regla de lectura: en esta muestra, la evidencia de $\hat{\alpha}_{FDGRX} > 0$ y significativa sugiere que el fondo entregó rendimiento en exceso promedio no explicado por su exposición al mercado (S&P 500), dentro del CAPM de un factor.
- Uso práctico: este resultado puede emplearse como insumo comparativo frente a otros fondos con mandato similar. La comparación relevante no es solo "quién rinde más", sino quién mantiene mejor desempeño ajustado por riesgo sistemático bajo el mismo benchmark, ventana y frecuencia.
- Límite metodológico: una alpha positiva en CAPM no prueba por sí sola "habilidad pura", porque puede capturar exposiciones a factores omitidos (por ejemplo, estilo growth, momentum o concentración sectorial). Si se busca atribución más fina de gestión, el siguiente paso natural es extender a un modelo multifactor.
- Mensaje final de la aplicación: en esta práctica, el CAPM en exceso permite separar tres piezas con sentido operativo: exposición sistemática ($\hat{\beta}$), rendimiento requerido (k_e) y desempeño residual ajustado por mercado ($\hat{\alpha}$). La calidad de la lectura depende de mantener supuestos explícitos y comparabilidad de diseño empírico.

Conclusiones.

- Aprendizaje central: una misma regresión CAPM en exceso puede responder preguntas distintas según el parámetro de interés. En esta práctica, ese mapa fue: $\hat{\beta}$ para exposición, k_e para tasa requerida y $\hat{\alpha}$ para desempeño ajustado por mercado.
- Mensaje metodológico: los resultados son útiles si se mantienen supuestos explícitos y consistencia de diseño (misma frecuencia, misma ventana y mismo benchmark). Cambiar esos elementos cambia la lectura económica de los coeficientes.
- Mensaje de decisión: en aplicaciones reales, el valor de los resultados no está en "predecir" retornos puntuales, sino en mejorar decisiones comparativas: cuánto riesgo sistemático se asume, qué rendimiento exigir y qué parte del desempeño queda sin explicar por el mercado.
- Criterio de prudencia: las estimaciones son evidencia de muestra (ex post), no garantías de desempeño futuro. Por eso, la interpretación correcta es condicional y debe acompañarse de juicio económico.