

Actividad semana 11

Integrantes:

-Llancari Nivin Meyli
-Vera Fonseca July
-Alejo Huamán Melissa
-Ccuro Minaya Lucia
-Flores Diaz, Christian

de Cox. Nos enfocaremos en la duración de la hospitalización (duraci_hospita_diaz) como tiempo y el fallecimiento (desenla_fallecido) como evento.

Instalación de paquetes (ejecutar solo si es necesario)

```
{r}
install.packages(c("tidyverse", "lubridate", "survival", "survminer", "gtsummary",
"broom", "here", "rio", "ggsurvfit"))
```

Carga de paquetes

```
{r}
library(tidyverse)
library(lubridate)
library(survival)
library(survminer)
library(gtsummary)
library(broom)
library(here)
library(rio)
library(ggsurvfit)

— Attaching core tidyverse packages
tidyverse 2.0.0 —
✓ dplyr    1.1.4   ✓ readr    2.1.5
✓ forcats  1.0.0   ✓ stringr  1.5.1
✓ ggplot2  3.5.2   ✓ tibble   3.2.1
✓ lubridate 1.9.4   ✓ tidyr    1.3.1
✓ purrr   0.1.4
— Conflicts
tidyverse_conflicts() —
```

Preparación de los datos

Filtrado de pacientes hospitalizados

Solo analizaremos pacientes hospitalizados (pac_fue_hospital == "Sí") ya que tienen datos de duración de hospitalización.

```
{r}
covid_data <- covid_data %>% filter(pac_fue_hospital == "Sí")
```

Recodificación del desenlace

Convertimos desenla_fallecido a numérico: 1 si falleció ("Sí"), 0 si no ("No").

```
{r}
covid_data <- covid_data %>%
  mutate(desenla_fallecido_num = if_else(desenla_fallecido == "Sí", 1, 0))
```

Conversión de variables de medición a numéricas y manejo de NA

Las variables clínicas tienen indicadores de medición (e.g., Saturación_O2_medida). Las convertimos a numéricas y asignamos NA cuando no se midieron.

```
{r}
# Función para convertir a numérico y setear NA si no fue medido
convert_to_numeric_na <- function(data, measure_var, value_var) {
  data[[value_var]] <- as.numeric(data[[value_var]])
  data[[value_var]] <- if_else(data[[measure_var]] == "No", NA_real_,
  data[[value_var]])
  return(data)
}

# Lista de parejas de variables (medida, valor)
measure_value_pairs <- list(
  c("Saturación_O2_medida", "Saturación_O2"),
  c("Temperatura_medida", "Temperatura"),
  c("Presión_arterial_media_medida", "Presión_arterial_media"),
  c("Dímero_D_medido", "Dímero_D"),
  c("Plaquetas_medidas", "Plaquetas"),
  c("INR_medido", "INR"),
  c("Nitrógeno_ureico_sangre_medido", "Nitrógeno_ureico_sangre"),
  c("Creatinina_medida", "Creatinina"),
  c("Sodio_medido", "Sodio"),
  c("Glucosa_medida", "Glucosa"),
```

Conversión de variables categóricas a factores

Convertimos las variables "Sí"/"No" a factores para su uso en modelos.

```
{r}
categorical_vars <- c("raza_negra", "raza_blanca", "asiatico", "latino",
                      "infacto_mioc", "enfer_vascu_periferica", "falla_card_conges"
                      ,
                      "enfer_cardiovascular", "demencia",
                      "enf_pulmonar_obstruc_cronica",
                      "diabetes_mell_complicada", "diabetes_mell_simple",
                      "enfer_renal",
                      "AII.CNS", "Pure.CNS", "Accidente_cerebrovascular",
                      "Convulsión",
                      "Síncope_previo", "Otra_afeccción_neurológica_previa",
                      "Otra_lesión_cerebral")

covid_data <- covid_data %>%
  mutate(across(all_of(categorical_vars), ~ factor(if_else(. == "Sí", "Sí", "No"))))
```

Análisis de Supervivencia

Método de Kaplan-Meier

Curva de supervivencia global

Estimamos la probabilidad de supervivencia (no fallecimiento) durante la hospitalización.

Método de Kaplan-Meier

Curva de supervivencia global

Estimamos la probabilidad de supervivencia (no fallecimiento) durante la hospitalización.

```
{r}
km <- survfit(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ 1, data =
covid_data)
```

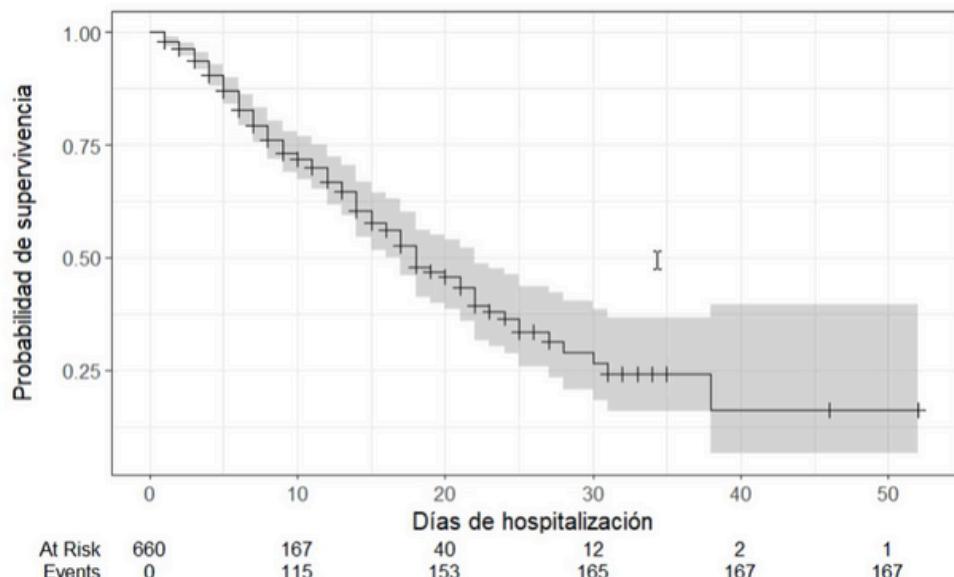
Gráfico de Kaplan-Meier

```
{r}
# Instalar el paquete si no lo tienes instalado
if (!requireNamespace("ggsurvfit", quietly = TRUE)) {
  install.packages("ggsurvfit")
}

# Cargar los paquetes necesarios
library(survival)
library(ggplot2)
library(ggsurvfit)

# Modelo de supervivencia Kaplan-Meier
surv_model <- survfit2(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ 1, data =
covid_data)

# Gráfico
ggsurvfit(surv_model) +
  labs(
    x = "Días de hospitalización",
    y = "Probabilidad de supervivencia")
```



Supervivencia a 30 días

```
{r}
summary(survfit(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ 1, data = covid_data), times = 30)
```

```
Call: survfit(formula = Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ 1, data = covid_data)

time n.risk n.event survival std.err lower 95% CI upper 95% CI
30      12      165     0.266  0.0508      0.182      0.386
```

Mediana del tiempo de supervivencia

```
{r}
survfit(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ 1, data = covid_data)
```

```
Call: survfit(formula = Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ 1, data = covid_data)

n events median 0.95LCL 0.95UCL
[1,] 660     167      18      16      22
```

Comparación entre grupos

Comparamos la supervivencia entre pacientes con y sin diabetes mellitus complicada.

```
{r}
survdiff(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ diabetes_mell_complicada, data = covid_data)
```

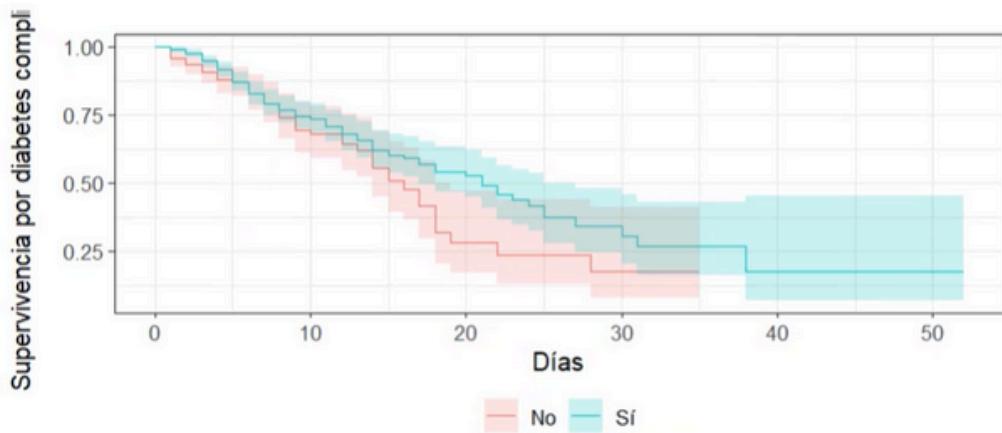
```
Call:
survdiff(formula = Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ diabetes_mell_complicada, data = covid_data)

N Observed Expected (O-E)^2/E (O-E)^2/V
diabetes_mell_complicada=No 181      53     43.1    2.273   3.18
diabetes_mell_complicada=Sí 479     114    123.9    0.791   3.18

Chisq= 3.2 on 1 degrees of freedom, p= 0.07
```

Gráfico estratificado por diabetes complicada

```
{r}
survfit2(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ diabetes_mell_complicada, data = covid_data) |>
ggsurvfit() +
labs(x = "Días", y = "Supervivencia por diabetes complicada") +
add_confidence_interval() +
add_risktable() +
add_pvalue(caption = "Log-rank {p.value}")
```



Modelo de Regresión de Cox

Modelo básico

Ajustamos un modelo de Cox con edad, severidad y diabetes complicada como covariables.

```
{r}
cox_model <- coxph(Surv(duraci_hospita_diaz, desenla_fallecido_num) ~ Edad +
severidad + diabetes_mell_complicada, data = covid_data)
```

Reporte del modelo

```
{r}
# Instalar el paquete si no lo tienes instalado
if (!requireNamespace("gtsummary", quietly = TRUE)) {
  install.packages("gtsummary")
}

# Cargar el paquete necesario
library(gtsummary)

# Suponiendo que ya tienes el modelo cox_model ajustado
cox_model %>%
 tbl_regression(exp = TRUE)
```

The package "broom.helpers" (>= 1.17.0) is required.
 Would you like to install it?

1: Yes

Análisis univariable con múltiples covariables

Generamos una tabla con modelos univariados para varias covariables.

```
{r}
# Instalar si no tienes el paquete gtsummary
if (!requireNamespace("gtsummary", quietly = TRUE)) {
  install.packages("gtsummary")
}

# Cargar los paquetes necesarios
library(gtsummary)
library(survival)
library(dplyr)

# Crear la tabla de análisis univariable
tabla_cox <- covid_data |>
 tbl_uvregression(
    method = coxph,
    y = Surv(duraci_hospita_diaz, desenla_fallecido_num),
    include = c(Edad, severidad, diabetes_mell_complicada, enfer_renal,
Saturación_O2),
    exponentiate = TRUE,
    conf.int = TRUE,
    hide_n = TRUE,
    pvalue_fun = ~ style_pvalue(.x, digits = 3),
    estimate_fun = ~ style_number(.x, digits = 2)
  ) |>
  bold_p(t = 0.05) |>
  modify_header(
    estimate = "***HR no ajustado***",
    p.value = "***Valor p***"
  )
```

Conclusión

Este análisis proporciona una visión inicial de la supervivencia en pacientes hospitalizados con COVID-19 y diabetes. Los pasos incluyen preparación de datos, estimación de curvas de supervivencia y modelado de riesgos. Puedes ajustar las covariables en el modelo de Cox o explorar otros grupos según tus necesidades.