

Tema V: Representación tabular.

Ejercicios Prácticos con `gtsummary`

Dr. Maicel Monzón Pérez

2025-03-24

Índice

Introducción	2
Instrucciones	3
Materiales	3
Tareas 1. Resumir variables en tablas simples con <code>tbl_summary</code>	4
Tareas 2. Describir variables por grupos con <code>tbl_summary</code>	4
Tareas 3. Describir variables por grupos y de manera global con <code>tbl_summary</code>	5
Tareas 4. Describir variables por grupos con <code>tbl_summary</code> e identificar diferencias.	6
Tareas 5. Crear Resumen por país y fortificación:	6
Tareas 6. Crear Tablas de contingencia con <code>tbl_cross()</code>	7
Tareas 7. Asociación univariada de variables con <code>tbl_uvregression</code>	7
Tareas 8. Asociación de variables con <code>tbl_regression</code> ajustada por múltiples variables	8
Conclusión	9

```
knitr::opts_chunk$set(echo = TRUE, warning = FALSE, message = FALSE)
library(gtsummary)
library(dplyr)
```

```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':
```

```
filter, lag
```

```
The following objects are masked from 'package:base':
```

```
intersect, setdiff, setequal, union
```

```
library(forcats)

# Cargar el dataset
datos <- read.csv("static/csv/datos_fortificacion.csv", stringsAsFactors = FALSE)
```

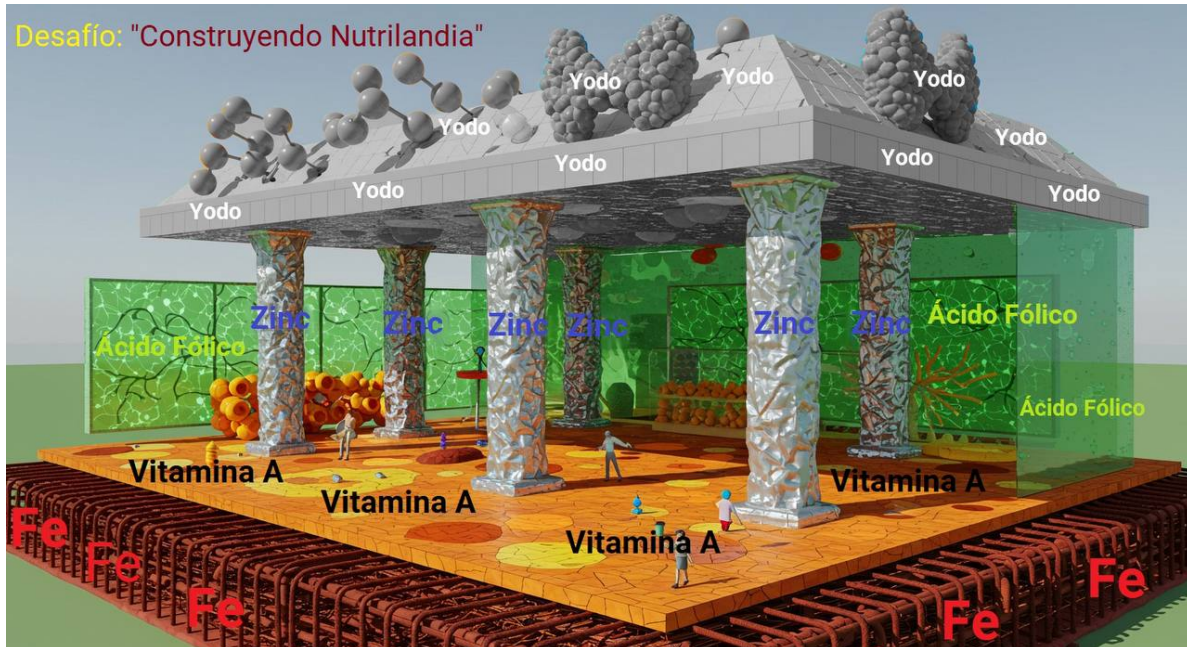
Introducción

Este ejercicio guiará a los estudiantes para explorar datos de fortificación de alimentos a gran escala usando la biblioteca `gtsummary`. El dataset incluye variables como:

- pais
- region
- alimento
- fortificado
- grupo_objetivo
- hierro_mg
- zinc_mg

Instrucciones

- Completa todas las tareas para obtener *5 puntos*.
- Usa el formato **R Markdown** o R script para entregar tus respuestas.



Materiales

- Datos de ejemplo para importar:
 - datos_fortificacion.csv

TIPS: use la función `view(nutricion_fortificacion)` y `glimpse(nutricion_fortificacion)` para ver la estructura del dataframe

Antes debemos importar los datos

```
library(tidyverse)
library(gtsummary)
datos <- read_csv("static/csv/datos_fortificacion.csv")
```

Tareas 1. Resumir variables en tablas simples con `tbl_summary`

-Describa las características fundamentales (variables resúmenes) de los encuestados en relación a variable como el área de residencia, el ingreso mensual y los alimentos ingeridos en la dieta?

Solución:

```
tbl <- datos %>%
  tbl_summary(include = c("area","ingreso_mensual","alimento"))
tbl %>% as_kable()
```

Characteristic	N = 500
area	
Rural	302 (60%)
Urbano	198 (40%)
ingreso_mensual	295 (210, 398)
alimento	
Aceite vegetal	82 (16%)
Arroz	77 (15%)
Azúcar	83 (17%)
Harina de trigo	81 (16%)
Leche en polvo	85 (17%)
Sal yodada	92 (18%)

Tareas 2. Describir variables por grupos con `tbl_summary`

-Describa las características fundamentales (variables resúmenes) del ingreso mensual y los alimentos ingeridos en la dieta según la viable área de residencia

Solución:

```
tbl <- datos %>%
  tbl_summary(
    include = c("ingreso_mensual","alimento"),
    by = "area")
tbl %>% as_kable()
```

Characteristic	Rural N = 302	Urbano N = 198
ingreso_mensual	300 (217, 406)	282 (209, 380)
alimento		
Aceite vegetal	57 (19%)	25 (13%)
Arroz	34 (11%)	43 (22%)
Azúcar	46 (15%)	37 (19%)
Harina de trigo	54 (18%)	27 (14%)
Leche en polvo	52 (17%)	33 (17%)
Sal yodada	59 (20%)	33 (17%)

Tareas 3. Describir variables por grupos y de manera global con `tbl_summary`

-Describa las características fundamentales (variables resúmenes) del ingreso mensual y los alimentos ingeridos en la dieta según la variable área de residencia y de manera general

Solución:

```
tbl <- datos %>%
  tbl_summary(
    include = c("ingreso_mensual", "alimento"),
    by = "area") %>%
  add_overall()

tbl %>% as_kable()
```

Characteristic	Overall N = 500	Rural N = 302	Urbano N = 198
ingreso_mensual	295 (210, 398)	300 (217, 406)	282 (209, 380)
alimento			
Aceite vegetal	82 (16%)	57 (19%)	25 (13%)
Arroz	77 (15%)	34 (11%)	43 (22%)
Azúcar	83 (17%)	46 (15%)	37 (19%)
Harina de trigo	81 (16%)	54 (18%)	27 (14%)
Leche en polvo	85 (17%)	52 (17%)	33 (17%)
Sal yodada	92 (18%)	59 (20%)	33 (17%)

Tareas 4. Describir variables por grupos con `tbl_summary` e identificar diferencias.

-Identifique diferencias entre el ingreso mensual y los alimentos ingeridos en la dieta por la variable área de residencia.

Solución:

```
tbl <- datos %>%
  tbl_summary(
    include = c("ingreso_mensual", "alimento"),
    by = "area") %>%
  add_p()

tbl %>% as_kable()
```

Characteristic	Rural N = 302	Urbano N = 198	p-value
ingreso_mensual	300 (217, 406)	282 (209, 380)	0.5
alimento			0.015
Aceite vegetal	57 (19%)	25 (13%)	
Arroz	34 (11%)	43 (22%)	
Azúcar	46 (15%)	37 (19%)	
Harina de trigo	54 (18%)	27 (14%)	
Leche en polvo	52 (17%)	33 (17%)	
Sal yodada	59 (20%)	33 (17%)	

Tareas 5. Crear Resumen por país y fortificación:

¿Cómo varía la fortificación y el ingreso por países?

Solución:

```
tbl <- datos %>%
  tbl_summary(
    by = pais,
    include = c(ingreso_mensual, fortificado)
  ) %>%
  add_difference()

tbl %>% as_kable()
```

Characteristic	Nepal N = 239	Nigeria N = 261	Difference	95% CI	p-value
ingreso_mensual	294 (212, 393)	296 (209, 403)	-4.3	-29, 21	0.7
fortificado	171 (72%)	186 (71%)	0.28%	-7.9%, 8.5%	>0.9

Tareas 6. Crear Tablas de contingencia con `tbl_cross()`

¿Existe una asociación significativa entre país y fortificación?

Solución:

```
tbl <- datos %>%
  tbl_cross(
    row = pais,
    col = fortificado,
    percent = "cell"
  ) %>%
  add_p()

tbl %>% as_kable()
```

	FALSE	TRUE	Total	p-value
pais				>0.9
Nepal	68 (14%)	171 (34%)	239 (48%)	
Nigeria	75 (15%)	186 (37%)	261 (52%)	
Total	143 (29%)	357 (71%)	500 (100%)	

Tareas 7. Asociación univariada de variables con `tbl_uvregression`

¿Cómo afecta el ingreso mensual a la probabilidad de fortificación?

Solución:

```
tbl <- datos %>%
  tbl_uvregression(
    include = c("pais", "ingreso_mensual", "hierro_mg"),
    y = "fortificado",
    method = glm,
    method.args = list(family = binomial),
    exponentiate = TRUE
  )

tbl %>% as_kable()
```

Characteristic	N	OR	95% CI	p-value
pais	500			
Nepal		—	—	
Nigeria		0.99	0.67, 1.45	>0.9
ingreso_mensual	500	1.00	1.00, 1.00	0.4
hierro_mg	500	1.04	0.99, 1.09	0.11

Tareas 8. Asociación de variables con `tbl_regression` ajustada por múltiples variables

¿Qué variables son significativas para predecir fortificación?

Solución:

```
modelo <- glm(fortificado ~ pais + ingreso_mensual + grupo_objetivo,
              data = datos, family = binomial)

tbl <- modelo %>% tbl_regression(exponentiate = TRUE)

tbl %>% as_kable()
```

Characteristic	OR	95% CI	p-value
pais			
Nepal	—	—	
Nigeria	0.98	0.67, 1.45	>0.9
ingreso_mensual	1.00	1.00, 1.00	0.4
grupo_objetivo			

Characteristic	OR	95% CI	p-value
Adolescentes	—	—	
Mujeres 15-49	0.75	0.46, 1.24	0.3
Niños 6-23m	0.64	0.40, 1.02	0.064

Conclusión

Este cuaderno cubre desde tablas básicas hasta análisis complejos con `gtsummary`. Los estudiantes deben: - Practicar con argumentos como `statistic`, `percent`, y `add_*`(). - Explorar el dataset para identificar patrones en fortificación alimentaria. - Aplicar conocimientos a contextos reales de salud pública.

Para más detalles, consulte la documentación de `gtsummary` y ejemplos de análisis de datos de encuestas .

