

Razón neutrófilos - linfocitos en el pronóstico de sobrevida en los pacientes con melanoma maligno cutáneo del Instituto Nacional De Enfermedades Neoplásicas 2005-2012

Bioestadística - Proyecto Final

Flor Castillo Silva

Gabriela Polo Mendoza

Diego Linares Trelles

Roosevelt Ubaldo Chavez

27 de marzo del 2021

Índice

1. Introducción	2
2. Objetivos	2
3. Desarrollo	2
3.1. Clasificación de variables	2
3.2. Paquetes utilizados	6
3.3. Carga y limpieza de la base de datos	6
3.4. Análisis Descriptivo	7
3.5. Análisis Bivariado	7
3.6. Análisis Multivariado	7
3.7. Análisis Gráfico	7
4. Conclusiones	7
5. Referencias	7

1. Introducción

2. Objetivos

3. Desarrollo

3.1. Clasificación de variables

3.1.1. Variables Cualitativas

3.1.1.1. Variables Dicotómicas

Variable	Nombre en el dataset	Valor
Género	genero	masculino=0, femenino=1
Antecedentes patológicos personales	antecedentes_patologicos	No=0, Sí=1
Hipertensión arterial	hipertension_arterial	No=0, Sí=1
Diabetes mellitus	diabetes_mellitus	No=0, Sí=1
Enfermedades del colágeno	enfermedades_del_colageno	No=0, Sí=1
Tuberculosis	tbc	No=0, Sí=1
Insuficiencia Renal Crónica	irc	No=0, Sí=1
Hepatitis	hepatitis	No=0, Sí=1
Asma	asma	No=0, Sí=1
Háb. nocivo tabaquismo	tabaquismo	No=0, Sí=1
Háb. nocivo alcohol	alcoholismo	No=0, Sí=1
Antecedentes familiares cáncer	antec_fam_cancer	No=0, Sí=1
Color de piel*	color_piel	Blanco=0, No blanco=1
Fase de crecimiento radial	fase_crec_radial	No=0, Sí=1
Fase de crecimiento vertical	fase_crec_vertical	No=0, Sí=1
Presencia de ulceración	presencia_ulceracion	No=0, Sí=1
Márgenes comprometidos	margenes_comprometidos	No=0, Sí=1
Microsatélites	microsatelites	No=0, Sí=1
LVI	lvi	No=0, Sí=1
PNI	pni	No=0, Sí=1
Ganglio centinela	gc	Positivo=1, Negativo=2
Ganglio centinela + Disección ganglionar	gc+_diseccion_ganglionar	No=0, Sí=1
Ampliación de márgenes	ampliacion_margenes	No=0, Sí=1
IFN por 3mo	ifnpor3mo	No=0, Sí=1
En tránsito	en_transito	No=0, Sí=1
Ganglios regionales	ganglios_regionales	No=0, Sí=1
Ganglios a distancia	ganglios_distancia	No=0, Sí=1
Metástasis cerebral	metastasis_cerebral	No=0, Sí=1
Metástasis a pulmón	metastasis_pulmon	No=0, Sí=1
Metástasis a hígado	metastasis_higado	No=0, Sí=1
Metástasis a mama	metastasis_mama	No=0, Sí=1
Metástasis retro peritoneal	metastasis_retro_peritoneal	No=0, Sí=1
Metástasis a bazo	metastasis_bazo	No=0, Sí=1
Metástasis gastrointestinal	metastasis_gastrointestinal	No=0, Sí=1

3.1.1.2. Variables políticas

Variable	Nombre en el dataset	Valores
Topografía	topografia	piel (valor único)

Variable	Nombre en el dataset	Valores
Morfología	morfologia	melanoma maligno, melanoma de diseminacion superficial, melanoma de celulas epitelioides, melanoma mixto de celulas epitelioides y fusiformes, melanoma lentiginoso acral, melanoma nodular, melanoma del lentigo, melanoma amelanotico, melanoma de celulas fusiformes, melanoma lentiginoso acral, nevus azul maligno, melanoma maligno acral, melanoma desmoplasico
Procedencia	procedencia	Amazonas=0, Ancash=1, ... , Ucayali=25
Oficia	oficia	agricultor=1, obrero=2, trabajo en oficina=3, casa=4, marino=5, jubilado=6, estudiante=7
Antecedentes cáncer	antec_cancer	No=0, Sí, cuello uterino=1, próstata=2, pulmón=3, mama=4, basocelular=5, tiroides=6, cáncer de estómago=7, carcinoma neuroendocrino=8, enfermedad de hogkin=9
Familiar antecedente cáncer	antec_fam_cancer	padre=1, madre=2, hermano=3, abuelos=4, tia=5, hija=6, prima=7
Tipo cáncer familiar	tipo_cancer_familiar	1=cancer de mama, 2=cancer de prostata, 3=cancer de cuello uterino , 4=cancer de pancreas, 5=pulmon, 6=melanoma, 7=linfoma, 8=cancer cerebral, 9=hueso, 10=estomago, 11=colon, 12=tiroides, 13=cancer de vejiga, 14=cancer hepatico, 15=cancer de faringe, 16=cancer de testiculo, 17=ovario
Diagnostico de consultorio externo	diagnostico_consultorio_externo	múltiples valores
Diagnostico	dx	múltiples valores
Diagnostico de anatomia patologica	diagnostico_anatomia_patologica	múltiples valores

Variable	Nombre en el dataset	Valores
Localización lesión	localizacion_lesion	1=cc, 2=tronco anterior, 3=tronco posterior, 4=miembro superior derecho brazo, 5=msd antebrazo, 6=msd mano, 7=msd dedos, 8=msi brazo, 9=msi antebrazo, 10=msi mano, 11=msi dedos, 12=mid muslo, 13=mid pierna, 14=mid pie, 15=mid dedos, 16= mii muslo, 17=mii pierna
Subtipos histológicos	subtipos_histologicos	extensión superficial=1, nodular=2, lentigo maligno=3, acral=4, desmoplástico=5, nevoide=6, amelanocítico=7, lentiginoso acral=8, no registrado=9, sarcomatoide=10
Relación neurofilos linfocitos	relacion_neurofilos_linfocitos	bajo<=3.0, intermedio=3.0 a 5.4, alto>5.4
Fase de crecimiento vertical	fase_crec_vertical	No=0, Sí=1, No registrado en historia clínica=2
Nivel de clark	nivel_clark	I=1, II=2, III=3, IV=4, V=5
Existencia de regresión	existencia_regresion	No=0, Sí < 50 %=1, Sí>50 %
TIL peritumoral	til_peri_tumoral	ausente=0, debil=1, moderado=2, intenso=3, muy intenso=4, no evaluado=5
TIL intratumoral	til_intra_tumoral	ausente=0, debil=1, moderado=2, intenso=3, muy intenso=4, no evaluado=5
Presencia de lesión melanocítica asociada	presencia_lesion melanocitica_asociada	No=0, Nevus melanocítico=1, nevus atípico=2
Tipo de célula predominante	tipo_celula_predominante	epiteloides=1, fusocelular=2, desmoplásico=3, no consignado=4, misto=5
T	t	in situ=0, T1a=1, T1b=2, T2a=3, T2b=4, T3a=5, T3b=6, T4a=7, T4b=8
N	n	N0=0, N1a=1, N1b=2, N2a=3, N2b=4, N2c=5, N3=6
M	m	M0=0, M1a=1, M1b=2, M1C=3

Variable	Nombre en el dataset	Valores
Staging	staging	in situ=0, IA=1, IB=2, IIA=3, IIB=4, IIC=5, IIIA=6, IIIB=7, IIIC=8, IV=9
Cirugía realizada	cirugia_realizada	reseccion local amp=1, rla+bgc+xeno injert=2, rla+diseccion inguinal=3, amputacion de dedos+bgc=4, amputacion de dedos+diseccion ganglionar=5, diseccion ganglionar=6, ampicion de margenes+bgc=7, ampicion de margenes+diseccion ganglionar=8, amputacion=9, biopsia de lesion=10, revision de laminas (no se opero)=11, ampliacion de margenes=12
Recidiva local	recidiva_local	No=0, Si=1, No se sabe=2
Metastasis	metastasis	No=0, Si=1, No se sabe=2
Seguimiento	seguimiento	libre de enfermedad=0, con enfermedad=1 no tiene seguimiento=2
Servicio de última atención	servicio_ultima_atencion	servicio de ultima atencion 0=stm,1=oncologia , 2=emergencia, 3=medicina paliativa, 4= cyc, 5 =radioterapia, 6=oftalmologia, 7= cx torax,8=gastroenterologia, 9=neurocirugia, 10 cirugia plastica, 11=uci, 12=servicio social, 13= , 14= ,15= , 16=
Evolución	evolucion	multiples valores
Observación	observacion	valores multiples

3.1.1.3. Variables ordinales

Variable	Nombre en el dataset	Valores
Edad	edad	0, ..., 98

3.1.2. Variables Cuantitativas

3.1.2.1. Variables continuas

Variable	Nombre en el dataset	Unidad
Leucocitos totales 1ra consulta	leucocitos_tot_1ra_consulta	-
Espesor del Breslow	espesor_mm_del_breslow	mm
Relación neurofilos linfocitos	relacion_neurofilos_linfocitos	adimensional

Variable	Nombre en el dataset	Unidad
Nro. de mitosis	nro_mitosis_mm2	mm ²
Num. absoluto de neutrofilos	numero_absoluto_neutrofilos	-
Num. absoluto de linfocitos	numero_absoluto_linfocitos	-

3.1.2.2. Variables discretas

Variable	Nombre en el dataset	Unidad
Año diagnostico	año_dx	año
Tiempo de enfermedad	tiempo_enfermedad_meses	meses
Fecha de diag. de anatomia patologica variable tiempo cero	fecha_diagnostico_anatomia_ patologi- ca_variable_tiempo_cero	fecha
Leucocitos totales de metastasis o ultimo control	leucocitos_totales_momento metasta- sis_o_ultimo_control	-
Fecha estudio ganglio	fecha_estudio_ganglio	fecha
Tiempo de recidiva	tiempo_recidiva_meses	meses
Fecha de Metastasis	fecha_metastasis	fecha
Fecha de ultimo seguimiento o fallecimiento	date_last_follow_up	fecha
Último control	ultimo_control	fecha

3.2. Paquetes utilizados

```
require(readxl) #para leer archivos en formato .xlsx
```

3.3. Carga y limpieza de la base de datos

```
# Cargamos la base de datos
data <- read_excel("database.xlsx")

# Revisamos si se cargó correctamente
head(data)

# Inspeccionamos brevemente la estructura de sus variables
str(data)
```

Dado que existen dos columnas con valores para la variable Relación de Neurofilos - Linfocitos, tabularemos la cantidad de NAs de ambas variables y cogeremos la columna que posea menor cantidad de datos faltantes.

```
# Reemplazar '.' por NA
cond1 <- data$relacion_neurofilos_linfocitos...35=='.' #condicion
data$relacion_neurofilos_linfocitos...35[cond1] <- NA # Valor
cond2 <- data$relacion_neurofilos_linfocitos...39=='.' #condicion
data$relacion_neurofilos_linfocitos...39[cond2] <- NA # Valor

table(is.na(data$relacion_neurofilos_linfocitos...35)) # Columna 1

##
## FALSE TRUE
## 618 264
```

```
table(is.na(data$relacion_neurofilos_linfocitos...39)) # Columna 2
```

##

```
## FALSE TRUE
```

```
##      326      556
```

Filtramos en función de nuestras variables de interés

```
data_limpia <- data[!cond1,]
```

```
paste(unique(tolower(data$morfologia)), collapse=', ')
```

```
## [1] "melanoma maligno, melanoma de diseminacion superficial, melanoma de celulas epitelioides, melanoma de celulas dendriticas"
```

3.4. Análisis Descriptivo

3.4.1. Clasificación de variables

3.5. Análisis Bivariado

3.6. Análisis Multivariado

3.7. Análisis Gráfico

4. Conclusiones

5. Referencias