

T  
519.53  
REY  
p.2



**ESCUELA SUPERIOR POLITECNICA DEL LITORAL**

**Instituto de Ciencias Matemáticas**

**"Estudio estadístico de la población infantil con cáncer en los  
principales centros urbanos del Ecuador"**

**TESIS DE GRADO**

**Previa a la obtención del Título de:  
INGENIERA EN ESTADISTICA INFORMATICA**

Presentada por:

***Sonnia Paola Reyes Ramos***



**GUAYAQUIL - ECUADOR**

**AÑO**

**2000**

# AGRADECIMIENTO

A Jehová

A mi familia

A mi mejor amigo

A todos mis maestros

Al Dr. Paulson





# DEDICATORIA



A MI MADRECITA  
POR SU VALENTÍA  
Y AMOR

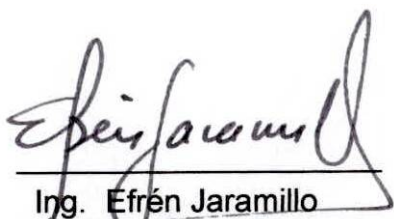
## TRIBUNAL DE GRADUACIÓN



Ing. Pablo Álvarez  
SUB-DIRECTOR DEL ICM



Ing. Margarita Martínez de Jordán  
DIRECTOR DE TESIS



Ing. Efrén Jaramillo  
MIEMBRO DEL TRIBUNAL



Mat. Eduardo Rivadeneira  
MIEMBRO DEL TRIBUNAL

# DECLARACIÓN EXPRESA

“ La responsabilidad del contenido de esta tesis de Grado, me corresponden exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”



---

Sonia Paola Reyes Ramos

## RESUMEN

El presente trabajo es un estudio de la población infantil con cáncer desde 1985 hasta 1996 que fueron atendidas en las instituciones especializadas de Quito y Guayaquil. Dichas instituciones especializadas son la Sociedad de Lucha contra el Cáncer, presente en las dos ciudades, y la Cruz Roja y el Hospital Vaca Ortiz en Quito.

La primera parte se centra en los aspectos generales del cáncer, como factores de riesgo, síntomas y tratamientos. A continuación se revisa estadísticas a nivel mundial y a nivel local de la población adulta y después de la población infantil.

Después se realiza estadística descriptiva de algunas variables de la población infantil que padece de esta enfermedad; como edad, sexo, lugar donde vive, estado actual. A continuación se realizan inferencias sobre independencia entre variables y sobre si el tiempo de supervivencia depende del tratamiento recibido.

Luego, se presentan tablas de supervivencia de cinco años para niños con cáncer y modelos de predicción para el número de pacientes atendidos en Solca de Guayaquil.

## INDICE GENERAL

	Pag.
RESUMEN	II
ÍNDICE GENERAL	III
ABREVIATURAS	IV
SIMBOLOGÍA	V
ÍNDICE DE TABLAS	VI
ÍNDICE DE GRÁFICOS	VII
ÍNDICE DE MATRICES	IX
INTRODUCCION	1
 I. ALGUNOS ASPECTOS DEL CÁNCER	 4
1.1 Definición	5
1.2 Posibles causas del cáncer	6
1.3 La detección es posible	12
1.3.1 La mamografía	12
1.3.2 El papanicolau	12
1.3.3 Otros métodos	13

1.4 Tratamientos para el cáncer	13
1.4.1 Cirugía	14
1.4.2 Radioterapia	15
1.4.3 Quimioterapia	16
1.4.4 Hormonoterapia	18
1.5 Estadísticas mundiales de la población con cáncer	18
1.5.1 El cáncer en los Estados Unidos	19
1.5.2 Cáncer maligno en la población europea. 1990	20
1.6 El cáncer en el Ecuador	21
1.6.1 Datos estadísticos de la población en Quito y Guayaquil	22
1.6.2 Los niños, un caso especial	32
 II. ESTADÍSTICA DESCRIPTIVA	 45
2.1 Sexo	47
2.2 Tipo de cáncer	48
2.3 Estado	51
2.4 Edad real	53
2.5 Provincia de nacimiento	54
2.6 Provincia donde vive	56
2.7 Tiempo de salida	58

2.8 Estadística descriptiva de la variable tiempo de supervivencia	
clasificada por el tipo de tratamiento recibido	60
2.8.1 Niños que no recibieron tratamiento	65
2.8.2 Niños que sólo recibieron quimioterapia	65
2.8.3 Niños que sólo recibieron radioterapia	66
2.8.4 Niños que sólo recibieron cirugía	67
2.8.5 Niños que recibieron quimioterapia y cirugía	68
2.8.6 Niños que recibieron quimioterapia y radioterapia	69
2.8.7 Niños que recibieron radioterapia y cirugía	69
2.8.8 Niños que recibieron los tres tratamientos	70
 III. ESTADÍSTICA INFERENCIAL	 72
3.1 Tablas de contingencia	73
3.1.1 Guayaquil	74
3.1.1.1 Lugar donde vive versus tipo de cáncer	74
3.1.1.2 Estado versus edad categorizada	79
3.1.1.3 Estado versus sexo	83
3.1.1.4 Estado versus tipo de cáncer	87
3.1.1.5 Tiempo de salida versus edad categorizada	91
3.1.1.6 Intervalos de tiempo de salida versus sexo	95
3.1.2 Guayaquil y Quito	100
3.1.2.1 Lugar donde vive versus tiempo de salida	100



3.1.2.2 Lugar donde vive versus estado	105
3.2 Análisis de varianza	109
3.2.1 La prueba de Kruskal Wallis	111
3.2.1.1 Prueba para todos los grupos	115
3.2.1.2 Prueba sin el grupo 000	116
3.2.1.3 Prueba con los grupos 10 y 100	118
3.2.1.4 Prueba con los grupos 110 y 101	119
3.2.1.5 Prueba con los grupos 001-011 y 111	120
3.2.1.6 Agrupación final	121
3.2.2 Resultados de las comparaciones y comparación del tiempo de supervivencia	122
IV TABLAS DE SUPERVIVENCIA	125
V MODELO PARA PREDECIR EL NÚMERO DE PACIENTES QUE SERÁN ATENDIDOS EN SOLCA	133
5.1 Los datos	134
5.2 ¿Qué método utilizaremos?	136
5.3 El método de regresión	143
5.3.1 Modelo de Regresión Multivariado	147
5.3.2 Verificación del modelo de regresión	160

GLOSARIO

APÉNDICE

BIBLIOGRAFÍA

## ABREVIATURAS

Estom	Carcinoma de estómago
HIG	Carcinoma en el hígado
HUE	Carcinoma en los huesos
LDH	Linfoma de hodkin
Leuc	Leucemia en general
Linf.	Linfomas
LLA	Leucemia linfoide aguda
LMA	Leucemia mieloide aguda
LNH	Linfomas que no son de hodkin
OJO	Cáncer o tumor en el ojo
OT	Otros carcinomas
PBL	Cáncer en las partes blandas
RIÑ	Cáncer en el riñón
SN	Cáncer en el sistema nervioso
VAR	Paciente sufre de varios carcinomas a la vez

## **SIMBOLOGÍA**

B	El estado del paciente es bueno hasta el momento
?	No se sabe el estado del paciente porque abandonó el tratamiento
F	El paciente falleció
SE	El paciente salió de la institución sin evidencia de actividad tumoral, pero no se sabe su estado actual.

## INDICE DE TABLAS

	Pag.
Tabla I	Factores de riesgos y síntomas de los diferentes carcinomas
	6
Tabla II	Carcinomas más comunes de mujeres. 1990
	20
Tabla III	Carcinomas más comunes de varones. 1990
	20
Tabla IV	Población Europea con cáncer
	20
Tabla V	Población Ecuatoriana con cáncer en porcentajes. 1985-1996
	23
Tabla VI	Población Ecuatoriana con cáncer. 1985-1996
	26
Tabla VII	Tabla de frecuencia de tipo de cáncer versus ciudad dónde viven los individuos en estudios
	29
Tabla VIII	Valores esperados de frecuencia
	30
Tabla IX	Desviaciones estándares
	30
Tabla X	Población infantil con cáncer en Ecuador por

	porcentajes según ciudad. 1985-1996	32
Tabla XI	Población ecuatoriana infantil con cáncer en porcentajes según tipo de cáncer, edad y ciudad. 1985-1996	34
Tabla XII	Población ecuatoriana infantil con cáncer según tipo de cáncer, edad y ciudad. 1985-1996	36
Tabla XIII	Tabla de frecuencia de tipo de cáncer versus edad de niños en estudios	37
Tabla XIV	Valores esperados de la frecuencia	38
Tabla XV	Desviaciones estándares	38
Tabla XVI	Tabla de frecuencia de tipo de cáncer versus ciudad de dónde proceden los niños en estudio	40
Tabla XVII	Valores esperados de la frecuencia	41
Tabla XVIII	Desviaciones estándares	41
Tabla XIX	Tabla de frecuencia de tipo de cáncer versus sexo del niño en estudio.	43
Tabla XX	Valores esperados de la frecuencia.	43
Tabla XXI	Desviaciones estándares.	44
Tabla XXII	Tabla de frecuencia de la variable sexo	47
Tabla XXIII	Tabla de frecuencia de la variable tipo de cáncer	49
Tabla XXIV	Tabla de frecuencia de la variable estado	51
Tabla XXV	Tabla de frecuencia de la variable edad real	53

Tabla XXVI	Tabla de frecuencia de la variable provincia de nacimiento	55
Tabla XXVII	Tabla de frecuencia de la variable provincia dónde vive	57
Tabla XXVIII	Tabla de frecuencia de la variable tiempo de salida	59
Tabla XXIX	Tabla de estadística descriptiva de la variable tiempo de supervivencia separadas por tipos de tratamientos	61
Tabla XXX	Tabla de contingencia de las variables lugar donde vive tipo de cáncer.	76
Tabla XXXI	Prueba $\chi^2$ para variables lugar dónde vive y tipo de cáncer	76
Tabla XXXII	Medidas direccionales para variables lugar dónde vive tipo de cáncer	77
Tabla XXXIII	Medidas simétricas para variables lugar dónde vive tipo de cáncer	77
Tabla XXXIV	Tabla de contingencias de las variables estado y edad categorizada	80
Tabla XXXV	Prueba $\chi^2$ para variables estado y edad categorizada	80
Tabla XXXVI	Medidas direccionales para variables estado y	



	edad categorizada.	81
Tabla XXXVII	Medidas simétricas para variables estado y edad categorizada.	82
Tabla XXXVIII	Tabla de contingencias de las variables estado y sexo	84
Tabla XXXIX	Prueba $\chi^2$ para variables estado y sexo	84
Tabla XL	Medidas direccionales para variables estado y sexo.	85
Tabla XLI	Medidas simétricas para variables estado y sexo.	86
Tabla XLII	Tabla de contingencias de las variables estado y tipo de cáncer.	88
Tabla XLIII	Prueba $\chi^2$ para variables estado y tipo de cáncer	88
Tabla XLIV	Medidas direccionales para variables estado y tipo de cáncer.	89
Tabla XLV	Medidas simétricas para variables estado y tipo de cáncer.	90
Tabla XLVI	Tabla de contingencias de las variables tiempo de salida y edad categorizada	92
Tabla XLVII	Prueba $\chi^2$ para variables tiempo de salida y edad categorizada.	93
Tabla XLVIII	Medidas direccionales para variables tiempo de salida y edad categorizada.	93



Tabla XLIX	Medidas simétricas para variables tiempo de salida y edad categorizada.	94
Tabla L	Tabla de contingencias de las variables intervalos de tiempo de salida y sexo	97
Tabla LI	Prueba $\chi^2$ para variables intervalos de tiempo de salida y sexo	98
Tabla LII	Medidas direccionales para variables intervalos de tiempo de salida y sexo.	98
Tabla LIII	Medidas simétricas para variables intervalos de tiempo de salida y sexo.	99
Tabla LIV	Tabla de contingencias de las variables lugar dónde vive y tiempo de salida	102
Tabla LV	Prueba $\chi^2$ para variables lugar dónde vive y tiempo de salida	102
Tabla LVI	Medidas direccionales para variables lugar dónde vive y tiempo de salida.	103
Tabla LVII	Medidas simétricas para variables lugar dónde vive y tiempo de salida.	104
Tabla LVIII	Tabla de contingencias de las variables lugar dónde vive y estado	106
Tabla LIX	Prueba $\chi^2$ para variables lugar dónde vive y estado.	107

Tabla LX	Medidas direccionales para variables lugar dónde vive y estado.	107
Tabla LXI	Medidas simétricas para variables lugar dónde vive y estado.	108
Tabla LXII	Nomenclatura de las diferente combinaciones de los tratamientos.	110
Tabla LXIII	Tabla de medias de rangos de la prueba de todos los grupos	115
Tabla LXIV	Estadísticos $\chi^2$	115
Tabla LXV	Tabla de medias de rangos de la prueba sin el grupo 000	116
Tabla LXVI	Estadísticos $\chi^2$	117
Tabla LXVII	Tabla de medias de rangos de la prueba de los grupos 10 y 100	118
Tabla LXVIII	Estadísticos $\chi^2$	118
Tabla LXIX	Tabla de medias de rangos de la prueba de los grupos 110 y 101	119
Tabla LXX	Estadísticos $\chi^2$	119
Tabla LXXI	Tabla de medias de rangos de la prueba de los grupos 001-011 y 111	120

Tabla LXXII	Estadísticos $\chi^2$	120
Tabla LXXIII	Tabla de medias de rangos con posibles agrupaciones finales	121
Tabla LXXIV	Tabla de supervivencia de niños con cáncer en Guayaquil	129
Tabla LXXV	Tabla de supervivencia de niños con cáncer en Quito	130
Tabla LXXVI	Número de paciente nuevos con tipo de cáncer más frecuentes de la ciudad de Guayaquil (1985-1996).	135
Tabla LXXVII	Valores del estadístico D para la prueba de Durbing Watson	137
Tabla LXXVIII	Comparación de números de pacientes pronosticados y real de 1997 y 1998	162

### Índice de gráficos

		Pag.
Gráfico 2.1	Sexo	47
Gráfico 2.2	Tipo de cáncer	50
Gráfico 2.3	Estado	52
Gráfico 2.4	Edad Real	54
Gráfico 2.5	Provincia de nacimiento	56
Gráfico 2.6	Provincia donde vive	58
Gráfico 2.7	Tiempo de salida	60
Gráfico 2.8	Gráfico de barras de tiempo de supervivencia según el tipo de tratamiento	64
Gráfico 3.1	Lugar donde vive y tipo de cáncer	78
Gráfico 3.2	Estado y edad categorizada	82

Gráfico 3.3	Estado y sexo	86
Gráfico 3.4	Estado y tipo de cáncer	90
Gráfico 3.5	Intervalos de tiempo de salida y edad categorizada	95
Gráfico 3.6	Intervalos de tiempo de salida y sexo	100
Gráfico 3.7	Lugar donde vive y tiempo de salida	104
Gráfico 3.8	Lugar donde vive y estado	108
Gráfico 3.9	Histograma: grupo con la mejor probabilidad de supervivencia	123
Gráfico 3.10	Histograma: Grupo con probabilidad de supervivencia intermedia	124
Gráfico 3.11	Histograma: Grupo con peor tiempo de supervivencia	124
Gráfico 4.1	Función de supervivencia	131
Gráfico 5.1	Autocorrelaciones de la variable mama	139
Gráfico 5.2	Autocorrelaciones parciales de la variable mama	139
Gráfico 5.3	Autocorrelaciones de la variable próstata	140
Gráfico 5.4	Autocorrelaciones parciales de la variable próstata	140

Gráfico 5.5	Autocorrelaciones de la variable mama con una diferenciación	141
Gráfico 5.6	Autocorrelaciones parciales de la variable mama con una diferenciación	142
Gráfico 5.7	Autocorrelaciones de la variable próstata con una diferenciación	142
Gráfico 5.8	Autocorrelaciones parciales de la variable próstata con una diferenciación	143
Gráfico 5.9	Comparación del valor real y la predicción del número de pacientes femeninas atendidas en Solca de Guayaquil en 1997	164
Gráfico 5.10	Comparación del valor real y la predicción del número de pacientes femeninas atendidas en Solca de Guayaquil en 1998	165
Gráfico 5.11	Comparación del valor real y la predicción del número de pacientes masculinos atendidos en Solca de Guayaquil en 1997	166
Gráfico 5.12	Comparación del valor real y la predicción del número de pacientes masculinos atendidos en Solca de Guayaquil en 1998	167



## Índice de Matrices

	Pag.
Matriz 5.1 Matriz de correlación estimada de número de pacientes nuevos masculinos	145
Matriz 5.2 Matriz de correlación estimada de número de pacientes nuevos femeninos	146
Matriz 5.3 Matriz de diseño de regresión multivariada	148
Matriz 5.4 Matriz de respuestas del modelo de regresión multivariada	149
Matriz 5.5 Matriz de parámetros desconocidos en regresión multivariada	149
Matriz 5.6 Matriz de errores en el modelo de regresión multivariada	150
Matriz 5.7 Matriz de diseño para hallar el número de pacientes que serán atendidos en Solca de Guayaquil	151
Matriz 5.8 Matriz de respuesta para los pacientes masculinos	152

Matriz 5.9	Matriz de respuesta para las pacientes femeninas	152
Matriz 5.10	Matriz de estimadores de los parámetros desconocidos para encontrar el número de pacientes masculinos	153
Matriz 5.11	Matriz de valores estimados de Y (número de pacientes masculinos con cáncer)	154
Matriz 5.12	Matriz de errores estimados	154
Matriz 5.13	Matriz de estimadores de los parámetros desconocidos para encontrar el número de pacientes masculinos	155
Matriz 5.14	Matriz de los valores estimados de Y (número de pacientes masculinos con cáncer)	156
Matriz 5.15	Matriz de los errores estimados	156
Matriz 5.16	Matriz del número estimado de pacientes masculinos para 1999	157
Matriz 5.17	Matriz del número estimado de pacientes femeninos para 1999	158
Matriz 5.18	Matriz de respuesta del número estimado de pacientes masculinos para 1997	159
Matriz 5.19	Matriz de respuesta del número estimado de pacientes femeninos para 1997	159
Matriz 5.20	Matriz de respuesta del número estimado de pacientes femeninos para 1998	159



Matriz 5.21	Matriz de respuesta del número estimado de pacientes masculinos para 1997	160
Matriz 5.22	Vector de índices de ajuste para pacientes masculinos	161
Matriz 5.23	Vector de índices de ajuste para pacientes femeninos	161
Matriz 5.24	Vector de medias cuadráticas del error para pacientes masculinos	162
Matriz 5.25	Vector de medias cuadráticas del error para pacientes femeninos	162

# INTRODUCCIÓN

Este trabajo de tesis es la presentación del análisis estadístico a nivel descriptivo e inferencial de la población infantil con cáncer atendidas en la Sociedad de Lucha contra el Cáncer de Quito y Guayaquil, la Cruz Roja de Quito y el Hospital Vaca Ortiz de la misma ciudad; en el periodo comprendido entre 1985 y 1996.

Los motivos fundamentales para centrar nuestro estudio en la población infantil son

1. El alto índice de abandono de tratamiento de los pacientes adultos diagnosticados con cáncer. Ya que, regularmente, los adultos se preocupan más por el cuidado y tratamiento de sus hijos, que por el suyo propio; el índice de abandono en la población infantil atendida en las instituciones es menor que el de la población adulta.
2. Los infantes que sufren de cáncer se podrían considerar como un grupo de la población de pacientes con características bien definidas; puesto

que los cánceres que padecen los niños son diferentes a la de los adultos; y tienen causas diferentes, como se podrá constatar más adelante.

3. El breve periodo de tiempo disponible para realizar mi estudio, que no permite recoger la información deseada, por medio del seguimiento respectivo. Cabe recalcar que las instituciones antes mencionadas no hacen seguimiento de los pacientes que abandonan el tratamiento para constatar su estado, por considerarlo muy costoso.

Los objetivos fundamentales son tener evidencia estadística si el tipo de carcinoma que se tenga se relaciona con factores externos como la edad, sexo o ciudad donde resida el niño; determinar si el tipo de tratamiento recibido por el infante influye en el tiempo de supervivencia del paciente, estudiar el comportamiento de la función de supervivencia durante los primeros años de tratamiento de los principales carcinomas que afectan a los niños; y, obtener un modelo para predecir el número de casos nuevos de pacientes con cáncer para los años posteriores.

Para llevar a cabo los objetivos señalados, utilizaremos teorías estadísticas para hacer inferencias sobre los datos como tablas de contingencias, análisis de varianza, regresión multivariada y tablas de supervivencia. Además se utilizó tablas de frecuencia y gráficos de barra para el análisis descriptivo.

Al final se resumen los resultados de los estudios estadísticos realizados y se recomiendan aspectos de funcionamiento que las instituciones auspiciantes deberían mejorar para realizar su labor más eficazmente.

## **CAPITULO I**

# **ALGUNOS ASPECTOS DEL CÁNCER**

# Capítulo 1

## 1. ALGUNOS ASPECTOS DEL CÁNCER

### 1.1 Definición

Cáncer es un grupo de enfermedades caracterizadas por el crecimiento y esparcimiento incontrolables de células anormales. Al no contrarrestar el esparcimiento incontrolable de estas células con algún método externo al cuerpo, estas células anormales pueden invadir órganos vitales, lo que desembocaría en la muerte del paciente.

A pesar de tener como característica común el crecimiento de las células anormales, lo cual desemboca en la muerte; el cáncer se puede considerar como varias enfermedades en una y las diferencias radican en las causas que lo provocan y los tratamientos para contrarrestarlo.



## 1.2 Posibles causas del cáncer

Los científicos John Higginson y Sir Richard Doll, en sus investigaciones realizadas en los años 60, han concluido que del 80 al 90% de los cánceres son causados por factores ambientales tales como exposición a agentes químicos específicos o virus, factores dietéticos y patrones reproductivos; mientras que sólo del 10 al 20% son causados por factores internos (hormonas, condiciones de inmunidad y mutaciones inherentes). Los factores que causan el cáncer pueden actuar juntos o en secuencia al iniciar o promover la carcinogénesis.

A continuación se muestra una tabla con los factores de riesgo y síntomas de algunos tipos de cáncer según estos dos científicos.

TABLA I

### FACTORES DE RIESGO Y SÍNTOMAS DE LOS DIFERENTES CARCINOMAS

Tipo	Factores de riesgo	Síntomas
Vejiga y riñón	Exposición a algunos químicos, como la naftalina.	Sangre al orinar Dolor o ardor al orinar.

	Excesivo consumo de cafeína	Incremento en la frecuencia al orinar.
	y/o azúcares artificiales.	
	Una historia de cistosomiasis.	
	Infecciones urinarias	
	frecuentes.	
	Fumar	
Mamas	Tener el primer hijo después de los 35 años.	Masas, engrosamiento u otros cambios físicos en las mamas.
	No tener hijos	Picazón, enrojecimiento y/o
	Historia familiar con este tipo de cáncer	ardor en los pezones no asociados con la lactancia.
	Consumo alto de cafeína o alcohol.	
	Dietas altas en calorías.	
	Diabetes.	
Utero	Más de 5 partos.	Sangrado entre menstruaciones.
	Primera relación sexual antes de los 18.	Periodos con abundante flujo.
		Menstruaciones dolorosas.
	Historial de gonorrea u otras enfermedades de transmisión sexual.	
	Muchos compañeros	



	sexuales.	
	Infertilidad	
Colon	Falta de fibra y calcio en la dieta.	Sangrado rectal
	Pólipos	Cambios radicales en los hábitos de defecación.
	Familiares con historial con este tipo de cáncer.	
	Estreñimiento o diarreas continuas	
	Dieta alta en grasas	
Endometrio	Nunca haber estado embarazada	Sangrado entre menstruaciones.
	Al pasar la menopausia	Periodos con abundante flujo.
	Historia familiar	Menstruaciones dolorosas.
	Diabetes	
	Obesidad	
	Hipertensión	
Laringe	Hábitos fuertes de fumar	Tos persistente
	Consumo de alcohol	Ronquera
Leucemia	Factores hereditarios	Palidez

	Exposición a radiación	Fatiga
	Infecciones virales crónicas	Pérdida de peso
		Infecciones repetidas
		Moretones
		Dolor de huesos y articulaciones
		Sangrado de la nariz
Pulmón	Fumar	Tos persistente
	Exposición a asbestos, níquel, cromados o material radiactivo.	Esputo con sangre
		Dolor de pecho
	Bronquitis crónica	
	Historial de tuberculosis	
	Exposición a pesticidas o herbicidas.	
Linfoma	Factores hereditarias	Nódulos linfáticos dilatados
	Disfunción en el sistema inmunológico.	Picazón
		Sudor nocturno
		Pérdida de peso y fiebres inexplicables
Boca y cuello	Irritación en la boca	Úlceras crónicas en la boca,
	Consumo excesivo de alcohol	lengua o garganta que no

	Fumar o mascar tabaco	sanar.
Ovario	No tener hijos	No se desarrolla síntomas
	Dieta alta en grasas	hasta que está en etapa avanzada
Próstata	Infecciones recurrentes en la próstata.	Fluido urinario débil o interrumpido.
	Dieta alta en grasa animal	Dolor continuo en la parte baja
	Alto consumo de leche, carnes y/o café.	de la espalda, pelvis y/o parte superior de las caderas.
	Alto consumo de testosterona en tratamientos para la impotencia	
	Vasectomía	
	Tener más de 50 años	
Piel	Exposición al sol, sobretodo a las personas con piel delicada.	Un tumor bajo la piel. Una ulceración que no sana Lunares que cambian de color o tamaño
	Historial de lunares	
	Lunares en los pies o en lugares irritados por la ropa.	Lesiones que parecen lunares
	Cicatrices de quemaduras	

	severas y ardor que no sana	
	Historial familiar	
Estómago	Anemia perniciosa	Indigestión y dolor después de
	Falta de fibra en la dieta	comer
	Dieta alta en grasa	Pérdida de peso.
	Gastritis crónica	
	Pólipos en el estómago.	
Testículo	Testículo ascendente	Ampollas
		Dilatación del testículo
		Engrosamiento del escroto
		Repentina aparición de fluido
		en el escroto
		Dolor en el escroto
		Dolor en la parte baja del
		abdomen

Todos tenemos riesgo de desarrollar cáncer. Ya que la mayoría de los casos de cáncer tienen causa externa, según la evidencia científica, podemos concluir que el cáncer puede ser prevenido.

### **1.3 La detección es posible**

Exámenes frecuentes conducidos por un profesional pueden resultar en la detección del cáncer de mama, colon, recto, útero, próstata, testículos, cavidad bucal y piel en una etapa temprana, donde el tratamiento es más exitoso.

#### **1.3.1 La mamografía**

La evaluación del uso de la mamografía como método de detección de cáncer de mama es complicada.

Los principales beneficios de chequeos continuos por medio de mamografías, sobretudo en mujeres con un alto riesgo, son: una mayor posibilidad de detectar a tiempo el cáncer de mama, seguridad de médicos y pacientes, y una relación favorable costo-beneficio.

#### **1.3.2 El papanicolaou**

El papanicolaou ha ayudado a la detección eficaz de lesiones preneoplásicas, displasias y carcinomas in situ en el cuello uterino. Al ser tratadas estas lesiones eficazmente, nunca evolucionan hasta la fase de carcinoma invasivo.

Ya que uno de los grupos más afectados por el cáncer al útero son las mujeres de un nivel socioeconómico bajo, se exige un mayor esfuerzo



social para prevenirlo por medio de los sistemas de información, con el fin de que todas las lesiones anormales sean tratadas.

Se recomienda que el intervalo de la realización de este examen sea de un año. Esta conclusión se basa en la historia natural del cáncer uterino, cuya evolución es de varios años.

### **1.3.3 Otros metodos**

Los carcinomas al colon y recto son muy difíciles de diagnosticar, ya que sus síntomas son muy comunes, aún en la población sin cáncer. Lo que se puede es tener un chequeo continuo, sobretodo las personas de grupos de alto riesgo, por medio de los marcadores tumorales como el antígeno carcinoembrionario.

Se recomienda a fumadores mayores de 45 años tomarse radiografías a los pulmones anualmente.

Se recomienda a hombres mayores de 60 años hacerse una cuidadosa palpación digital anualmente, para facilitar la detección del cáncer en las primera etapas, donde aún puede ser curado.

## **1.4 Tratamientos para el cáncer**

Los tratamientos ortodoxos o convencionales hasta ahora para el cáncer son la cirugía, radiación, quimioterapia y hormonoterapia.

### 1.4.1 Cirugía

Hay diferentes tipos de cirugías:

**Cirugía curativa:** terapia locorregional que puede curar sólo aquellos procesos localizados en su tejido originario y en sus ganglios regionales.

**Cirugía preventiva:** es el método de elección para las lesiones in situ en todas las superficies epiteliales del cuerpo como la piel, boca, cuello uterino; y premalignas como pólipos gastrointestinales, adenomas de tiroides o mama.

**Cirugía diagnóstica:** como biopsias, laparotomías exploradoras, laparoscopías.

**Cirugía paliativa:** pocas veces se utiliza esta técnica para curar o prevenir algún síntoma: resección de lesiones que causan hemorragia, amputación de miembros con sarcomas dolorosos o sangrantes, extirpación de mamas ulceradas, descompresión de la médula afectada por el tumor.

**Cirugía reductora:** extirpar la mayor cantidad posible de masa tumoral con la esperanza de que la quimioterapia o la radioterapia puedan



detener el crecimiento o curar los pequeños restos de tumores. (Se aplica más a tumores infantiles).

Cirugía de las recidivas tumorales: se realizan cuando los ganglios están clínicamente afectados después de que el tumor primario se ha sometido a un tratamiento curativo con cirugía o radiaciones.

#### **1.4.2 Radioterapia**

El empleo de radiaciones ionizantes en la práctica clínica fue introducido al inicio de los años 30, cuando el radio se utilizó como fuente de energía letal para destruir el cáncer. En un principio, el tratamiento consistía en implantar el radio directamente en el seno del tumor o en una cavidad corporal, usando una fórmula especial para calcular la dosis tumoral. La cantidad de radio empleado, cuantificada en miligramos o gramos, se incrementó a la vez que las normas y procedimientos de protección al personal se perfeccionaron.

Otros isótopos artificiales fueron introducidos en los años 50, tales como cobalto 60 y cesio 137, los cuales permitieron incrementar la dosificación en tumores profundos y con menor absorción de las estructuras óseas próximas.

Con el progreso de la física, se dispuso de haces de radiación electromagnética de alta energía, generadas por pulsos de electrones

acelerados en generadores electrostáticos. Actualmente se utilizan aceleradores lineales de baja, media y alta energía y braquiterapia con implantes localizados de cobalto 60, cesio 137, yodo 125 iridio 192; en vez de las unidades de cobalto, ya que estas son más lentas y requieren recalibraciones frecuentes para compensar la pérdida de actividad radiactiva.

Los efectos secundarios de este tratamiento se manifiestan 2 o 3 semanas terminado el tratamiento y duran por años. Los principales son:

1. Dolores de cabeza constantes
2. Pérdida de sueño.
3. Dolor de huesos.
4. Fiebre.
5. Lesiones en la piel, sobretodo en la zona donde se recibió directamente el rayo.

#### **1.4.3 Quimioterapia**

La quimioterapia evolucionó en los años 40 con la aplicación terapéutica de la mostaza nitrogenada. Su concepto fue idéntico al de la quimioterapia antibacteriana, es decir, destruir las células cancerosas

como a las bacterias. La quimioterapia ha demostrado con el tiempo ser más efectiva en cánceres neoplásicos, mientras que los casos de cánceres más frecuentes ha sido menos eficaz.

La combinación de la cirugía, la radioterapia y la quimioterapia ha aumentado las posibilidades de vida en las últimas décadas. La quimioterapia adyuvante es el tratamiento adicional que se da tras haber extirpado por completo la enfermedad macroscópica en pacientes sin signos de residuos de tumor. El objetivo de dicho tratamiento es controlar las posibles micro metástasis residuales, locales o a distancia. La quimioterapia neoadyuvante es un concepto que se introdujo en los ochenta, y consiste en empezar el tratamiento en el periodo preoperatorio y postoperatorio para prevenir la diseminación y crecimiento de las células cancerosas.

Los efectos secundarios de la quimioterapia se enlistan a continuación:

1. Náuseas
2. Vómitos
3. Afecciones a la médula ósea, lo que causa anemia
4. Infertilidad

#### **1.4.4 Hormonoterapia**

A través del tiempo se ha observado que ciertas manipulaciones hormonales podían influir en forma favorable en el crecimiento neoplásico surgido en órgano dependientes de hormonas, tales como próstata, mama, ovario y tiroides. Las hormonas también pueden ejercer un resultado favorable en tumores que se originan en órganos independientes de hormonas, como las leucemias y los linfomas.

Hay dos maneras de alterar el fluido hormonal:

1. Hormoterapia ablativa: extirpación de órganos productores de hormonas o destrucción de su funcionamiento.
2. Hormonoterapia aditiva: mediante la administración de hormonas o sustancia similares.

#### **1.5 Estadísticas mundiales de la población con cáncer.**

El cáncer es una enfermedad que afecta a toda la población a nivel mundial, tanto niños como adultos, hombres y mujeres.

A continuación presentamos algunos dato a nivel de Europa y Estados Unidos

### 1.5.1 El cáncer en los Estados Unidos

Los cinco tipos de cáncer más comunes en la población femenina de los Estados Unidos. 1990 (los datos están dados en porcentajes)

TABLA II

#### CARCINOMAS MÁS COMUNES EN MUJERES. 1990

Localización del cancer	Porcentaje
Mama	31.12
Colon y recto	13.55
Pulmón	12.6
Utero	8.87
Ovario	4.36
Otros	29.5

Los cinco tipos de cáncer más comunes en la población masculina de los Estados Unidos. 1990



TABLA III  
CARCINOMAS MÁS COMUNES EN VARONES. 1990

Localizacion del cancer	Porcentaje
Prostata	28.9
Pulmon	16.3
Colon y recto	11.81
Vejiga	6.24
Linfoma no hodkin	3.95
Otros	32.8

Se puede concluir que los tipos más comunes de cáncer en los Estados Unidos varían según el sexo. Esto se debe a que los órganos más sensibles a esta enfermedad son los que caracterizan la sexualidad del individuo, como la próstata en los varones y las mamas en las mujeres.

#### 1.5.2 Cáncer maligno en la población europea. 1990

TABLA IV  
POBLACIÓN EUROPEA CON CÁNCER. 1990

Hombres		Mujeres	
Localización	Porcentaje	Localización	Porcentaje
Pulmón	20.70	Mamas	27.77

Colon	13.18	Colon	14.76
Próstata	12.30	Pulmón	5.60
Vejiga	7.39	Útero	5.13
Estómago	6.59	Estómago	5.02
Otros	39.86	Otros	41.72

En Europa, a diferencia de Estados Unidos, predomina el cáncer de pulmón en los hombres, seguido del de colon. En cambio en el sexo femenino, el cáncer de mama ocupa un lugar principal en ambas regiones.

### **1.6 El cáncer en el Ecuador**

En el Ecuador ha habido un incremento notorio de la población atendida por esta enfermedad. Desde el año 1985 hasta el año 1996, ha habido un incremento del 50.85% en la cantidad de varones atendidos en Solca de Guayaquil, mientras que el incremento del número de mujeres atendidas en esta institución, en el mismo periodo, es del 50.67%.

El cáncer es una enfermedad que se la puede curar si se la trata a tiempo. Los escasos recursos, la falta de preocupación y de información han sido factores que han ayudado a la poca o ninguna prevención por parte de los afectados.



Otro factor que influye mucho en el aumento de la mortalidad por este mal es la ubicación geográfica de los centros con recursos para diagnosticarlo y tratarlo, ya que éstos son localizados mayormente en las ciudades principales del país. El presente trabajo se basa en la información que nos han proporcionado los mismos. Estos centros son:

En Quito, contamos con el apoyo de la Cruz Roja y La Sociedad de Lucha contra el Cáncer de Quito. Solicitamos ayuda al hospital de niños Vaca Ortiz, lastimosamente se nos prohibió el acceso a las historias clínicas de sus pacientes.

En Guayaquil, contamos con el apoyo de La Sociedad de Lucha contra el Cáncer de Guayaquil.

#### **1.6.1 Datos estadísticos de la población en Quito y Guayaquil**

A continuación se presenta la información de los carcinomas que más afectan a los ecuatorianos de las principales ciudades en porcentajes de la población total con cáncer, agrupados por ciudades y sexo

TABLA V

## POBLACIÓN ECUATORIANA CON CANCER EN PORCENTAJES. 1985-1996

Hombres			
Quito		Guayaquil	
Estómago	17	Estómago	13
Próstata	14	Piel	10
Piel	13	Próstata	8
Leucemia	7	Linfomas	7
Pulmón	5	Leucemia	7
Linfomas	3	Pulmón	7
Otros	41	Otros	48
Mujeres			
Quito		Guayaquil	
Cervix	23	Cervix	49
Mamas	12	Mamas	14
Piel	10	Piel	6
Estómago	10	Estómago	4
Leucemia	4	Leucemia	3
Vesícula	4	Ovario	3
Pulmón	2	Linfomas	2
Otros	35	Otros	19

El carcinoma más común en los hombres ecuatorianos es el de estómago, seguido por el de próstata y piel. El cáncer de estómago puede ser evitado con costumbres alimenticias adecuadas, según el estudio realizado por John Higginson y Sir Richard Doll, característica que casi nadie toma en cuenta en nuestro país al momento de alimentarse, ya que no existe apropiada información y educación nutricional en la población.

En las mujeres ecuatorianas se nota un gran porcentaje de cáncer uterino y mamario, siendo el primero el que cobra más vidas. La causa fundamental que no permite evitar estas muertes es la falta de información y de revisiones periódicas preventivas, tales como papanicolaus y mamografías.

A simple vista, se puede observar que hay algunas diferencias entre Quito y Guayaquil, con respecto a los tipos de carcinomas más comunes. El porcentaje de casos de cáncer de cervix es mucho mayor en Guayaquil que en Quito, al contrario el porcentaje de casos de cáncer de estómago y de piel femenino es mayor en Quito que en Guayaquil.

Todos estos aspectos observados nos lleva a hacernos la pregunta, ¿influye la región donde se viva el tipo de cáncer que puede afectar a un individuo?

Para contestar esta interrogante, procedemos a plantear la hipótesis nula: "El tipo de cáncer que tiene el individuo es independiente de la ciudad donde vive"

#### Versus

Hipótesis alternativa: "El tipo de cáncer que tiene el individuo es dependiente de la ciudad donde vive"

Para realizar esta prueba de hipótesis,

1. Hallamos una muestra de la población constituida por todas las personas con cáncer de Quito y Guayaquil, las cuales suman un total de 35116 casos recolectados desde 1985 hasta 1996 en las instituciones antes mencionadas, repartidos de la siguiente manera:

TABLA VI

## POBLACIÓN ECUATORIANA CON CANCER. 1985-1996

Mujeres			
Quito		Guayaquil	
Piel	1150	Piel	604
Pulmón	230	Linfomas	235
Mamas	1380	Mamas	1522
Estómago	1150	Estómago	392
Leucemia	460	Leucemia	352
Vesícula	460	Ovario	329
Cervix	2645	Cervix	5335
Otros	4024	Otros	2073
Total	11498	Total	10862
Hombres			
Piel	965	Piel	533
Linfomas	223	Linfomas	355
Pulmón	371	Pulmón	348
Estómago	1262	Estómago	710
Leucemia	520	Leucemia	372
Próstata	1040	Próstata	453

Otros	3044	Otros	2561
Total	7425	Total	5331

2. El tamaño de la muestra es obtenido de las fórmulas:

$$\hat{\sigma}_p^2 = \frac{(1-f) * p * (1-p)}{n-1} \quad \text{Fórmula 1.1}$$

$$\ell = Z_{\alpha/2} \hat{\sigma}_p \quad \text{Fórmula 1.2}$$

$$\ell^2 = Z_{\alpha/2}^2 \hat{\sigma}_p^2 \quad \text{Fórmula 1.3}$$

Donde n es el tamaño de la muestra, f es la división de n sobre N (N es el tamaño de la población). P es la proporción de individuos que tienen una característica específica (que vivan en una ciudad, que tengan un cáncer específico, etc.); y,

$\sigma_p^2$  : es el estimador de la varianza del estimador de la proporción p.

$\ell$  : es el error asociado al cálculo.

$Z_{\alpha/2}^2$  : es el valor del estadístico normal al cuadrado, con un  $\alpha/2$  como nivel de significancia



Al despejar la fórmula 1.3 obtenemos que  $n$  es igual a:

$$n = \frac{p^*(1-p) + \frac{\ell^2}{Z_{\alpha/2}^2}}{\frac{\ell^2}{Z_{\alpha/2}^2} + \frac{p^*(1-p)}{N}} \quad \text{Fórmula 1.4}$$

Trabajamos a un nivel de significancia del 0.05, con un error de 0.02, el  $N$  que es igual a 35116 y con la proporción que maximiza la varianza. Para hallar dicha proporción, encontramos las raíces de la derivada de 1.1 para  $p$  (proporción); de lo cual tenemos que la proporción que maximiza la varianza es igual a 0.5 obtuvimos un tamaño de muestra igual a 2000.

3. Procedemos a sacar una muestra aleatoria de la población original.

4. A continuación colocamos todos los valores de dicha muestra en las casillas de las tablas de contingencia que se muestran a continuación:



TABLA VII

TABLA DE FRECUENCIA DE TIPO DE CÁNCER VERSUS CIUDAD  
DONDE VIVEN LOS INDIVIDUOS EN ESTUDIO

	0	1	Total
1	130	114	244
2	279	228	507
3	93	78	171
4	50	33	83
5	51	38	89
6	37	29	66
7	45	49	94
8	405	341	746
Total	1090	910	2000

TABLA VIII

## VALORES ESPERADO DE LAS FRECUENCIAS

	0	1
1	132.98	111.02
2	276.315	230.685
3	93.195	77.805
4	45.235	37.76
5	48.505	40.496
6	35.97	30.03
7	51.23	42.77
8	406.57	339.43

TABLA IX

## DESVIACIONES ESTÁNDARES

	0	1
1	-0.258	0.283
2	0.162	-0.177
3	-0.02	0.022
4	0.708	-0.775

5	0.358	-0.392
6	0.172	-0.188
7	-0.87	0.953
8	-0.078	0.085

Los valores de las columnas están dados por:

0 representa a Quito

1 representa a Guayaquil

Los valores de las filas están dados por:

1 representa al cáncer estomacal

2 representa al cáncer en el aparato reproductor

3 representa al cáncer en la piel

4 representa la leucemia

5 representa al cáncer pulmonar

6 representa a los linfomas

7 representa al cáncer mamario

8 representa a los demás cánceres

5. Al utilizar el estadístico correspondiente a esta prueba (e.d.:  $\chi^2$  con una distribución ji cuadrada) con un nivel de significancia de 5 por ciento, obtuvimos los siguientes resultados:

	Valor	Grados de libertad	P
Estadístico de prueba $\chi^2$	3.333	7	0.8535

Al tener un valor p mayor a 0.05, no hay suficiente evidencia estadística para rechazar la hipótesis nula, lo que implica que la ciudad donde se viva no influye en el tipo de cáncer que los individuos bajo estudio padecen.

### 1.6.2 Los niños: un caso especial

Los datos recopilados de las instituciones colaboradoras, demuestran que la aparición del cáncer en los niños es poco frecuente. A continuación, se muestra una tabla con el porcentaje de niños tratados con relación al total de pacientes tratados en dichas instituciones, agrupados por sexo y por ciudad.

TABLA X  
POBLACIÓN INFANTIL CON CÁNCER EN ECUADOR POR PORCENTAJE  
SEGÚN CIUDAD. 1985-1996

Niños	
Guayaquil	Quito
4.05%	2.46%
Niñas	

Guayaquil

Quito

8.84%

5.24%

Los estudios recientes han dado indicio de que la mayoría de los carcinomas infantiles se desarrollan por influencias químicas, físicas o biológicas que actúan en la placenta materna de manera fortuita. Las anomalías cromosómicas pueden estar asociadas con una predisposición a padecer leucemias y otras malformaciones congénitas. La alteración de regiones específicas de diferentes cromosomas también se asocia con tumores como el retinoblastoma y el nefroblastoma. Otras causas asociadas con una alta incidencia de malignidad, particularmente con linfomas y leucemias, son los inmunotrastornos congénitos primarios.

Es justo destacar que los tipos de carcinomas que afectan a los niños, son diferentes a los que afectan a los adultos.

En Ecuador, los casos donde la muerte del niño ocurre en los primeros meses de tratamiento, es por un diagnóstico tardío de la enfermedad, a causa de ignorancia o poca prevención por parte de los padres, según lo muestran las historias clínicas que he estado investigando.

A continuación presentamos una lista de los porcentajes de carcinomas que afectan a los niños en los principales centros urbanos del Ecuador, clasificado por edades, tomadas desde el año 1985 a 1996.

TABLA XI

POBLACIÓN ECUATORIANA INFANTIL CON CÁNCER EN PORCENTAJE  
SEGÚN TIPO DE CÁNCER, EDAD Y CIUDAD. 1985-1996

	Guayaquil			Quito		
	0-4	5-9	10-14	0-4	5-9	10-14
Niños						
Leucemia	53	53	40	45	43	44
Ganglios linfáticos	21	26	30	10	25	19
Sistema nervioso	9	17	11	9	14	9
Ojo	11	1	0	9	0	0
Otros	5	3	20	27	18	28
Niñas						
Leucemia	57	52	51	45	47	44
Ganglios linfáticos	10	16	21	6	16	15
Sistema nervioso	10	14	9	13	13	0
Riñón	17	6	0	6	6	0
Otros	2	4	16	13	18	35
Ojo	4	8	4	17	0	6



Al tener la población infantil, cabe hacer las preguntas: ¿influye la edad en el tipo de cáncer que adquiere el niño?, ¿influye el sexo?, ¿qué podemos alegar sobre la relación entre la ciudad donde viven y el tipo de cáncer de los niños del Ecuador? Para contestar estas preguntas, tendremos que plantear diferentes hipótesis y hacer las respectivas pruebas:

Hipótesis nula: "El tipo de cáncer que tiene el niño es independiente de la edad"

Versus

Hipótesis alternativa: "El tipo de cáncer que tiene el niño es dependiente de la edad"

Para este propósito sacamos la muestra de la población que se encuentra descrita a continuación:



TABLA XII  
POBLACIÓN ECUATORIANA INFANTIL CON CÁNCER SEGÚN TIPO DE  
CÁNCER, EDAD Y CIUDAD. 1985-1996

	Guayaquil			Quito		
Niños						
Leucemia	95	85	40	46	42	37
Ganglios	38	42	30	10	24	16
Sistema nervioso	16	27	11	9	14	8
Ojo	20	2	0	9	0	0
Otros	9	5	20	28	17	24
Total	180	160	100	102	97	84
Niñas						
Leucemia	120	65	69	54	65	58
Ganglios	21	20	28	7	22	20
Sistema nervioso	21	18	12	16	18	0
Ojo	40	13	22	23	33	46
Otros	8	10	5	20	0	8
Total	211	125	135	120	138	131

Sacamos el tamaño de muestra por medio de la ecuación 1.3, con valores de error igual a 0.04, una proporción y un nivel de significancia del mismo valor que las pruebas anteriores y un tamaño de la población igual a 1583. El tamaño de muestra obtenido es de 437 individuos.

Al colocar los valores de esta muestra en las tablas de contingencia y hacer pruebas de hipótesis con un estadístico  $\chi^2$  al cinco por ciento de significancia, se obtiene:

TABLA XIII

TABLA DE FRECUENCIA DE TIPO DE CÁNCER VERSUS EDAD DE NIÑOS EN ESTUDIO

	1	2	3	Total
1	97	82	52	231
2	23	32	23	78
3	13	29	10	52
4	28	7	3	38
5	15	8	15	38
Total	176	158	103	437

TABLA XIV

## VALORES ESPERADO DE LAS FRECUENCIAS

	1	2	3
1	93.03	83.52	54.446
2	31.41	28.2	18.384
3	20.94	18.8	12.256
4	15.3	13.74	8.957
5	15.3	13.74	8.957

TABLA XV

## DESVIACIONES ESTÁNDARES

	1	2	3
1	0.41	-0.17	-0.33
2	-1.5	0.715	1.08
3	-1.74	2.352	-0.64
4	3.25	-1.82	-1.99
5	-0.08	-1.55	2.02

Las columnas se representan de la siguiente manera:

1: edad entre 0 y 4 años

2: edad entre 4 y 9 años

3: edad entre 10 y 14 años

Las filas se representan de la siguiente manera:

1: leucemia

2: cáncer en los ganglios linfáticos

3: cáncer en el sistema nervioso

4: cáncer en el ojo

5: otros cánceres.

	Valor	Grados de libertad	P
Estadístico de prueba $\chi^2$	37.471	8	0.000

Al tener un valor p menor a 0.05, rechazamos la hipótesis nula, lo que implica que hay evidencia estadística de que la edad influye en el tipo de cáncer que los niños ecuatorianos tienen.

Hipótesis nula: "El tipo de cáncer que tiene el niño es independiente de la ciudad de donde procede"

Versus

Hipótesis alternativa: "El tipo de cáncer que tiene el niño es dependiente de la ciudad de donde procede"

Al utilizar el mismo tamaño de muestra, tablas de contingencia y hacer pruebas de hipótesis con un estadístico  $\chi^2$  con diez por ciento de significancia, se obtiene:

TABLA XVI

TABLA DE FRECUENCIA DE TIPO DE CÁNCER VERSUS CIUDAD DE DONDE PROCEDEN LOS NIÑOS EN ESTUDIO

	0	1	Total
1	73	158	231
2	33	45	78
3	20	32	52
4	15	23	38
5	25	13	38
Total	166	271	437

TABLA XVII

## VALORES ESPERADO DE LAS FRECUENCIAS

	0	1
1	87.7	143.25
2	29.6	48.371
3	19.8	32.247
4	14.4	23.565
5	14.4	23.565

TABLA XVIII

## DESVIACIONES ESTÁNDARES

	0	1
1	-1.574	1.23
2	0.619	-0.49
3	0.056	-0.04
4	0.149	-0.12
5	2.781	-2.18

Las columnas se representan por

1: Guayaquil



0: Quito

Las filas se representan por

1: leucemia

2: cáncer en los ganglios linfáticos

3: cáncer en el sistema nervioso

4: cáncer en el ojo

5: otros carcinomas.

	Valor	Grados de libertad	P
Estadístico de prueba $\chi^2$	17.126	4	0.002

Al tener un valor p menor a mi nivel de significancia (0.05), rechazamos la hipótesis nula, lo que implica que hay evidencia estadística de que la ciudad donde se viva influye en el tipo de cáncer que los niños en estudio tienen.

Hipótesis nula: "El tipo de cáncer que tiene el niño es independiente del sexo"

Versus

Hipótesis alternativa: "El tipo de cáncer que tiene el niño es dependiente del sexo"

Al utilizar el mismo tamaño de muestra que en las pruebas anteriores, las tablas de contingencia y hacer pruebas de hipótesis con un estadístico  $\chi^2$ , al mismo porcentaje de significancia, se obtiene:

TABLA XIX

TABLA DE FRECUENCIA DE TIPO DE CÁNCER VERSUS SEXO DEL NIÑO EN ESTUDIO

	0	1	Total
1	120	111	231
2	39	39	78
3	24	28	52
4	30	8	38
5	11	27	38
Total	224	213	437

TABLA XX

VALORES ESPERADO DE LAS FRECUENCIAS

	0	1
1	118	112.59
2	40	38.018
3	26.7	25.346

4	19.5	18.522
5	19.5	18.522

TABLA XXI

## DESVIACIONES ESTÁNDARES

	0	1
1	0.146	-0.15
2	-0.155	0.159
3	-0.514	0.527
4	2.384	-2.45
5	-1.921	1.97

	Valor	Grados de libertad	P
Estadístico de prueba $\chi^2$	19.868	4	0.001

Al tener un valor p menor a 0.05, rechazamos la hipótesis nula, lo que implica que hay evidencia estadística de que el sexo influye en el tipo de cáncer que los niños ecuatorianos tienen.

## **CAPÍTULO II**

# **ESTADÍSTICA DESCRIPTIVA**

## **Capítulo 2**

### **2. ESTADÍSTICA DESCRIPTIVA**

En base a la información obtenida en una de las instituciones colaboradoras se formó una base de datos de niños con cáncer que han sido atendidos en La Sociedad de Lucha contra el Cáncer en Guayaquil desde 1985 hasta 1994. Dicha base de datos forma parte del Anexo 1.

Por medio de la información recopilada se puede hacer un estudio de frecuencias de las variables importantes.

2.1 Sexo

Esta variable puede tomar dos valores Masculino y Femenino. La tabla de frecuencia de esta variable se muestra a continuación

TABLA XXII

TABLA DE FRECUENCIA DE LA VARIABLE SEXO

	Frecuencia	Frecuencia relativa	Frecuencia relativa Acumulada
Femenino	263	42.4	42.4
Masculino	357	57.6	100
Total	620	100	

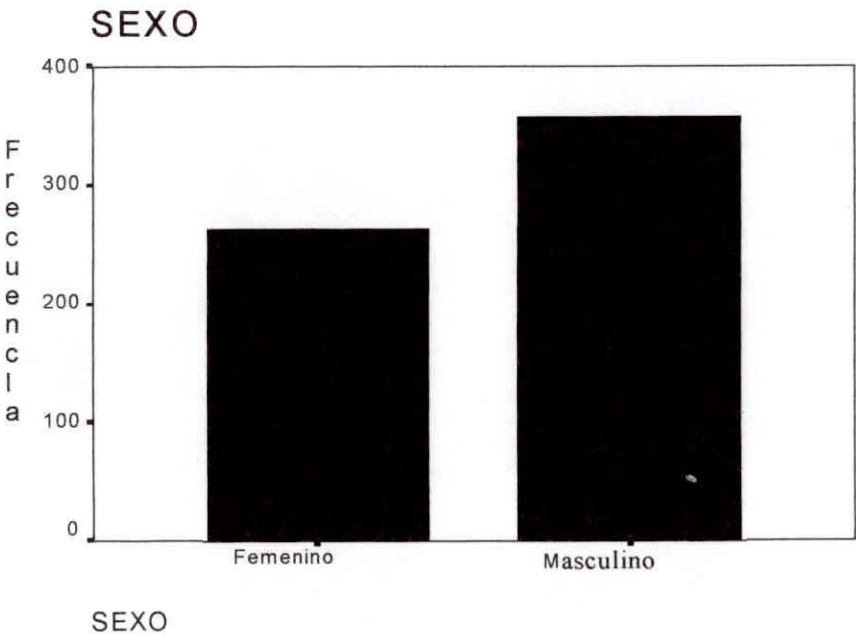


Gráfico 2.1



Del total de la población de 620 infantes, el 57.6 % son niños, mientras que el resto son niñas.

## 2.2 Tipo De Cancer

Esta variable puede tomar 13 valores que corresponden a 13 lugares diferentes en donde se han desarrollado células cancerígenas en los niños de estudio o a un tipo de carcinoma particular por su forma y características. Esta división es:

HIG: células cancerígenas o tumor maligno en el hígado.

HUE: células cancerígenas o tumor maligno en los huesos.

LDH: linfoma de hodkin.

LLA: leucemia linfoide aguda

LMA: leucemia mieloide aguda.

LNH: linfoma no hodkin.

OJO: células cancerígenas o tumor maligno en el ojo.

OT: otros

PBL: células cancerígenas o tumor maligno en las partes blandas.

RIÑ: células cancerígenas o tumor maligno en los riñones.

SN: células cancerígenas o tumor maligno en el sistema nervioso.

VAR: células cancerígenas o tumor maligno en varios órganos del cuerpo.

TABLA XXIII

TABLA DE FRECUENCIAS DE LA VARIABLE TIPO DE CÁNCER

	Frecuencia	Frecuencia relativa	Frecuencia relativa Acumulada
HIG	3	0.5	0.5
HUE	29	4.7	5.2
LDH	70	11.3	16.5
LLA	329	53.1	69.5
LMA	13	2.1	71.6
LNH	8	1.3	72.9
OJO	43	6.9	79.8
OT	9	1.5	81.3
PBL	22	3.5	84.8
RIÑ	38	6.1	91
SN	54	8.7	99.7
VAR	2	0.3	100
Total	620	100	

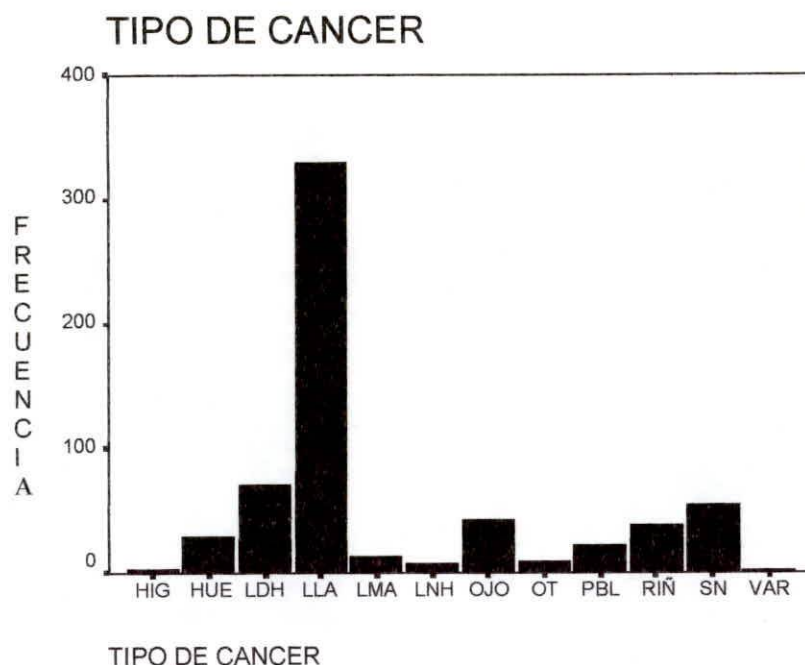


Gráfico 2.2

Más de la mitad de la población de niños sufre de leucemia linfocítica aguda.  
(329 casos de 620)

El segundo tipo de cáncer que más afecta a estos niños es el linfoma de hodkin (70 casos de 620), seguido de los carcinomas en el sistema nervioso (54 casos de 620) y en el ojo (43 casos de 620).

### 2.3 Estado

Esta variable describe el estado actual en que se encuentran los niños bajo estudio. Su codificación se describe a continuación:

?: dejó el tratamiento en Solca y no se sabe su estado actual

B: se encuentra en buenas condiciones hasta al actualidad.

F: fallecido

SE: dejó el tratamiento sin evidencia de actividad tumoral.

TABLA XXIV

TABLA DE FRECUENCIA DE LA VARIABLE ESTADO

	Frecuencia	Frecuencia relativa	Frecuencia relativa Acumulada
?	295	47.6	47.6
B	89	14.4	61.9
F	223	36	97.9
SE	13	2.1	100
Total	620	100	

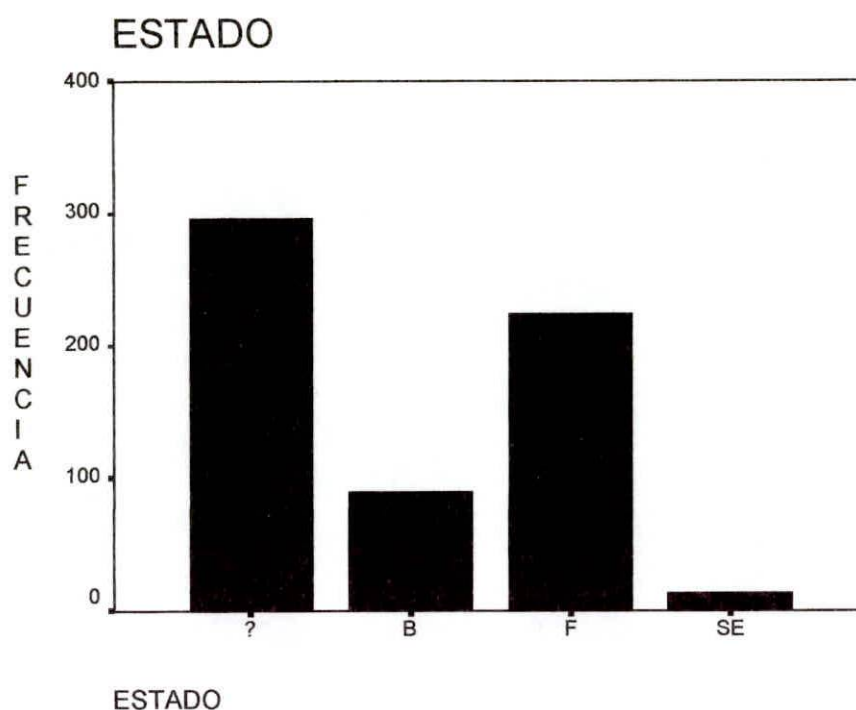


Gráfico 2.3

Hay una alta incidencia de abandono en el tratamiento(Del 49.7%)

De los niños que continuaron el tratamiento, el 71% de éstos han fallecido, mientras que el 29% de ellos están todavía con vida.

Se nota la marcada ineficiencia de la recopilación de datos en esta variable, puesto que sólo del 0.04% de los niños que dejaron el tratamiento se sabe su estado. Además no existe ningún tipo de seguimiento para saber el porqué ni el estado de los niños que han abandonado el tratamiento.

## 2.4 Edad Real

Esta variable describe la edad en que se diagnostica el carcinoma en el paciente.

TABLA XXV

TABLA DE FRECUENCIA DE LA VARIABLE EDAD REAL

	Frecuencia	Frecuencia relativa	Frecuencia relativa Acumulada
1	43	6.9	7.2
2	75	12.1	19.8
3	78	12.6	32.8
4	75	12.1	45.4
5	39	6.3	51.9
6	44	7.1	59.3
7	37	6	65.5
8	39	6.3	72
9	25	4	76.2
10	30	4.8	81.2
11	36	5.8	87.3
12	35	5.6	93.1
13	22	3.5	96.8
14	19	3.1	100



Total	597	96.3	
Sin edad	23	3.7	
	620	100	

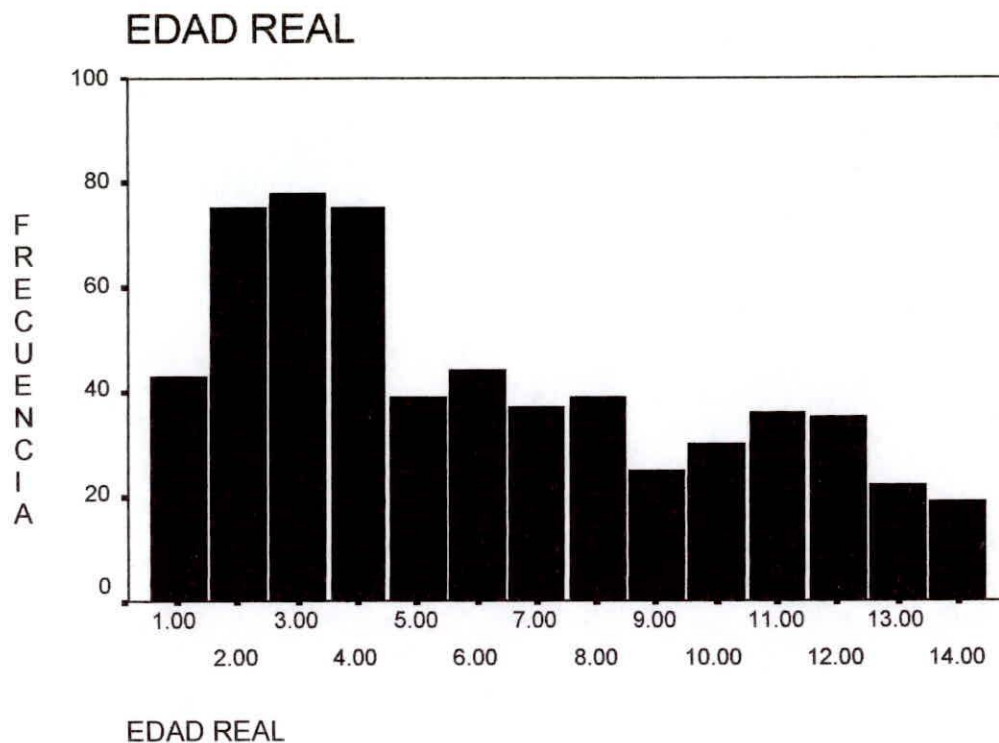


Gráfico 2.4

## 2.5 Provincia De Nacimiento

A continuación se detalla las provincias de nacimiento de los niños tratados en Solca de Guayaquil durante el periodo desde el año 1985 hasta 1994

TABLA XXVI

## TABLA DE FRECUENCIA DE LA VARIABLE PROVINCIA DE NACIMIENTO

	Frecuencia	Frecuencia relativa	Frecuencia relativa Acumulada
No Se Sabe	16	2.6	2.6
Azuay	6	1	3.5
Bolivar	5	0.8	4.4
Cañar	1	0.2	4.5
Chimborazo	4	0.6	5.2
Cotopaxi	1	0.2	5.3
El Oro	29	4.7	10
Esmeraldas	15	2.4	12.4
Guayas	406	65.5	77.9
Loja	4	0.6	78.5
Los Rios	43	6.9	85.5
Manabi	79	12.7	98.2
Morona	1	0.2	98.4
Napo	1	0.2	98.5
Oriente	1	0.2	98.7
Pichincha	8	1.3	100
Total	620	100	

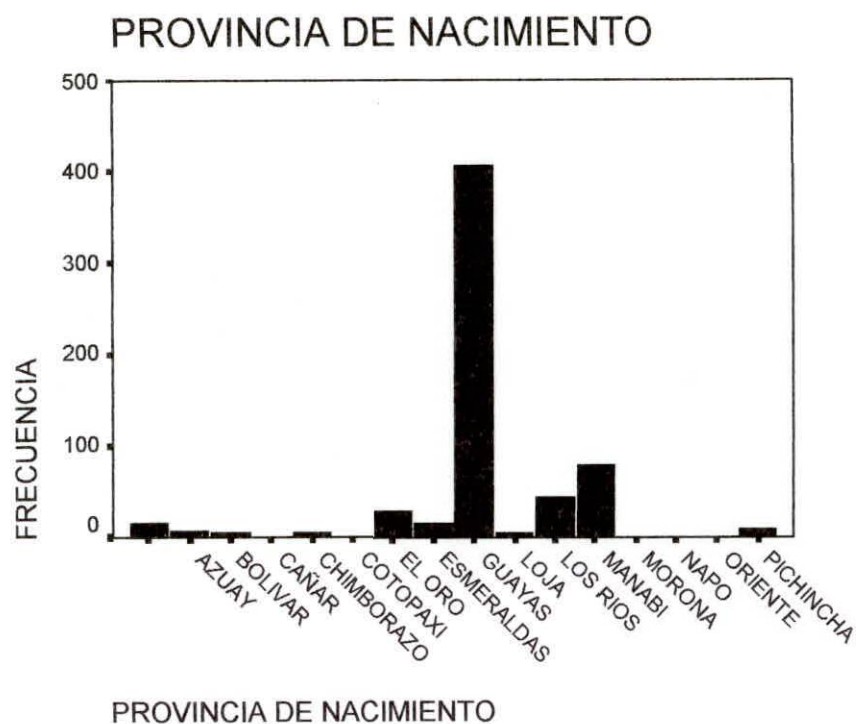


Gráfico 2.5

Hay una marcada diferencia entre las provincias donde nacieron los pacientes, más de la mitad de la población nace en la provincia del Guayas (65.5%), seguida por la provincia de Manabí con un 6.9% y la provincia de Los Ríos con un 4.7%.

## 2.6 Provincia Donde Vive

A continuación se detalla las provincias donde viven los niños tratados en Solca de Guayaquil durante el periodo desde el año 1985 hasta 1994

TABLA XXVII

## TABLA DE FRECUENCIA DE LA VARIABLE PROVINCIA DONDE VIVE

	Frecuencia	Porcentaje	Porcentaje acumulado
Otros	20	3.4	3.4
Azuay	5	0.8	4.2
Bolivar	6	1	5.2
Chimborazo	2	0.3	5.5
Cotopaxi	1	0.2	5.6
El Oro	24	3.9	9.5
Esmeraldas	12	1.9	11.5
Guayas	427	68.9	80.3
Loja	5	0.8	81.1
Los Rios	48	7.7	88.9
Manabi	62	10	98.9
Morona	1	0.2	99
Napo	1	0.2	99.2
Oriente	1	0.2	99.4
Pichincha	4	0.6	100
Total	620	100	

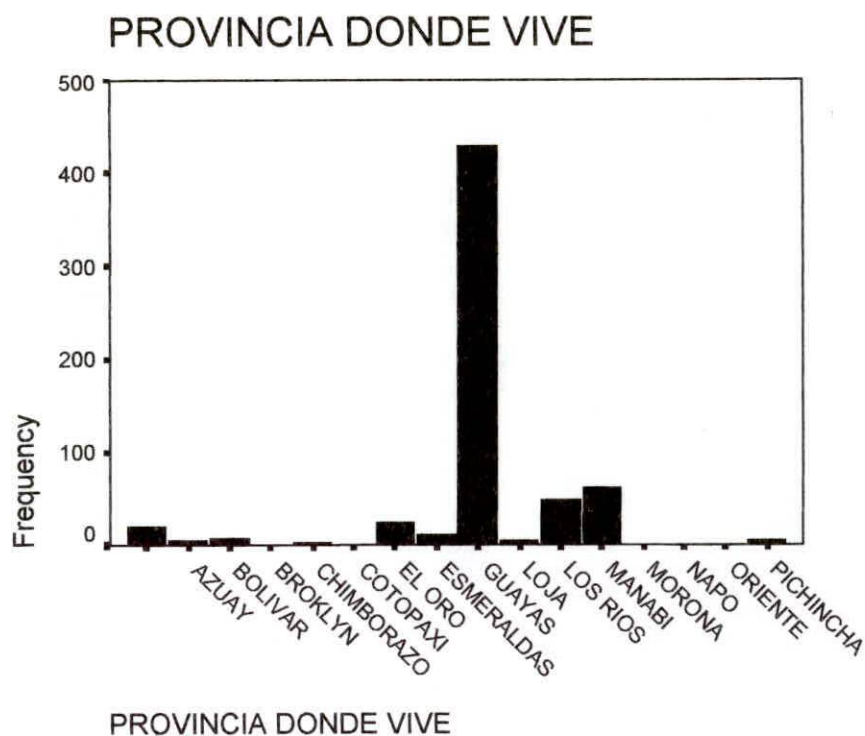


Gráfico 2.6

Hay una marcada diferencia entre las provincias donde nacieron los pacientes, más de la mitad de la población nace en la provincia del Guayas (68.9%), seguida por la provincia de Manabí con un 10% y la provincia de Los Ríos con un 7.7%.

## 2.7 Tiempo De Salida

Se define el tiempo de salida como la variable que mide el tiempo en años que le tomó al paciente salir del tratamiento dado por Solca de Guayaquil, no importa la condición de su salida.

TABLA XXVIII

## TABLA DE FRECUENCIA DE LA VARIABLE TIEMPO DE SALIDA

Tiempo de salida	Frecuencia	Porcentaje	Porcentaje acumulado
0	344	55.5	58.8
1	63	10.2	69.6
2	39	6.3	76.2
3	29	4.7	81.2
4	15	2.4	83.8
5	42	6.8	90.9
6	13	2.1	93.2
7	11	1.8	95
8	7	1.1	96.2
9	10	1.6	97.9
10	8	1.3	99.3
11	3	0.5	99.8
12	1	0.2	100
Total	585	94.4	
No encontrados	35	5.6	
	620	100	



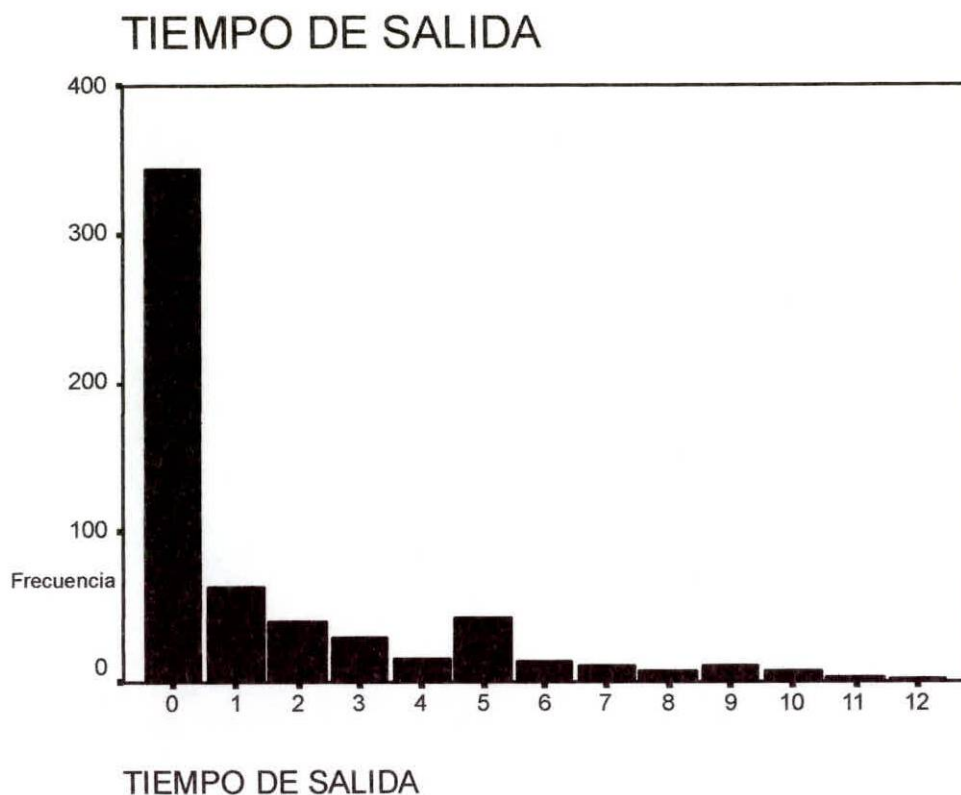


Gráfico 2.7

## **2.8 Estadística descriptiva de la variable tiempo de supervivencia clasificada por el tipo de tratamiento recibido**

En el siguiente recuadro se analizará las estadísticas descriptivas de los tiempos de supervivencia de los niños de la base de datos de las instituciones de Quito y Guayaquil colaboradoras, clasificada por las diferentes combinaciones de los tratamientos que recibieron.

TABLA XXIX

TABLA DE ESTADÍSTICA DESCRIPTIVA DE VARIABLE TIEMPO DE SUPERVIVENCIA SEPARADAS POR TIPOS DE TRATAMIENTOS

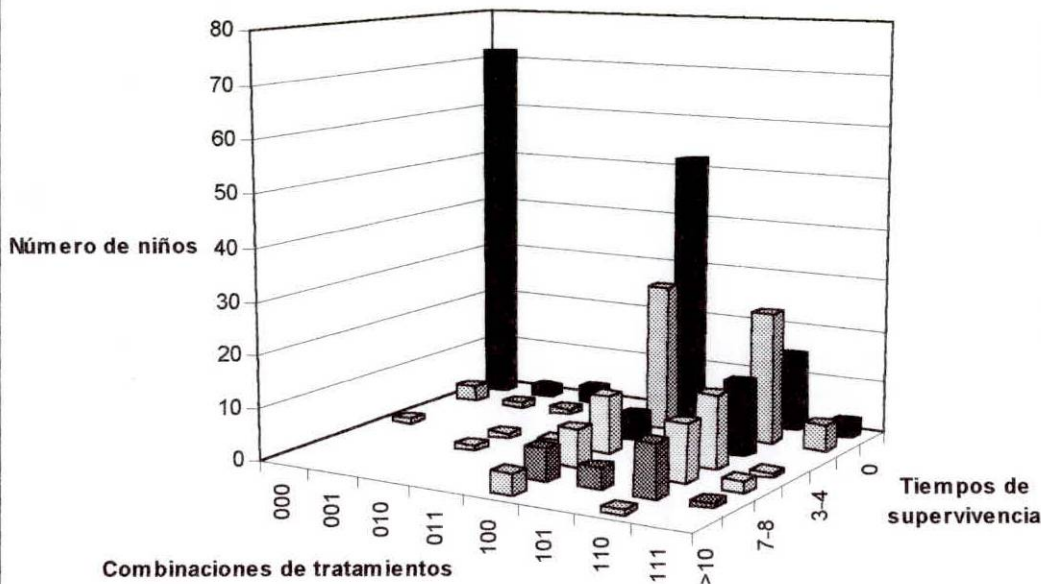
	000		100		010		001	
Media	0.11		2.45		3.91		3.91	
Intervalo de confianza del 95% para la media	-0.03	,0.24	1.81	, 3.09	-12.9	, 20.71	-12.9	, 20.71
Mediana	0		1		0		0	
Varianza	0.36		11.86		45.77		45.77	
Desviación estándar	0.6		3.44		6.77		6.77	
Mínimo	0		0		0		0	
Máximo	5		13.39		11.72		11.72	
Rango	5		13.39		11.72		11.72	
Sesgo	7.5		1.42		1.73		1.73	
Kurtosis	60.32		0.97		-1.21*		*	

Número de individuos	76		113		6		3	
----------------------	----	--	-----	--	---	--	---	--

	101		110		110		111	
Media	5.5		3.88		3.5		3.08	
Intervalo de confianza del 95% para la media	1.79	, 9.21	3.21	4.55	-28.27	35.27	0.79	, 5.37
Mediana	7.5		3		3.5		1	
Varianza	19.71		10.3		12.5		12.99	
Desviación estándar	4.44		3.21		3.54		3.6	
Mínimo	0		0		1		0	
Máximo	10		12.05		6		10	
Rango	10		12.05		5		10	
Sesgo	-0.44		0.55		0.81		0.95	
Kurtosis	-2.16*		-0.83*		3.09		-0.74*	
Número de individuos			8		90		12	

\* Este valor es un estimador del coeficiente de kurtosis. El intervalo de confianza de este coeficiente contiene el valor cero, por lo tanto se acepta la hipótesis que el coeficiente de kurtosis es 0.

Gráfico 2.8 Gráfico de barras de tiempos de supervivencia según el tipo de tratamientos



	000	001	010	011	100	101	110	111
>10					4		1	
9-10					6	4	10	1
7-8			1		7		11	2
5-6	1		1	1	11	1	14	1
3-4					5		14	
1-2	3	1	1	1	28	1	25	5

### **2.8.1 Niños que no recibieron tratamiento**

El 24.52% (76 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que no recibieron tratamiento alguno.

Más de la mitad de estos 76 niños sobrevivieron 0 días, y el tiempo de vida esperado de estos niños es de 0.11 años, es decir 1 mes y 7 días, y con una desviación de 7 meses y 6 días.

Mínimo viven 0 años y como máximo 5.

La distribución de este grupo de niños es asimétrica y con una picudez elevada, teniendo la mayor cantidad de datos a la izquierda de los valores de la variable tiempo de supervivencia, y eso se confirma al analizar los datos, ya que 72 niños sobrevivieron 0 años y 4 sobrevivieron más de 1 año, los cuales son considerados como valores aberrantes.

### **2.8.2 Niños que sólo recibieron quimioterapia**

El 36.45% (113 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que sólo recibieron quimioterapia.



Más de la mitad de estos 113 niños sobrevivieron 1 día, y el tiempo de vida esperado de estos niños es de 2.45 años, es decir 2 años 5 meses y 10 días, y con una desviación de 3 años 5 meses y 9 días.

Mínimo viven 0 años y como máximo 13 años y 20 días.

La distribución de este grupo de niños es asimétrica y con una picudez baja, teniendo la mayor cantidad de datos a la izquierda de los valores de la variable tiempo de supervivencia, y eso se confirma al analizar los datos, ya que 52 niños sobrevivieron 0 años, 17 niños sobrevivieron 1 año, 11 niños vivieron hasta el segundo año, 2 niños hasta el tercero, 3 hasta el cuarto, 4 niños sobrevivieron al quinto año, 7 niños al sexto año, 5 niños al séptimo, 2 niños al octavo y 2 al noveno año, el número de niños que sobrevivieron a los 10 años de tratamiento o más se los consideran como valores aberrantes, y son 8.

### **2.8.3 Niños que sólo recibieron radioterapia**

El 1.94% (6 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que sólo recibieron radioterapia.

El tiempo de vida esperado de estos niños es de 3.91 años, es decir 3 años 10 meses y 26 días, y con una desviación de 6 años 9 meses y 5 días.

Mínimo viven 0 años y como máximo 11 años 8 meses y 18 días.

La distribución de este grupo de niños es asimétrica sesgada a la izquierda y con una picudez baja, lo que implica que los datos están dispersos, y eso se confirma al analizar los datos, puesto que 3 niños sobrevivieron 0 años, 1 niño sobrevivió 2 años, 1 sobrevive 6 y el restante sobrevive hasta el año 8.

#### **2.8.4 Niños que sólo recibieron cirugía**

El 0.96% (3 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que sólo se les realizó cirugía.

El tiempo de vida esperado de estos niños es de 3.91 años, es decir 3 años 10 meses y 26 días, y con una desviación de 6 años 9 meses y 5 días.

Mínimo viven 0 años y como máximo 11 años 8 meses y 18 días.

La distribución de este grupo de niños es asimétrica sesgada a la izquierda y con una picudez baja, lo que implica que los datos están dispersos, y eso se confirma al analizar los datos, puesto que 2 niños sobrevivieron 0 años y el tiempo de supervivencia del otro es considerado un valor aberrante, puesto que es 12 años.

#### **2.8.5 Niños que recibieron quimioterapia y cirugía**

El 2.58% (8 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que sólo se les trató con quimioterapia y cirugía.

El tiempo de vida esperado de estos niños es de 5.5 años, es decir 5 años y 5 meses, y con una desviación de 4 años 5 meses y 8 días.

Mínimo viven 0 años y como máximo 10 años.

La distribución de este grupo de niños es asimétrica y sesgada levemente hacia la derecha y con una picudez baja, lo que implica que los datos están dispersos. Al analizar los datos notaremos que esto se cumple, ya que 2 niños no sobrevivieron al primer año, 1 no lo hizo al segundo, 1 al séptimo año; mientras que 3 niños han sobrevivido al noveno año de tratamiento y 1 al décimo.

### **2.8.6 Niños que recibieron quimioterapia y radioterapia**

El 29.03% (90 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que sólo se les trató con quimioterapia y radioterapia.

El tiempo de vida esperado de estos niños es de 3.88 años, es decir 3 años 10 meses y 16 días, y con una desviación de 3 años 2 meses y 15 días.

Mínimo viven 0 años y como máximo 12 años y 18 días.

La distribución de este grupo de niños es sesgada levemente hacia la izquierda y con una picudez baja, lo que implica que los datos están dispersos. Al analizar los datos notaremos que esto se cumple, ya que 25 niños sobrevivieron de 0 a 1 año, 25 lo hicieron hasta el año segundo y tercero, 13 al cuarto y quinto año, 10 al sexto y séptimo, 13 al octavo y noveno año, y 4 sobrevivieron más de 10 años.

### **2.8.7 Niños que recibieron radioterapia y cirugía**

El 0.65% (2 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que sólo se les trató con radioterapia y cirugía.



El tiempo de vida esperado de estos niños es de 3.5 años, es decir 3 años y 6 meses, con una desviación de 3 años 6 meses y 12 días.

Mínimo viven 1 año y como máximo 6 años.

Puesto que sólo se disponen de dos datos en este grupo, el coeficiente de kurtosis hace creer que se tiene la varianza de una variable normal estándar ( $\sigma^2 = 1$ ) y el sesgo da a entender que la población está ligeramente sesgada a la izquierda. Estas características no pueden atribuirse a la distribución de esta población puesto que se cuenta con pocos individuos, los cuales tienen tiempos de supervivencia de 1 y 6 años.

#### **2.8.8 Niños que recibieron los tres tratamientos juntos**

El 3.87% (12 niños) de la muestra de niños tratados entre 1985 y 1994, de los cuales se sabe su estado actual, pertenecen a los que se les trató con los tres tipos de tratamiento.

El tiempo de vida esperado de estos niños es de 3.08 años, es decir 3 años y 1 mes, y con una desviación de 3 años 7 meses y 7 días.

Mínimo viven 0 años y como máximo 10 años.

La distribución de este grupo de niños es sesgada levemente hacia la izquierda y con una picudez baja, lo que implica que los datos están dispersos. Al analizar los datos notaremos que esto se cumple, ya que 3 no sobreviven en el primer año de tratamiento, 4 no lo hacen al segundo año, 1 no sobrevive hacia el tercer año, 1 no sobrevive al sexto año, 1 no lo hace al séptimo, al igual que en el octavo y décimo año de tratamiento.



## **CAPÍTULO III**

# **ESTADÍSTICA INFERENCIAL**

## Capítulo 3

### 3. Estadística Inferencial

#### 3.1 Tablas de contingencia

Al ver las variables de una manera univariada es determinante saber la relación entre ellas y si tienen algún efecto entre sí.

A continuación presentamos el análisis bivariado de las variables más importantes.

Para este efecto, utilizaremos 3 pruebas:

1. Prueba  $\chi^2$

Mide la dependencia entre la variable con los valores en las columnas y la que tiene los valores en las filas. La hipótesis nula de esta prueba es que estadístico  $\chi^2$  es igual a cero, al contrario de la hipótesis alternativa.

## 2. Medidas direccionales

Una medida de asociación que refleja la reducción proporcional en el error al predecir la variable dependiente por medio de la variable independiente. El valor de este coeficiente se encuentra entre 0 y 1, los cuales significan una baja y alta relación de dependencia que tiene la variable dependiente con respecto a la independiente, respectivamente.

## 3. Medidas simétricas

Coeficiente de contingencia: es una medida de asociación de variables basada en el estadístico  $\chi^2$ . El valor de este coeficiente se encuentra entre 0 y 1, los cuales significan baja y alta asociación entre las variables respectivamente. La hipótesis nula de esta prueba es que el coeficiente de contingencia es igual a cero, al contrario de la hipótesis alternativa.

### 3.1.1 Guayaquil

#### 3.1.1.1 Lugar Donde Vive Vs. Tipo De Cancer

Queremos saber si el lugar donde se vive influye en el tipo de cáncer que tenga el infante.

La variable lugar donde vive toma dos valores: Guayaquil y No Guayaquil, lo que implica que la muestra se divide según el criterio si viven o no en Guayaquil.

La variable tipo de cáncer se divide en:

HUE: carcinoma a los huesos.

LDH: linfoma de hodkin.

LLA: leucemia linfoide aguda.

OJO: carcinoma de ojo.

PBL: carcinoma en las partes blandas.

RIÑ: carcinoma renal.

SN: carcinoma en el sistema nervioso.

OTROS: carcinoma en otras partes del cuerpo.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"El tipo de cáncer del infante es independiente del lugar donde vive"

vs.

"El tipo de cáncer del infante es dependiente del lugar donde vive"

A continuación la tabla de contingencia

TABLA XXX

TABLA DE CONTINGENCIA DE LAS VARIABLES LUGAR DONDE VIVE Y TIPO DE CÀNCER

	TIPO DE CANCER									
	Real	HUE	LDH	LLA	OJO	OT	PBL	RIÑ	SN	Total
Gquil	Esperado	9	27	165	18	18	15	18	22	292
No	Real	13.7	33	154.9	20.3	16.5	10.4	17.9	25.4	292
Gquil	Esperado	20	43	164	25	17	7	20	32	328
	Real	15.3	37	174.1	22.7	18.5	11.6	20.1	28.6	328
Total	Esperado	29	70	329	43	35	22	38	54	620

TABLA XXXI

PRUEBA  $\chi^2$  PARA VARIABLES LUGAR DONDE VIVE Y TIPO DE CÀNCER

Valor	Grados de libertad	Valor p
11.816	7	.107

Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que las variables son independientes.

TABLA XXXII

MEDIDAS DIRECCIONALES PARA VARIABLES LUGAR DONDE  
VIVE Y TIPO DE CÁNCER

		Valor	Error
Lambda	Simétrico	0.017	0.033
	Lugar donde vive dependiente	0.034	0.066
	Tipo de cancer dependiente	0.00	0.00
Goodman and	Lugar donde vive dependiente	0.019	0.011
Kruskal	Tipo de cancer dependiente	0.003	0.002

La variable LUGAR DONDE VIVE se puede predecir por medio de la variable TIPO DE CANCER con 1.9% a 3.4% de exactitud.

La variable TIPO DE CANCER se puede predecir por medio de la variable LUGAR DONDE VIVE con 0% a 0.3% de exactitud.

TABLA XXXIII

MEDIDAS SIMÉTRICAS PARA VARIABLES LUGAR DONDE VIVE Y  
TIPO DE CÁNCER

	Valor	Valor P
Coefficiente de contingencia	0.137	0.107



Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que el coeficiente de contingencia es estadísticamente nulo.

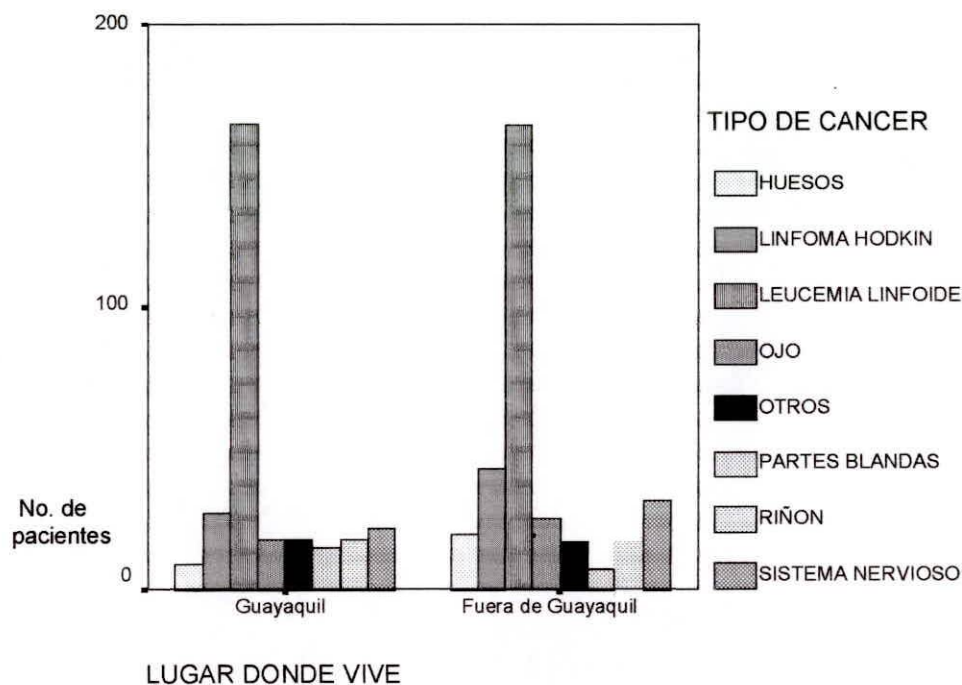


Gráfico 3.1

Las variables LUGAR DONDE VIVE y TIPO DE CANCER son independientes.

### 3.1.1.2 Estado vs. Edad categorizada

Queremos saber si la edad del infante influye en el estado en que se encuentra en los actuales momentos.

La variable edad categorizada se divide en 3 grupos:

1. De 0 a 4 años.
2. De 5 a 9 años.
3. De 10 a 14 años.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"La edad del infante es independiente del estado en que se encuentre"

vs.

"La edad del infante es dependiente del estado en que se encuentre"

A continuación la tabla de contingencia

TABLA XXXIV

TABLA DE CONTINGENCIA DE LAS VARIABLES ESTADO Y EDAD CATEGORIZADA

			Edad categorizada			
			1	2	3	Total
Estado	?	Real	129	88	79	296
		Esperado	140.4	87.8	67.8	296
	B	Real	41	31	17	89
		Esperado	42.2	26.4	20.4	89
	F	Real	124	65	46	235
		Esperado	111.4	69.7	53.8	235
Total		Real	294	184	142	620
		Esperado	294	184	142	620

TABLA XXXV

PRUEBA  $\chi^2$  PARA VARIABLES ESTADO Y EDAD CATEGORIZADA

Valor	Grados de libertad	Valor p
7.041	4	0.134

Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que las variables son independientes.

TABLA XXXVI

MEDIDAS DIRECCIONALES PARA VARIABLES ESTADO Y EDAD CATEGORIZADA

		Valor	Error
Lambda	Simétrico	0.00	0.00
	ESTADO Dependiente	0.00	0.00
	EDAD Dependiente	0.00	0.00
Goodman and Kruskal tau	ESTADO Dependiente	0.007	0.005
	EDAD Dependiente	0.006	0.004

La variable ESTADO se puede predecir por medio de la variable EDAD con 0% a 0.7% de exactitud.

La variable EDAD se puede predecir por medio de la variable ESTADO con 0% a 0.3% de exactitud.

TABLA XXXVII

MEDIDAS SIMÉTRICAS PARA VARIABLES ESTADO Y EDAD CATEGORIZADA

	Valor	Valor p.
Coefficiente de contingencia	0.106	0.134

Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que el coeficiente de contingencia es estadísticamente nulo.

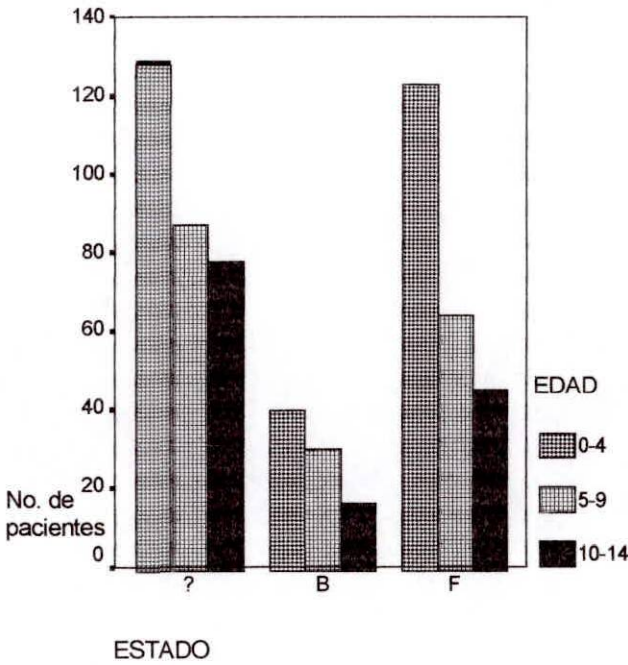


Gráfico 3.2

Las variables EDAD y ESTADO son independientes.

### 3.1.1.3 Estado vs. Sexo

Queremos saber si el sexo del infante influye en el estado en que se encuentra en los actuales momentos.

La variable sexo se divide en 2 grupo, femenino (F) y masculino (M).

La variable estado se divide en:

?: Estado desconocido.

F: Fallecido.

B: Bien.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"El sexo del infante es independiente del estado en que se encuentre"

vs.

"El sexo del infante es dependiente del estado en que se encuentre"



A continuación la tabla de contingencia

TABLA XXXVIII

TABLA DE CONTINGENCIA DE LAS VARIABLES ESTADO Y SEXO

			SEXO		Total
			F	M	
ESTADO	?	REAL	116	180	296
		ESPERADO	125.6	170.4	296
	B	REAL	40	49	89
		ESPERADO	37.8	51.2	89
	F	REAL	107	128	235
		ESPERADO	99.7	135.3	235
Total		REAL	263	357	620
		ESPERADO	263	357	620

TABLA XXXIX

PRUEBA  $\chi^2$  PARA VARIABLES ESTADO Y SEXO

Valor	Grados de libertad	Valor p
2.429	2	0.297

Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que las variables son independientes.

TABLA XL

MEDIDAS DIRECCIONALES PARA VARIABLES ESTADO Y EDAD CATEGORIZADA

		Valor	Error
Lambda	Simétrico	0.00	0.00
	Estado dependiente	0.00	0.00
	Sexo dependiente	0.00	0.00
Goodman and	Estado dependiente	0.003	0.003
Kruskal tau	Sexo dependiente	0.004	0.005

La variable ESTADO se puede predecir por medio de la variable EDAD con 0% a 0.7% de exactitud.

La variable EDAD se puede predecir por medio de la variable ESTADO con 0% a 0.3% de exactitud.

"La edad del infante es dependiente del tiempo que le tomó para salir de la institución"

A continuación la tabla de contingencia

TABLA XLVI

TABLA DE CONTINGENCIA DE LAS VARIABLES EDAD CATEGORIZADA E INTERVALOS DE SALIDA

		Edad categorizada				
			1	2	3	Total
Intervalos de Tiempos de Salida	0	Real	180	85	78	343
		Esperado	162.6	101.8	78.6	343
	1	Real	28	28	11	67
		Esperado	31.8	19.9	15.3	67
	2	Real	33	21	18	72
		Esperado	34.1	21.4	16.5	72
	3	Real	7	11	10	28
		Esperado	13.3	8.3	6.4	28
	4	Real	46	39	25	110
		Esperado	52.2	32.6	25.2	110
		Real	294	184	142	620
	Total	Esperado	294	184	142	620

Las variables ESTADO y TIPO DE CANCER son dependientes.

### **3.1.1. 5 Tiempo de salida vs. Edad categorizada**

Queremos saber si la edad del infante influye en el tiempo que le tomó salir de la institución.

La variable edad se divide en 3 grupos como antes mencionamos y el tiempo de salida se lo ha agrupado de la siguiente manera:

1. Tiempo de salida menor a 1 año, valor =0.
2. Tiempo de salida entre 1 año y 3 años.
3. Tiempo de salida entre 4 años y 6 años.
4. Tiempo de salida entre 7 años y 9 años.
5. Tiempo de salida mayor o igual a 10 años.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"La edad del infante es independiente del tiempo que le tomó para salir de la institución"

vs.

TABLA XLV

MEDIDAS SIMÉTRICAS PARA VARIABLES ESTADO Y TIPO DE CÁNCER

	Valor	Valor p.
Coeficiente de contingencia	0.359	0.000

Ya que trabajamos a un nivel de significancia del 95%, se rechaza a la hipótesis nula, lo que implica que el coeficiente de contingencia no es estadísticamente nulo y existe una asociación considerable entre estas dos variables.

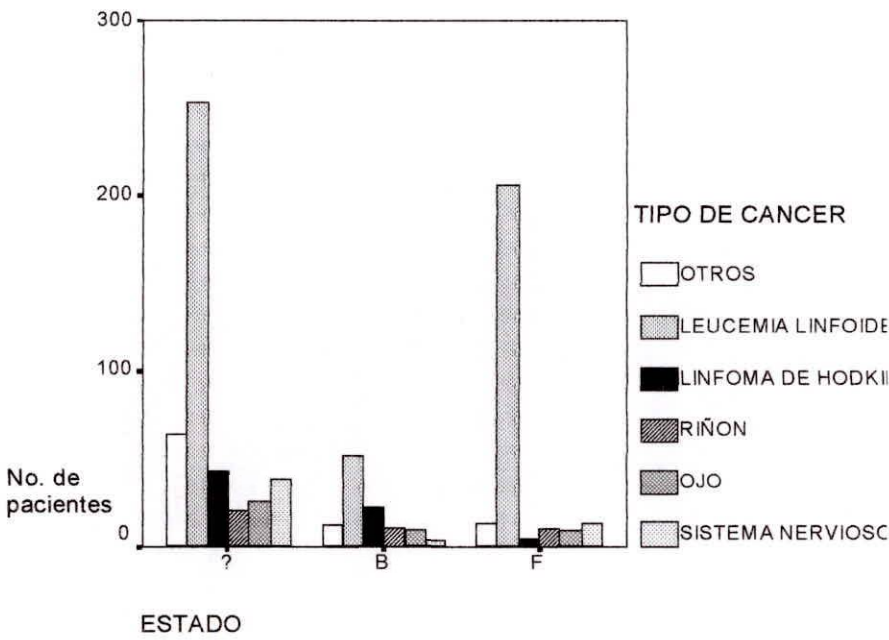


Gráfico 3.4

Ya que trabajamos a un nivel de significancia del 95%, se rechaza la hipótesis nula, lo que implica que las variables son dependientes.

TABLA XLIV

MEDIDAS DIRECCIONALES PARA VARIABLES ESTADO Y TIPO DE CÁNCER

		Valor	Error
Lambda	Simétrico	0.00	0.00
	Estado dependiente	.080	.027
	Tipo de cancer dependiente	0.151	0.049
Goodman and	Estado dependiente	0.000	0.000
Kruskal tau	Tipo de cancer dependiente	0.081	0.016

La variable ESTADO se puede predecir por medio de la variable TIPO DE CANCER con 0.8% a 15.1% de exactitud.

La variable TIPO DE CÁNCER se puede predecir por medio de la variable ESTADO con 0% a 0.48 de exactitud.



TABLA XLII

TABLA DE CONTINGENCIA DE LAS VARIABLES ESTADO Y TIPO DE CÁNCER

			Tipo de cancer								
			HUE	LDH	LLA	OJO	OT	PBL	RIÑ	SN	Tota
Estado	?	Real	20	39	123	24	23	17	16	34	296
		Esperado	13.8	33.4	157	20.5	16.7	10.5	18.1	25.8	296
	B	Real	5	22	34	9	3	4	9	3	89
		Esperado	4.2	10	47.2	6.2	5	3.2	5.5	7.8	89
	F	Real	4	9	172	10	9	1	13	17	235
		Esperado	11	26.5	125	16.3	13.3	8.3	14.4	20.5	235
Total		Real	29	70	329	43	35	22	38	54	620
		Esperado	29	70	329	43	35	22	38	54	620

TABLA XLIII

PRUEBA  $\chi^2$  PARA VARIABLES ESTADO Y TIPO DE CÁNCER

Valor	Grados de libertad	Valor p
91.507	14	0.000

#### 3.1.1.4 Estado vs. Tipo de cancer

Queremos saber si el tipo de cáncer que padece el infante influye en el estado en que se encuentra en los actuales momentos.

Las variables tipo de cáncer y estado toman los mismos valores de la pruebas previamente realizadas.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"El tipo de cáncer que padece el infante es independiente del estado en que se encuentre"

vs.

"El tipo de cáncer que padece el infante es dependiente del estado en que se encuentre"

A continuación la tabla de contingencia

TABLA XLI

MEDIDAS SIMÉTRICAS PARA VARIABLES ESTADO Y SEXO

	Valor	Valor P
Coefficiente de contingencia	0.062	0.297

Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que el coeficiente de contingencia es estadísticamente nulo.

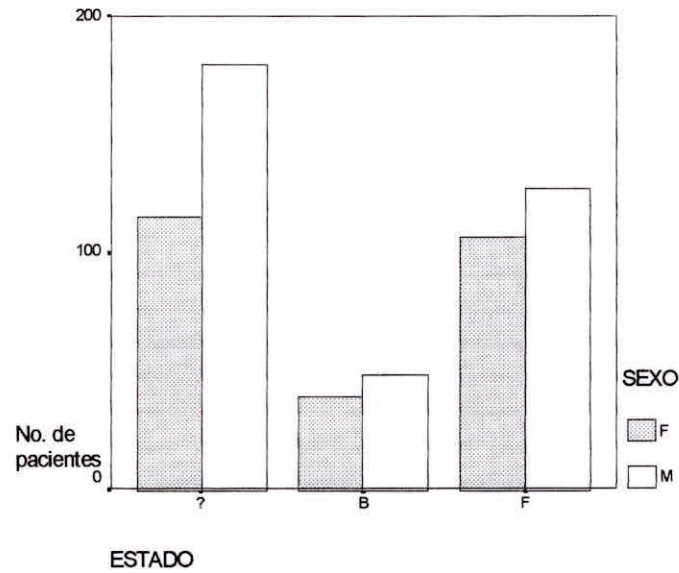


Gráfico 3.3

Las variables ESTADO y SEXO son independientes.

TABLA XLVII

PRUEBA  $\chi^2$  PARA LAS VARIABLES EDAD CATEGORIZADA E INTERVALOS DE SALIDA

Valor	Grados de libertad	Valor p
17.611	8	0.024

Ya que trabajamos a un nivel de significancia del 95%, se rechaza la hipótesis nula, lo que implica que las variables son dependientes.

TABLA XLVIII

MEDIDAS DIRECCIONALES PARA VARIABLES EDAD CATEGORIZADA E INTERVALOS DE SALIDA

		Valor	Error
Lambda	Simétrico	0.007	0.01
	Tiempo de salida dependiente	0.000	0.000
	Edad dependiente	0.012	0.01
Goodman and Kruskal tau	Tiempo de salida dependiente	0.009	0.01
	Edad dependiente	0.015	0.01

La variable TIEMPO DE SALIDA se puede predecir por medio de la variable EDAD con 0% a 0.9% de exactitud.

La variable EDAD se puede predecir por medio de la variable TIEMPO DE SALIDA con 1.2% a 1.5 de exactitud.

TABLA XLIX

MEDIDAS SIMÉTRICAS PARA VARIABLES EDAD CATEGORIZADA  
E INTERVALOS DE SALIDA

	Valor	Valor p.
Coefficiente de contingencia	0.166	0.024

Ya que trabajamos a un nivel de significancia del 95%, se rechaza a la hipótesis nula, lo que implica que el coeficiente de contingencia no es estadísticamente nulo y existe una asociación considerable entre estas dos variables.

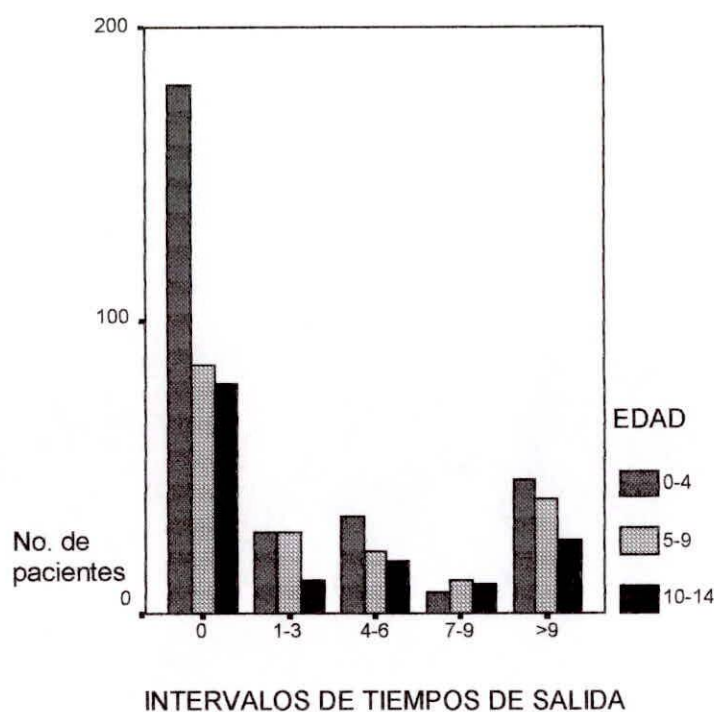


Gráfico 3.5

Las variables EDAD y TIEMPO DE SALIDA son dependientes.

### 3.1.1.6 Intervalos de tiempos de salida vs. Sexo

Queremos saber si el sexo del infante influye en el tiempo que le tomó salir de la institución.

La variable sexo toma dos valores, femenino y masculino; y el tiempo de salida se lo ha agrupado de la siguiente manera:



1. Tiempo de salida menor a 1 año, valor =0.
2. Tiempo de salida entre 1 año y 3 años, valor =1.
3. Tiempo de salida entre 4 años y 6 años, valor = 2.
4. Tiempo de salida entre 7 años y 9 años, valor =3.
5. Tiempo de salida mayor o igual a 10 años, valor =4.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"El sexo del infante es independiente del tiempo que le tomó para salir de la institución"

vs.

"El sexo del infante es dependiente del tiempo que le tomó para salir de la institución"

A continuación la tabla de contingencia

TABLA L

TABLA DE CONTINGENCIA DE LAS VARIABLES SEXO E INTERVALOS DE SALIDA

			Sexo		
			F	M	Total
Intervalos de tiempos De salida	0	Real	143	200	343
		Esperado	146	197.5	343
	1-3	Real	22	45	67
		Esperado	28.4	38.6	67
	4-6	Real	34	38	72
		Esperado	30.5	41.5	72
	7-9	Real	8	20	28
		Esperado	11.9	16.1	28
	>10	Real	56	54	110
		Esperado	46.7	63.3	110
Total		Real	263	357	620
		Esperado	263	357	620

TABLA LI  
PRUEBA  $\chi^2$  PARA LAS VARIABLES SEXO E INTERVALOS DE SALIDA

Valor	Grados de libertad	Valor p
8.718	4	0.069

Ya que trabajamos a un nivel de significancia del 95%, se acepta la hipótesis nula, lo que implica que las variables son independientes.

TABLA LII  
MEDIDAS DIRECCIONALES PARA VARIABLES SEXO E INTERVALOS DE SALIDA

		Valor	Error
Lambda	Simétrico	0.004	0.02
	Tiempo de salida dependiente	0.000	0.000
	Sexo dependiente	0.008	0.04
Goodman and	Tiempo de salida dependiente	0.003	0.000
Kruskal tau	Sexo dependiente	0.014	0.01

La variable TIEMPO DE SALIDA se puede predecir por medio de la variable SEXO con 0% a 0.03% de exactitud.

La variable EDAD se puede predecir por medio de la variable TIEMPO DE SALIDA con 0.08% a 1.4% de exactitud.

TABLA LIII

MEDIDAS SIMÉTRICAS PARA VARIABLES SEXO E INTERVALOS DE SALIDA

	Valor	Valor p.
Coefficiente de contingencia	0.118	0.069

Ya que trabajamos a un nivel de significancia del 95%, se acepta a la hipótesis nula, lo que implica que el coeficiente de contingencia es estadísticamente nulo y no existe una asociación considerable entre estas dos variables.

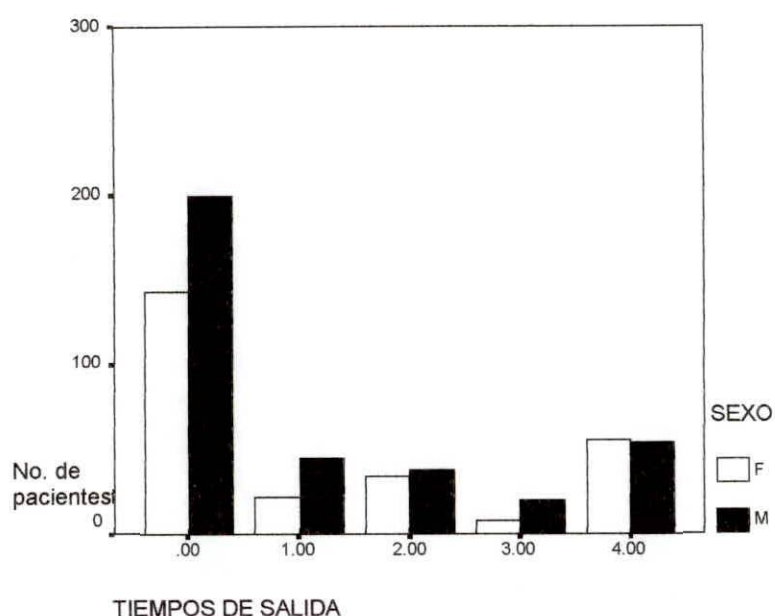


Gráfico 3.6

Ya que el poder de predicción de una variable por otra es sumamente bajo y se aceptaría la hipótesis de independencia a un 95% de nivel de significancia, se prefiere declarar las variables TIEMPO DE SALIDA y SEXO independientes.

### 3.1.2. Guayaquil y Quito

#### 3.1.2.1 Lugar donde vive vs. Tiempo de salida

Queremos saber si el lugar donde vive el infante influye en el tiempo que le tomó salir de la institución.

La variable LUGAR DONDE VIVE para esta prueba toma 3 valores:

1. Guayaquil.
2. Quito
3. Otros.

La variable tiempo de salida toma los mismos valores que en las pruebas anteriores.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"El lugar donde vive el infante es independiente del tiempo que le tomó para salir de la institución"

vs.

"El lugar donde vive el infante es dependiente del tiempo que le tomó para salir de la institución"

A continuación la tabla de contingencia



TABLA LIV

TABLA DE CONTINGENCIA DE LAS VARIABLES LUGAR DONDE VIVE E INTERVALOS DE SALIDA

			Tiempo de salida					
			0	1-3	4-6	7-9	>10	Total
Lugar Donde Vive	Gquil	Real	182	64	28	13	5	292
		Esperado	176	64.8	22.1	19.9	8.7	292
	Quito	Real	81	48	18	27	12	186
		Esperado	112	41.3	14.1	12.7	5.5	186
	Otros	Real	224	67	15	15	7	328
		Esperado	198	72.8	24.8	22.4	9.8	328
	Total	Real	487	179	61	55	24	806
		Esperado	487	179	61	55	24	806

TABLA LV

PRUEBA  $\chi^2$  PARA LAS VARIABLES LUGAR DONDE VIVE E INTERVALOS DE SALIDA

Valor	Grados de libertad	Valor p
51.287	8	0.000

Ya que trabajamos a un nivel de significancia del 95%, se rechaza la hipótesis nula, lo que implica que las variables son dependientes.

TABLA LVI

MEDIDAS DIRECCIONALES PARA VARIABLES LUGAR DONDE VIVE E INTERVALOS DE SALIDA

		Valor	Error
Lambda	Simétrico	.038	0.012
	Lugar donde vive dependiente	0.063	0.021
	Tiempo de salida dependiente	0.000	0.000
Goodman and	Lugar donde vive dependiente	0.028	0.008
Kruskal tau	Tiempo de salida dependiente	0.022	0.007

La variable TIEMPO DE SALIDA se puede predecir por medio de la variable LUGAR DONDE VIVE con 2.8% a 6.3% de exactitud.

La variable LUGAR DONDE VIVE se puede predecir por medio de la variable TIEMPO DE SALIDA con 0% a 2.2% de exactitud.

TABLA LVII

MEDIDAS SIMÉTRICAS PARA VARIABLES LUGAR DONDE VIVE E  
INTERVALOS DE SALIDA

	Valor	Valor p.
Coefficiente de contingencia	0.245	0.000

Ya que trabajamos a un nivel de significancia del 95%, se rechaza a la hipótesis nula, lo que implica que el coeficiente de contingencia no es estadísticamente nulo y existe una asociación considerable entre estas dos variables

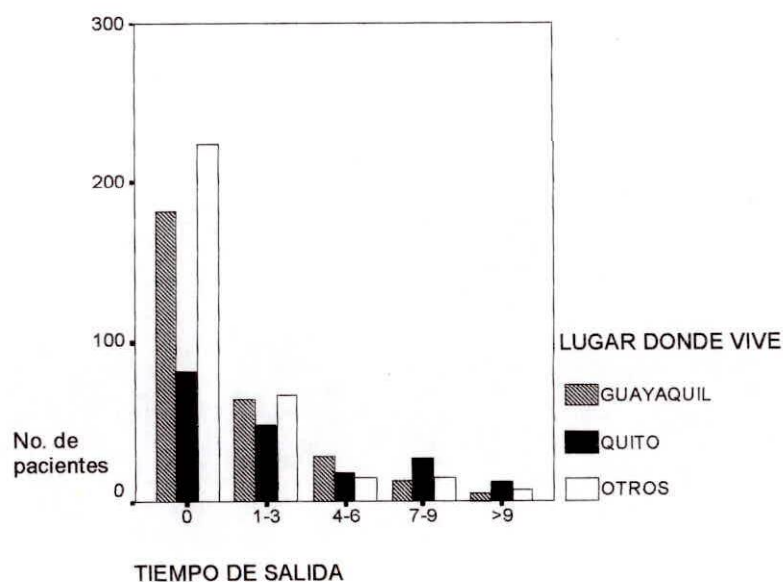


Gráfico 3.7

Las variables TIEMPO DE SALIDA y LUGAR DONDE VIVE son dependientes.

### **3.1.2.2 Lugar donde vive vs. Estado**

Queremos saber si el lugar donde vive el infante influye en el estado en el que salió de la institución.

La variable LUGAR DONDE VIVE para esta prueba toma 3 valores:

1. Guayaquil.
2. Quito
3. Otros.

La variable estado toma los mismos valores que en las pruebas anteriores.

Para poder comprobar estadísticamente si hay alguna relación entre estas variables, realizamos una tabla de contingencia con la hipótesis nula de

"El lugar donde vive el infante es independiente del estado en que se encontraba al salir de la institución"

vs.

"El lugar donde vive el infante es dependiente del estado en que se encontraba al salir de la institución"

A continuación la tabla de contingencia

TABLA LVIII

TABLA DE CONTINGENCIA DE LAS VARIABLES LUGAR DONDE VIVE Y ESTADO

		Estado				
			?	B	F	Total
Lugar Donde Vive	Gquil	Real	126	52	114	292
		Esperado	160.5	39.1	92.4	292
	Quito	Real	131	19	36	186
		Esperado	102.2	24.9	58.8	186
	Otros	Real	186	37	105	328
		Esperado	180.3	44	103.8	328
Total		Real	443	108	255	806
		Esperado	443	108	255	806

TABLA LIX

PRUEBA  $\chi^2$  PARA LAS VARIABLES LUGAR DONDE VIVE Y ESTADO

Valor	Grados de libertad	Valor p
36.376	4	0.000

Ya que trabajamos a un nivel de significancia del 95%, se rechaza la hipótesis nula, lo que implica que las variables son dependientes.

TABLA LX

MEDIDAS DIRECCIONALES PARA VARIABLES LUGAR DONDE VIVE Y ESTADO

			Valor	Error
Lambda	Simétrico		0.029	0.021
	Lugar donde vive	Dependiente	0.05	0.036
	Estado	Dependiente	0.000	0.000
Goodman and	Lugar donde vive	Dependiente	0.021	0.007
Kruskal tau	Estado	Dependiente	0.03	0.009

La variable ESTADO se puede predecir por medio de la variable LUGAR DONDE VIVE con 0% a 3% de exactitud.



La variable LUGAR DONDE VIVE se puede predecir por medio de la variable TIEMPO DE SALIDA con 2.1% a 5% de exactitud.

TABLA LXI  
MEDIDAS SIMÉTRICAS PARA VARIABLES LUGAR DONDE VIVE Y ESTADO

	Valor	Valor p
Coefficiente de contingencia	0.208	0.000

Ya que trabajamos a un nivel de significancia del 95%, se rechaza a la hipótesis nula, lo que implica que el coeficiente de contingencia no es estadísticamente nulo y existe una asociación considerable entre estas dos variables.

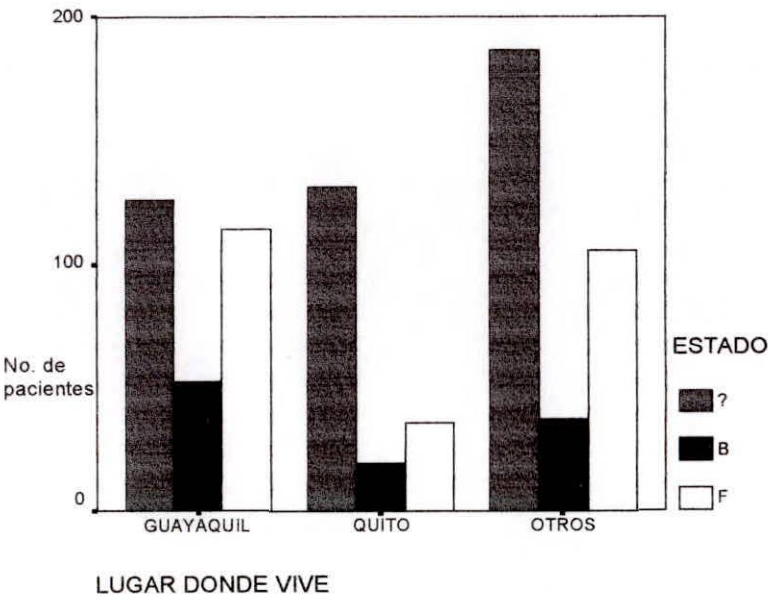


Gráfico 3.8

Las variables ESTADO y LUGAR DONDE VIVE son dependientes.

### **3.2 Análisis de varianza**

Como se mostró en el capítulo introductorio, los tratamientos principales para detener y/o erradicar el cáncer son quimioterapia, radioterapia y cirugía.

Los doctores examinan la situación de cada paciente y dependiendo del tipo de cáncer que padezca, de cuán avanzado esté y de su estado de salud, determinan qué tipo de combinación de estos tres tratamientos básicos se necesita, según el experto en oncología clínica de Solca de Guayaquil, el Dr. Guillermo Paulson.

Al encontrarnos ante esta situación, surge una curiosidad: ¿habrá algún tipo de combinación de los tratamientos que sea más eficaz o que permita al paciente vivir más tiempo? Al examinar la muestra que se tiene de niños que han padecido cáncer, podremos clasificarlos según las diferentes combinaciones de tratamiento que han recibido. Para este procedimiento, utilizaremos a la parte de la muestra donde sabemos certeramente si los pacientes han muerto o todavía se encuentran bien, esto da un total de 310 niños tratados con cáncer entre 1986 y 1994 de Quito y Guayaquil.

Antes de analizar esta clasificación, codificaremos todas estas combinaciones en forma binaria, de la siguiente manera:

TABLA LXII

NOMENCLATURA DE LAS DIFERENTES COMBINACIONES DE LOS TRATAMIENTOS

Tipo de combinación	Quimioterapia	Radioterapia	Cirugía
Sin tratamiento	0	0	0
Sólo cirugía	0	0	1
Sólo radioterapia	0	1	0
Cirugía y radioterapia	0	1	1
Sólo quimioterapia	1	0	0
Quimioterapia y cirugía	1	0	1
Quimioterapia y radioterapia	1	1	0
Los tres tratamientos juntos	1	1	1

Para poder analizar si las diferentes combinaciones de los tratamientos afectan el tiempo de supervivencia de los niños, necesitaremos realizar un análisis de varianza. Para realizar un análisis de varianza normal, necesitamos comprobar dos supuestos:

- Debe de haber homocedasticidad entre los datos de los diferentes factores, es decir que las varianzas de los datos en los diferentes factores deben ser estadísticamente iguales
- Los errores deben tener una distribución normal.

Para este efecto, tenemos 8 grupos, es decir tres diferentes factores (Quimioterapia, radioterapia y cirugía) con 2 niveles cada uno (si recibió o no el tratamiento).

Al realizar la prueba de Levene, cuya hipótesis nula es que sí hay homocedasticidad entre los diferentes grupos, obtenemos un p de valor 0, por lo que rechazamos la hipótesis nula y aceptamos que las varianzas en los diferentes grupos es diferente.

Al realizar la prueba de Kolmogorov Smirnov para comprobar el segundo supuesto de normalidad de los datos, encontramos que nuestro p es igual a cero, por lo tanto este supuesto tampoco se cumple.

Por lo tanto procedemos a realizar un análisis de varianza no paramétrico.

### **3.2.1 La prueba de Kruskal Wallis**

La prueba de Kruskal Wallis, es una prueba que sirve para comparar dos o más grupos divididos por los factores de una variable y saber si estos grupos tienen el mismo efecto o no en otra variable.



Para este efecto se debe cumplir que en cada grupo debe de haber al menos 5 individuos y ser muestras independientes.

La hipótesis nula de esta prueba es que los efectos de los  $k$  grupos son iguales (a cada efecto se lo denota como  $\tau_i$ ), por lo que

$$H_0: \tau_1 = \tau_2 = \tau_3 = \dots = \tau_k$$

Versus la hipótesis alterna: algún  $\tau_i$  es diferente de los demás.

Primero se debe ordenar a cada individuo según su tiempo de supervivencia y asignarle el número que le corresponda según el orden en que cayó, indiferentemente de a qué grupo pertenezca. Una vez que cada individuo tiene su propio rango, se halla la suma de los rangos total y por grupos. Las fórmulas están dadas a continuación:

Suma de rangos por grupos

$$R_j = \sum_{i=1}^{n_j} r_{ij}$$

Fórmula 3.1

Media de rangos por grupo

$$R_j = \frac{R_j}{n_j}$$

Fórmula 3.2

El estadístico utilizado para esta prueba es el H, y su fórmula está dada por:

$$H = \left( \frac{12}{n(n+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} \right) - 3(n+1)$$

Fórmula 3.3

Donde k es el número de tratamientos,  $n_j$  desde 1 hasta k son los diferentes tamaños de los grupos, y n es la suma de los individuos.

Se rechaza la hipótesis nula si H es mayor a un valor  $\chi^2$  con k-1 grados de libertad y a un nivel de significancia  $\alpha$ , y se la acepta si H es menor a este mismo valor.



Al analizar las estadísticas descriptivas de la muestra divididas en los 8 grupos, nos damos cuenta que el grupo 011 y el 001 tienen 2 y 3 individuos respectivamente, por lo que no se cumple un supuesto para realizar la prueba de Kruskal Wallis. Procedemos a agrupar estos dos tratamientos, verificando si estos son homogéneos por medio de la prueba de Mann Whitney para probar que ambos grupos de datos son de la misma población (hipótesis nula), de donde obtenemos la siguiente información:

Estadístico U	2
Valor P	0.553617

Con el valor p de 0.553617 podemos aceptar la hipótesis que los grupos 001 y 011 son de la misma población, por lo que lo agruparemos en uno solo para realizar las pruebas de Kruskal Wallis con los demás grupos.

Para un nivel de significancia de 0.05 y con 7 tratamientos o grupos, se realizan las siguientes pruebas con los siguientes grupos:

### 3.2.1.1 Prueba para todos los grupos

TABLA LXIII

TABLA DE MEDIAS DE RANGOS DE LA PRUEBA DE TODOS LOS GRUPOS

Grupo	n <sub>j</sub>	Media de rangos
000	76	80.81578827
010	6	159.5833282
001 - 011	5	175.6000061
100	113	156.7566376
101	8	217.5625
110	90	206.422226
111	12	182.9583282
Total	310	155.5

TABLA LXIV

ESTADÍSTICO  $\chi^2$ 

	Valor
Estadístico $\chi^2$	98.2373657
Grados de libertad	6

P

0.000

Se rechaza la hipótesis, lo implica que al menos un grupo tiene el tiempo de supervivencia de sus integrantes diferentes a los de los demás.

Al comparar la media de los rangos de cada grupo con la media de rangos total, podemos observar los valores de las medias de los rangos que están más alejadas de la media total de rangos, y así excluirlo. Este procedimiento se repite hasta que se acepte la hipótesis nula de homogeneidad en el valor del tiempo de supervivencia entre los grupos que participan en la prueba.

Para la siguiente prueba eliminaremos el grupo 000.

### 3.2.1.2 Prueba sin el grupo 000

TABLA LXV

TABLA DE MEDIAS DE RANGOS DE LA PRUEBA SIN EL GRUPO 000

Grupo	nj	Media de rangos
001 - 011	5	116.099998
010	6	103.75

100	113	99.6991119
101	8	151.875
110	90	137.738892
111	12	117.875
Total	234	117.5

TABLA LXVI  
ESTADÍSTICO  $\chi^2$

	Valor
Estadístico $\chi^2$	18.9514408
Grados de libertad	5
P	0.00196258

Se rechaza la hipótesis, lo que implica que todavía existe al menos un grupo que tiene el tiempo de supervivencia de sus integrantes diferentes a los de los demás; aunque podemos notar que el p comienza a subir, lo que implica que el grupo 000 realmente tenía tiempos de supervivencia bastante diferentes a los demás.

Al analizar los valores de las medias de los rangos, nos damos cuenta que los grupos 111 y 001-011 tienen una media de rangos muy cercana a la media de rangos total. Por esta razón dividiremos a los grupos en tres: los grupos que tienen la media de rangos menor

a la media de rangos total, los que la tienen mayor y los que la tienen aproximadamente igual.

### 3.2.1.3 Prueba con los grupos 10 y 100 (Media de rangos de grupo menor a la media total)

TABLA LXVII

TABLA DE MEDIAS DE RANGOS DE LA PRUEBA CON LOS GRUPOS 10 Y 100

<i>Grupo</i>	<i>n<sub>j</sub></i>	Media de rangos
10	6	61.33333206
100	113	59.92920303
Total	119	60

TABLA LXVIII

ESTADÍSTICO  $\chi^2$

	Valor
Estadístico $\chi^2$	0.010524675
Grados de libertad	1

P

0.918288507

El p de la prueba es sumamente alto, por lo que se concluye que los tiempos de salidas de estos grupos son realmente homogéneos.

### 3.2.1.4 Prueba con los grupos 110 y 101 (Media de rangos de grupo mayor a la media total)

TABLA LXIX

TABLA DE MEDIAS DE RANGOS DE LA PRUEBA CON LOS GRUPOS 110 Y 101

Grupo	nj	Media de rangos
101	8	58.9375
110	90	48.6611099
Total	98	49.5

TABLA LXX

ESTADÍSTICO  $\chi^2$ 

Estadístico $\chi^2$	0.9730128
Grados de libertad	1
P	0.32392994



El p de esta prueba es alto, por lo que se acepta la hipótesis de que estos dos grupos son homogéneos con respecto al tiempo de supervivencia de sus integrantes.

### 3.2.1.5 Prueba con los grupos 001-011 y 111 (Media de rangos de grupo aproximadamente igual a la media total)

TABLA LXXI

TABLA DE MEDIAS DE RANGOS DE LA PRUEBA CON LOS GRUPOS 111 Y 001-011

Grupo	Nj	Media de rangos
001-011	5	8.69999981
111	12	9.125
Total	17	

TABLA LXXII

ESTADÍSTICO  $\chi^2$

Estadístico $\chi^2$	0.02632258
Grados de libertad	1
P	0.87111501

Ya que tenemos un p alto, aceptamos la hipótesis que ambos grupos tienen efectos iguales.

### 3.2.1.6 Agrupación final

Para poder tener grupos más concentrados según los efectos de cada combinación de tratamientos en el tiempo de supervivencia, se realizaron diferentes pruebas entre los grupos con media parecida a la total y los grupos con media de rangos mayor a la media total, y con los grupos de menor media que la total. Los resultados se muEstran en la tabla siguiente:

TABLA LXXIII

TABLA DE MEDIAS DE RANGOS DE CON POSIBLES  
AGRUPACIONES FINALES

Tipo de grupos con respecto a la media de rangos total	Grupos a los cuales se les realizó la prueba	P encontrado en la prueba
Con media menor a la a total y con media igual la total	010,100,011-001	0.897
	010,100,111	0.593
	010,100,111,011-001	0.753
Con media mayor a la total y con media igual a la total	110,101,011-001	0.529
	110,101,111	0.341
	110,101,111,011-001	0.499

A pesar de que la unión de los grupos con media de rangos igual a la media de rangos total es válida para cualquiera de los otros grupos con medias mayor o menor que el de la media total, preferimos unirlos al grupo con media menor a la del total, puesto que para estas pruebas los p's obtenidos son mucho más altos.

### **3.2.2 Resultados de las agrupaciones y comparación del tiempo de supervivencia**

Después de hacer todo este estudio podremos agrupar los datos en tres grupos con diferentes tiempo de supervivencia:

1. El grupo conformado por los que recibieron las combinaciones de quimioterapia y ya sea con radioterapia (110) o con cirugía (101), los cuales estadísticamente tienen el mismo efecto en el tiempo de supervivencia proporcionado al individuo, es el que tiene un tiempo de supervivencia esperado mayor al resto de combinaciones de tratamientos. El tiempo de supervivencia esperado para este grupo es de 4.01 años y una desviación estándar de 3.33 años.

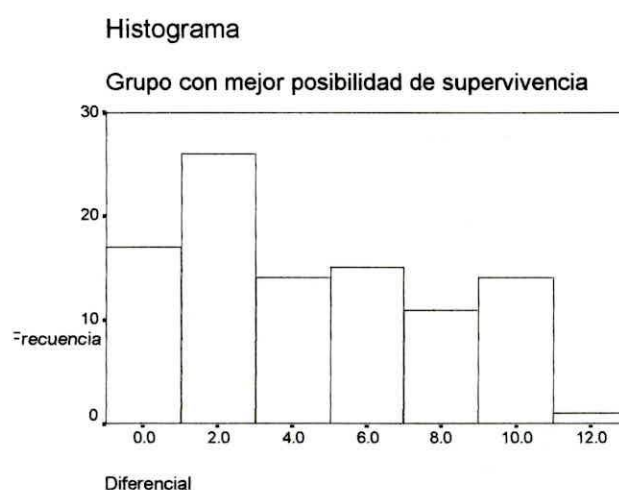


Gráfico 3.9

2. El grupo conformado por los que recibieron tratamiento con sólo radioterapia (010), radioterapia y cirugía (011), sólo cirugía (001), sólo quimioterapia (100) o los tres tratamientos juntos (111) es el segundo grupo con mayor tiempo esperado de supervivencia, el cual es de 2.56 años con una desviación estándar de 3.49 años

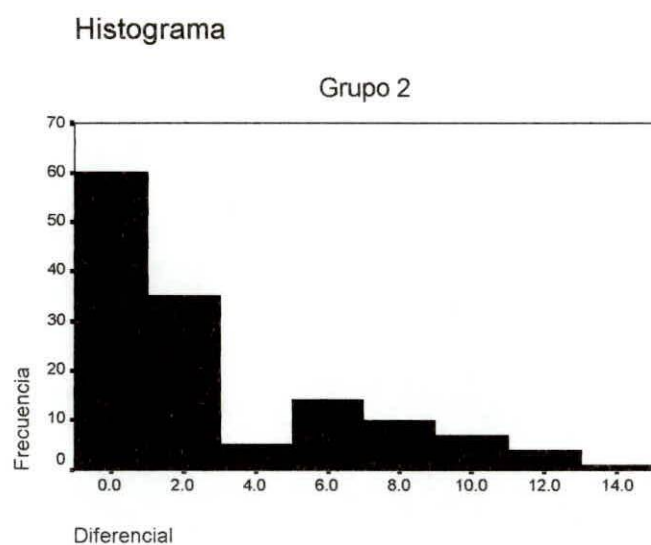


Gráfico 3.10

3. El grupo que no recibe tratamiento (000), el cual tiene el tiempo de supervivencia esperado más bajo de todos, el cual es de 0.11 años con una desviación estándar de 0.6 años.

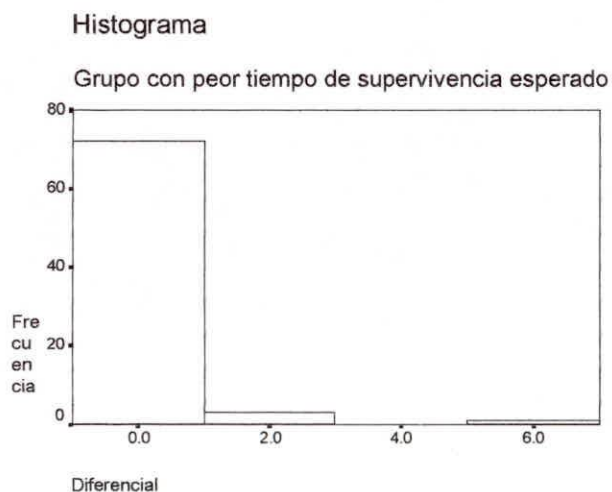


Gráfico 3.10

## **CAPÍTULO IV**

### **TABLAS DE SUPERVIVENCIA**



## Capítulo 4

### 4. TABLAS DE SUPERVIVENCIA

Necesitamos hacer un análisis más profundo de supervivencia de los niños que continuaron su tratamiento.

Al construir una tabla de supervivencia de niños con cáncer, tendremos que saber si otros factores; como la edad, sexo o el tipo de cáncer que padezcan, afectan el estado en que se encuentra o encontró el niño al momento de su egreso de la institución.

Tomando las conclusiones del capítulo anterior, se muestra que estadísticamente la edad y el sexo no representan una gran influencia en el estado actual del niño, al contrario del tipo de cáncer que padezcan y de

donde provienen, por lo que se tendrá que construir tablas de supervivencia clasificándolas por tipo de cáncer y por lugar donde viven.

Junto con esta premisa y tomando los valores conocidos como Fallecidos (F) y Bien (B), construiremos cinco diferentes tablas de supervivencia para los tipos de carcinomas siguientes:

- Linfoma de hodkin
- Leucemia linfoide

Se escogió estos dos tipos de carcinoma porque la cantidad acumulada de casos encontrados en estado Fallecido y Bien es suficientemente grande, en comparación con la de los otros, para que el error de supervivencia sea menor al 10%.

Con respecto al lugar de donde residen los niños de la muestra , se escogió la ciudad de Guayaquil y Quito por tener menos índices de datos perdidos.

Para entender de mejor forma la tabla supervivencia, a continuación se presenta una explicación de cada una de las variables y probabilidades que se utilizan para construirla

$x$ : número de años que transcurren desde que el niño empezó su tratamiento.

$l_x$ : número de niños vivos al tiempo  $x$  de tratamiento.

$d_x$ : número de niños fallecidos al tiempo  $x$  de tratamiento.

$q_x$ : probabilidad que un niño que está vivo al tiempo de tratamiento  $x$ , no sobreviva al tiempo  $x+1$ .

$p_x$ : probabilidad que un niño que está vivo al tiempo de tratamiento  $x$ , sobreviva al tiempo  $x+1$ .

$l_{x+1}/l_0$ : probabilidad de que un niño sobreviva al  $x+1$  del año de tratamiento.

$D_x/l_0$ : probabilidad de que un niño fallezca al  $x$  año de tratamiento

A continuación presentamos las tablas de supervivencia de los casos antes expuestos

TABLA LXXIV

## TABLAS DE SUPERVIVENCIA DE NIÑOS CON CÁNCER EN GUAYAQUIL

## Linfoma de hodkin

x	$l_x$	$d_x=l_x-l_{x+1}$	$q_x=d_x/l_x$	$p_x=1-q_x=l_{x+1}/l_x$	$l_{x+1}/l_0$	$d_x/l_0$	Error#
0	26	2	0.0769	0.9231	0.9231	0.0769	0.0532
1	24	1	0.0417	0.9583	0.8846	0.0385	0.0643
2	23	0	0	1	0.8846	0	0.0643
3	23	1	0.0435	0.9565	0.8462	0.0385	0.0732
4	22	0	0	1	0.8462	0	0.0732
5.0+	22	0	0	1	**	**	**

## Leucemia linfoide

x	$l_x$	$d_x=l_x-l_{x+1}$	$q_x=d_x/l_x$	$p_x=1-q_x=l_{x+1}/l_x$	$l_{x+1}/l_0$	$d_x/l_0$	Error#
0	205	129	0.6293	0.3707	0.3707	0.6293	0.0337
1	76	17	0.2237	0.7763	0.2878	0.0829	0.0316
2	59	14	0.2373	0.7627	0.2195	0.0683	0.0289
3	45	6	0.1333	0.8667	0.1902	0.0293	0.0274
4	39	4	0.1026	0.8974	0.1707	0.0195	0.0263
5.0+	35	1	0.0556	0.9444	**	**	**

TABLA LXXV  
TABLAS DE SUPERVIVENCIA DE NIÑOS CON CÁNCER EN QUITO

Leucemia linfoide

x	$l_x$	$d_x=l_x-l_{x+1}$	$q_x=d_x/l_x$	$p_x=1-q_x=l_{x+1}/l_x$	$l_{x+1}/l_0$	$d_x/l_0$	Error#
0	52	12	0.230769	0.769230769	0.76923	0.230769	0.059
1	40	5	0.125	0.875	0.67308	0.096154	0.0657
2	35	4	0.114286	0.885714286	0.59615	0.076923	0.0687
3	31	3	0.096774	0.903225806	0.53846	0.057692	0.0698
4	28	1	0.03571	0.964285714	0.51923	0.01923	0.07
5.0+	27	27	1	0	**	**	**

# Para encontrar el error utilizamos la fórmula del estimador de la desviación estándar de la proporción, el cual está dado por la fórmula

$$\sqrt{\frac{p(1-p)}{n-1}}$$

Fórmula 4.1



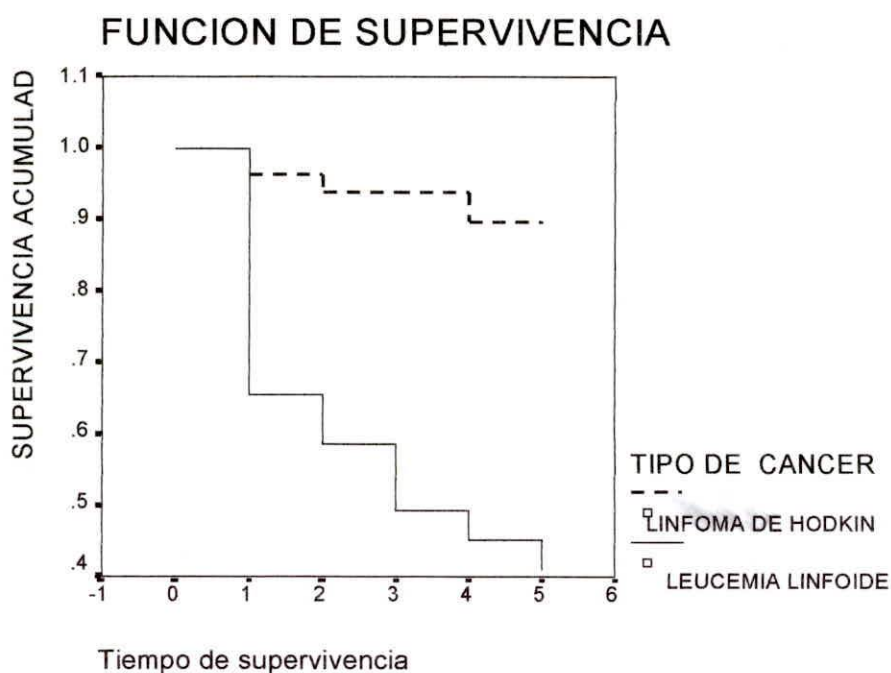


Gráfico 4.1

Tanto en la tabla de supervivencia como en el gráfico, se nota claramente lo mortal que es la leucemia en los niños, ya que sólo sobreviven el 17.07% de los niños al quinto año de tratamiento, al contrario que el linfoma de hodkin que tiene una probabilidad de 84.62 de supervivencia al quinto año.

Se nota una gran diferencia entre el porcentaje de supervivencia de niños con leucemia en la ciudad de Guayaquil y en la ciudad de Quito, al quinto año de tratamiento. En Quito, el porcentaje de niños que viven al quinto año de tratamiento es de 51.92%. mientras que en Guayaquil sólo es de 17.07%.



El porcentaje de mortalidad más elevado lo encontramos al primer año de tratamiento. En Guayaquil, al primer año de tratamiento mueren aproximadamente 62.93% de los niños con leucemia y 7.69% de los niños con linfoma de hodkin; y en Quito fallecen 23.07% de los niños con leucemia en el primer año de tratamiento.

## **CAPÍTULO V**

### **MODELO PARA PREDECIR NÚMERO DE PACIENTES QUE SERÁN ATENDIDOS EN SOLCA**

## Capítulo 5

### 5. MODELO PARA PREDECIR NÚMERO DE PACIENTES QUE SERÁN ATENDIDOS EN SOLCA

En este capítulo, nos dedicaremos por completo a encontrar modelos que nos faciliten la predicción del número de casos que se tendrán en años futuros.

Esta información es de vital importancia, ya que ayudará a las instituciones oncológicas especializadas a programar sus recursos a futuro, para poder atender a la población afectada que está en constante crecimiento.

#### 5.1 Los datos

Los datos utilizados para encontrar estos modelos matemáticos son el número de pacientes atendidos por en Solca de Guayaquil\*, desde el año 1985 hasta 1996, clasificados por tipos de cáncer. A continuación se muestra una lista:

TABLA LXXVI

NÚMERO DE PACIENTES NUEVOS CON TIPOS DE CÁNCER MÁS  
FRECIENTES DE LA CIUDAD DE GUAYAQUIL (1985-1996)

Hombres												
Año	85	86	87	88	89	90	91	92	93	94	95	96
Piel	51	59	44	39	52	50	43	47	28	21	37	46
Linfomas	34	21	28	32	31	45	27	31	32	12	21	33
Pulmón	22	23	37	27	28	23	25	33	36	40	26	27
Estómago	55	65	53	56	53	43	54	67	64	65	56	74
Leucemia	10	17	33	29	28	18	31	27	53	40	51	47
Próstata	23	27	22	16	21	35	40	31	55	65	71	63
Otros	157	182	197	207	201	217	213	216	223	252	256	241
Total	352	394	414	406	414	431	433	452	491	495	518	531

Mujer												
Año	85	86	87	88	89	90	91	92	93	94	95	96
Piel	44	61	53	58	49	64	47	39	46	43	47	40
Linfomas	16	22	20	17	19	19	15	18	20	24	19	25
Mamas	98	107	102	89	106	124	112	129	151	152	177	196
Estómago	19	40	21	26	30	28	30	35	43	54	35	35

Leucemia	15	14	20	24	15	21	23	26	28	81	53	47
Ovario	15	23	25	25	19	26	25	21	43	44	34	33
Cervix	416	434	379	430	447	435	430	457	422	457	463	529
Otros	127	136	153	156	146	179	192	202	217	156	200	212
Total	750	837	773	825	831	896	874	927	970	1011	1038	1130

## 5.2 ¿Qué método utilizaremos?

En estadística hay dos métodos para calcular tendencias de valores a futuro en función de valores pasados: series de tiempo y regresión. Para utilizar la primera, se necesita que la correlación entre los errores de predicción sean diferentes de cero, es decir que estos errores estén correlacionados; mientras que para la segunda se necesita que dicha correlación sea igual a cero.

Para poder determinar qué método utilizar, realizamos la prueba de Durbin Watson, la cual tiene como hipótesis nula que la correlación de los errores es igual a cero, en contraste con la hipótesis alterna que especifica que dicha correlación es diferente de cero. Para establecer si se rechaza o no dichas hipótesis, se calcula un estadístico de nombre D, que está dado por la siguiente fórmula:

$$D = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad \text{Fórmula 5.1}$$

Donde  $e_t$  es el error en el tiempo  $t$  ( $e_t = Y_t - \hat{Y}_t$ ).

En nuestro caso, para  $n=12$  y a un 95% de confianza, tenemos los límites para rechazar o aceptar la hipótesis nula dados por:

Si  $D > 1.33$ , entonces aceptamos la hipótesis nula.

Si  $D < 1.02$ , entonces aceptamos la hipótesis alterna.

Si  $1.02 \leq D \leq 1.33$ , entonces no se puede concluir.

Haciendo la prueba para las variables tipo de cáncer para hombres y para mujeres obtuvimos los siguientes valores para  $D$

TABLA LXXVII

VALORES DEL ESTADÍSTICO  $D$  PARA LA PRUEBA DE DURBIN WATSON

Mujeres

Tipo de cáncer	D	Se concluye que...
Piel	2.288	$\rho=0$
Linfomas	2.101	$\rho=0$
Mamas	1.094	No se puede concluir
Estómago	2.209	$\rho=0$
Leucemia	1.940	$\rho=0$
Ovarios	1.818	$\rho=0$
Cervix	2.279	$\rho=0$
Otros	2.044	$\rho=0$



## Hombres

Tipo de cáncer	D	Se concluye que ...
Piel	1.797	$\rho=0$
Linfomas	2.029	$\rho=0$
Pulmón	1.791	$\rho=0$
Estómago	1.839	$\rho=0$
Leucemia	2.241	$\rho=0$
Próstata	1.193	No se puede concluir
Otros	1.721	$\rho=0$

En los carcinomas de mama y de próstata no se puede llegar a una conclusión certera, por lo tanto hallaremos la autocorrelaciones y la autocorrelaciones parciales del error por medio de bandas de confianza, y así demostraremos si son o no nulas.

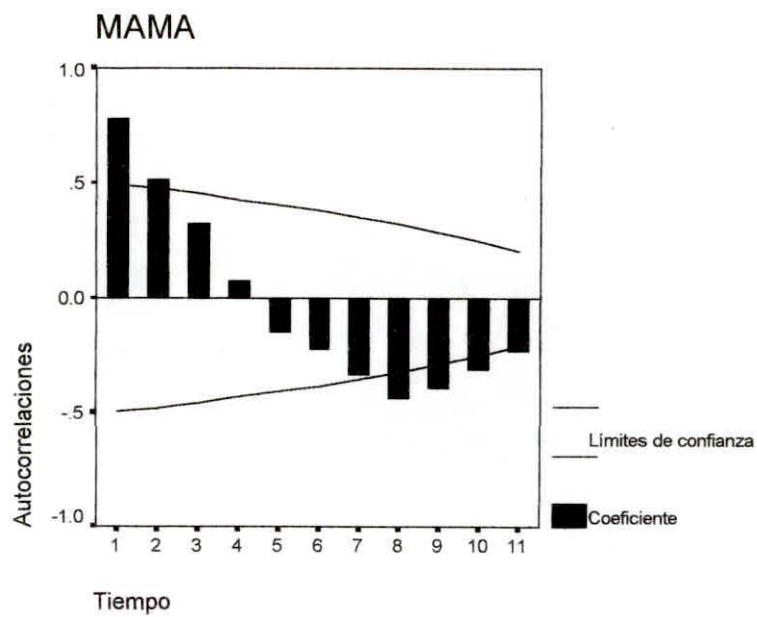


Gráfico 5.1

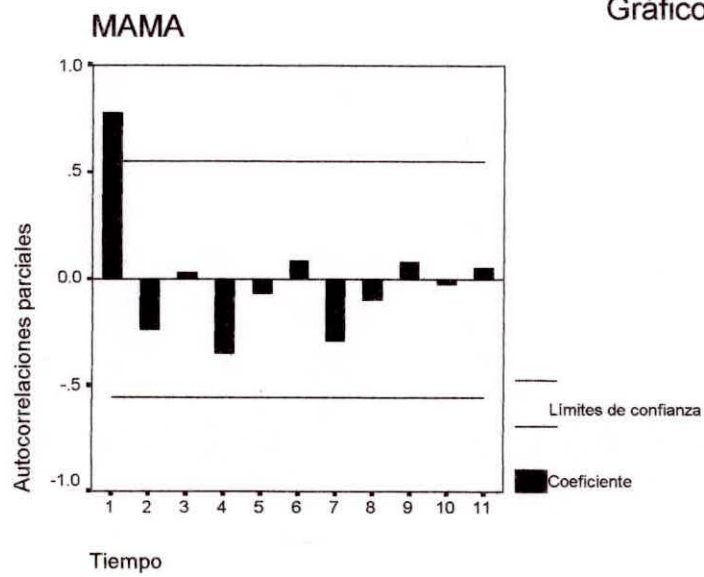


Gráfico 5.2

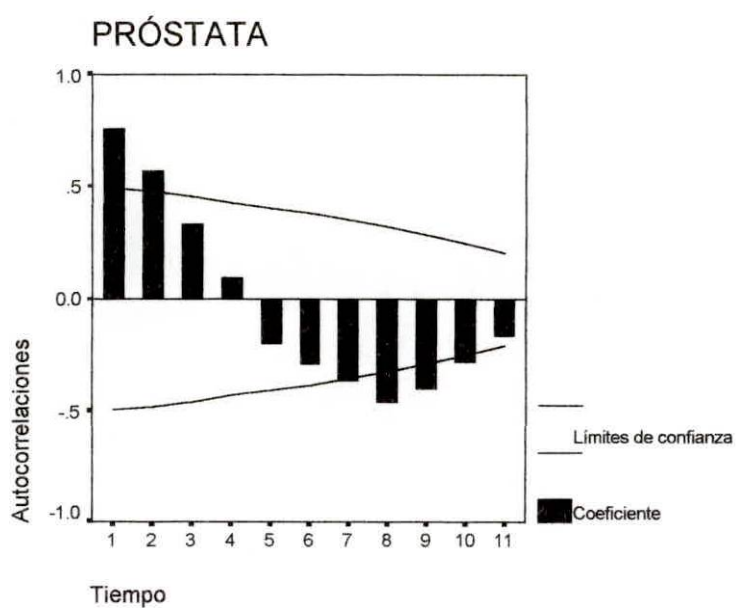


Gráfico 5.3

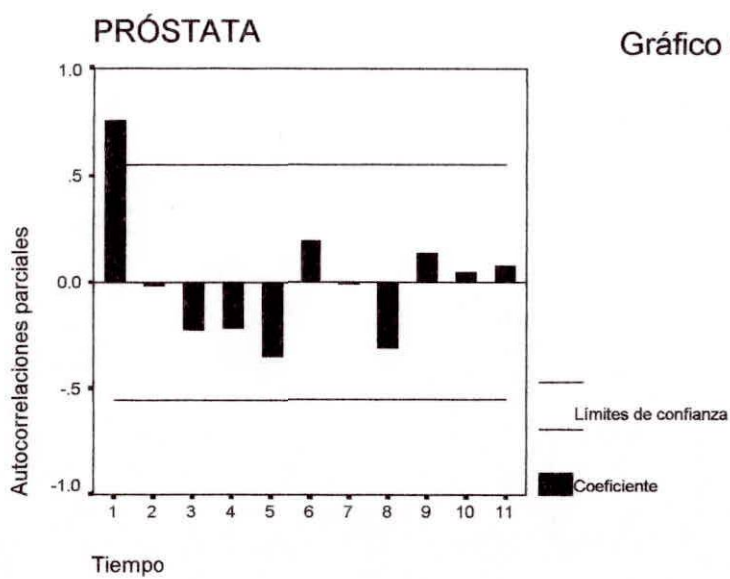


Gráfico 5.4

Al observar los gráficos de las autocorrelaciones de los errores, tanto de la variable próstata como de la variable mama, observamos que hay

evidencia de que el modelo necesita diferenciación, por lo que diferenciamos una vez a los valores, es decir los nuevos valores a utilizar ya no son  $Y_t$ , si no  $d_t = Y_t - Y_{t-1}$ , y se grafican a las autocorrelaciones y la autocorrelaciones parciales de los errores de estos nuevos valores diferenciados. Así obtenemos las bandas de confianza que a continuación se exponen

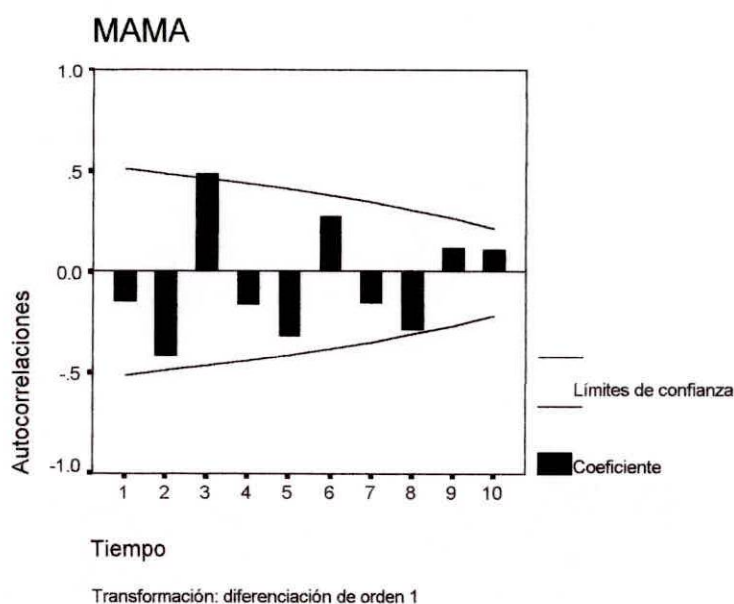


Gráfico 5.5

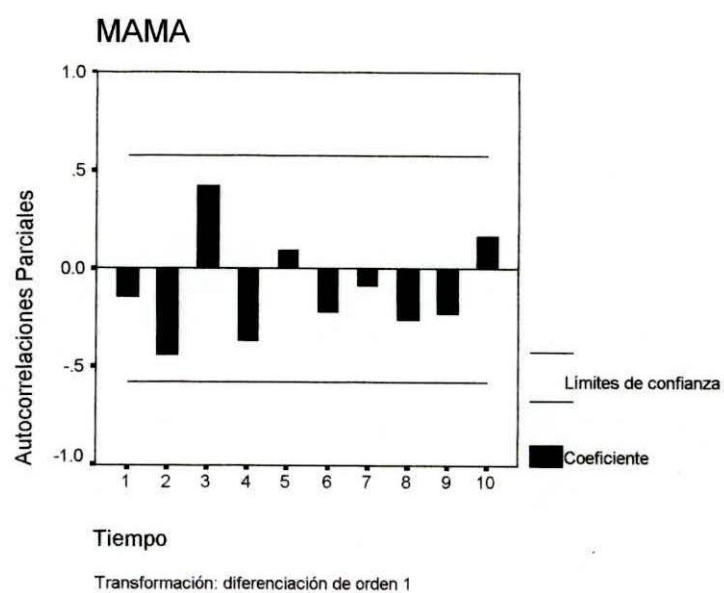


Gráfico 5.6

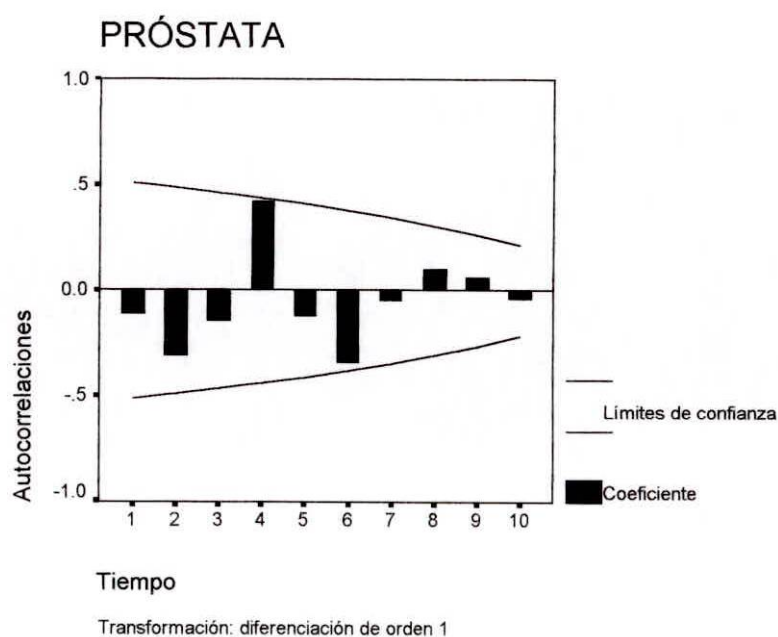


Gráfico 5.7

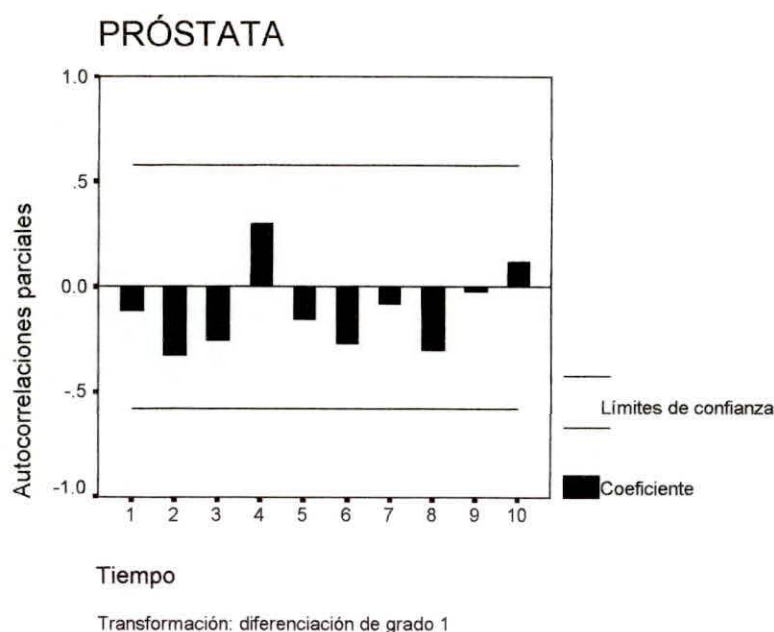


Gráfico 5.8

Puesto que al diferenciar las autocorrelaciones y las autocorrelaciones parciales son nulas, implica que no se puede utilizar métodos de series temporales para predecir los valores del número de pacientes carcinomas de mama o de próstata. Por lo tanto, utilizaremos el método de regresión para predecir el número de personas que serán atendidas en Solca de Guayaquil.

### 5.3 El método de regresión

Ya que el cáncer es una enfermedad con propiedades metastásicas, para poder determinar modelos que den resultados cercanos a la realidad, tendremos que investigar primero si existe correlación entre la aparición



de nuevos casos de cierto tipo de cáncer (al cual se le puede llamar tipo I) y la aparición de nuevos casos de otro tipo de cáncer (al cual lo llamaremos tipo II).

Ya que los datos obtenidos son sólo de 12 años, utilizaremos el estimador de máxima verosimilitud del coeficiente de correlación  $\rho$ , el cual se lo representa por  $r$ , y su fórmula es.

La predicción sólo se podrá realizar a nivel Guayaquil, ya que en Quito no se dispone de los datos por año.

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Fórmula 5.2

Donde

$X_i$ : es el número de pacientes nuevos del cáncer tipo 1 en el año  $i$ .

$Y_i$ : es el número de pacientes nuevos del cáncer tipo 2 en el año  $i$ .

$\bar{X}$  : es la media muestral del número de pacientes nuevos de tipo 1.

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \quad \text{Fórmula 5.3}$$

$\bar{Y}$  : es la media muestral del número de pacientes nuevos de tipo 2.

$$\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n} \quad \text{Fórmula 5.4}$$

$Y, n$  es el número de datos que se tengan, es decir 12.

Utilizando esta fórmula, se obtiene los diferentes  $r$ 's para las diferentes combinaciones de tipos de cáncer, y así obtenemos las matrices de correlación estimadas para hombres y para mujeres:

**Matriz 5.1** Matriz de correlación estimada de número de pacientes nuevos masculinos

	Piel	Linfomas	Pulmón	Estómago	Leucemia	Próstata
Piel	1	0.44	-0.71	-0.22	-0.7	-0.62
Linfomas		1	-0.42	-0.42	-0.32	-0.41
Pulmón			1	0.33	0.52	0.27
Estomago				1	0.42	0.43

Leucemia	1	0.77
Próstata		1

**Matriz 5.2** Matriz de correlación estimada de número de pacientes nuevos femeninos

	Piel	Linfomas	Mamas	Estomago	Leucemia	Ovario	Cervix
Piel	1	-0.12	-0.5	-0.25	-0.43	-0.22	-0.44
Linfomas		1	0.62	0.64	0.59	0.59	0.53
Mamas			1	0.55	0.7	0.69	0.76
Estomago				1	0.7	0.78	0.41
Leucemia					1	0.78	0.51
Ovario						1	0.28
Cervix							1

Al analizar la matriz de correlación entre los diferentes tipos de cáncer, observamos que estas variables tienen una alta correlación, por lo que se concluye que sus valores no pueden ser predecidos por un modelo lineal univariado.

Se descarta también la utilización de series temporales, puesto que al hallar las autocorrelaciones del error, estas son estadísticamente nulas.

A continuación se describe un modelo multivariado de regresión que se aplicará en este capítulo para calcular los valores de los futuros pacientes que padezcan de los diferentes tipos de cáncer en Solca de Guayaquil.

### 5.3.1 Modelo de Regresión Multivariado.

Si consideramos que tenemos  $n$  individuos a los cuales se les mide  $m$  características (respuestas), podemos obtener un modelo relacionando dichas respuestas o variables a explicar ( $Y_1, Y_2, \dots, Y_m$ ) con  $r$  variables de explicación ( $Z_1, Z_2, \dots, Z_r$ ). De esta manera, para cada individuo se asume un modelo de regresión de la siguiente forma:

$$Y_1 = \beta_{01} + \beta_{11} Z_1 + \dots + \beta_{r1} Z_r + \varepsilon_1$$

$$Y_2 = \beta_{02} + \beta_{12} Z_1 + \dots + \beta_{r2} Z_r + \varepsilon_2$$

.

.

.

$$Y_m = \beta_{0m} + \beta_{1m} Z_1 + \dots + \beta_{rm} Z_r + \varepsilon_m$$

El error denotado por  $\varepsilon = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m]$  tiene  $E(\varepsilon) = 0$  y  $\text{Var}(\varepsilon) = \Sigma$ .

Entonces, el error asociado con diferentes características puede estar correlacionado.

Si establecemos una notación paralela a la de la regresión clásica obtenemos que  $[Z_{j0}, Z_{j1}, \dots, Z_{jr}]$  son los valores de las variables de explicación del  $j$ -ésimo individuo,  $Y_j = [Y_{j1}, Y_{j2}, \dots, Y_{jm}]$  las variables a explicar del  $j$ -ésimo individuo y  $\varepsilon_j = [\varepsilon_{j1}, \varepsilon_{j2}, \dots, \varepsilon_{jm}]$  los errores. Entonces obtenemos la matriz de diseño:

**Matriz 5.3** Matriz de diseño de regresión multivariada.

$$\mathbf{Z}_{n \times (r+1)} = \begin{pmatrix} Z_{10} & Z_{11} & \dots & Z_{1r} \\ Z_{20} & Z_{21} & \dots & Z_{2r} \\ \vdots & \vdots & & \vdots \\ Z_{n0} & Z_{n1} & \dots & Z_{nr} \end{pmatrix}$$

Las otras matrices del modelo de regresión se encuentran descritas a continuación:

**Matriz 5.4** Matriz de respuestas del modelo de regresión multivariada

$$Y_{n \times m} = \begin{pmatrix} Y_{11} & Y_{12} & \dots & Y_{1m} \\ Y_{21} & Y_{22} & \dots & Y_{2m} \\ \vdots & \vdots & & \vdots \\ Y_{n1} & Y_{n2} & \dots & Y_{nm} \end{pmatrix}$$

$$= \begin{pmatrix} Y_{(1)} & Y_{(2)} & \dots & Y_{(m)} \end{pmatrix}$$

**Matriz 5.5** Matriz de los parámetros desconocidos en regresión multivariada

$$\beta_{(r+1) \times m} = \begin{pmatrix} \beta_{01} & \beta_{02} & \dots & \beta_{0m} \\ \beta_{11} & \beta_{12} & \dots & \beta_{1m} \\ \vdots & \vdots & & \vdots \\ \beta_{r1} & \beta_{r2} & \dots & \beta_{rm} \end{pmatrix}$$

$$= \begin{pmatrix} \beta_{(1)} & \beta_{(2)} & \dots & \beta_{(m)} \end{pmatrix}$$



**Matriz 5.6** Matriz de los errores en el modelo de regresión multivariada

$$\begin{aligned} \mathbf{\varepsilon}_{n \times m} &= \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} & \dots & \varepsilon_{1m} \\ \varepsilon_{21} & \varepsilon_{22} & \dots & \varepsilon_{2m} \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ \varepsilon_{n1} & \varepsilon_{n2} & \dots & \varepsilon_{nm} \end{pmatrix} \\ &= \begin{pmatrix} \varepsilon_{(1)} & \varepsilon_{(2)} & \dots & \varepsilon_{(m)} \end{pmatrix} \end{aligned}$$

El modelo de regresión multivariado lineal está definido de la siguiente manera:

$$\mathbf{Y}_{n \times m} = \mathbf{Z}_{n \times (r+1)} * \boldsymbol{\beta}_{(r+1) \times m} + \boldsymbol{\varepsilon}_{n \times m}$$

Donde  $\mathbf{Y}$ ,  $\mathbf{Z}$ ,  $\boldsymbol{\beta}$ ,  $\boldsymbol{\varepsilon}$  son elementos del espacio matrices, y  $E(\varepsilon_{(i)}) = 0$ ,  $\text{Cov}(\varepsilon_{(i)}, \varepsilon_{(j)}) = \sigma_{ik} \mathbf{I}$ ;  $i, k = 1, 2, \dots, m$

Las  $m$  características del  $j$ -ésimo individuo tienen una matriz de covarianza  $\Sigma = \{\sigma_{ik}\}$ , pero las características de diferentes individuos tienen un valor de correlación igual a cero. En este caso  $\boldsymbol{\beta}$  es un

parámetro desconocido, la matriz de diseño  $Z$  tiene en su  $j$ -ésima fila a los valores  $[z_{j0}, z_{j1}, \dots, z_{jr}]$ .

Al aplicar este modelo a nuestros datos, consideramos a la cantidad de pacientes por tipo de cáncer como las características a predecir; por lo tanto  $m$  es igual a 7 en el caso de pacientes masculinos y 8 en el caso de los pacientes femenino. El tamaño de muestra es igual a 12, ya que son 12 años a estudiar. Se considera a  $r$  como 1, puesto que la única variable de explicación es el año.

Entonces las matrices  $Z$  y  $Y$  quedan de la siguiente manera:

**Matriz 5.7** Matriz de diseño para hallar el número de pacientes que serán atendidos en Solca de Guayaquil

$$Z = \begin{pmatrix} 1 & 1985 \\ 1 & 1986 \\ 1 & 1987 \\ 1 & 1988 \\ 1 & 1989 \\ 1 & 1990 \\ 1 & 1991 \\ 1 & 1992 \\ 1 & 1993 \\ 1 & 1994 \\ 1 & 1995 \end{pmatrix}$$

**Matriz 5.8** Matriz de respuesta para los pacientes masculinos

$$Y = \begin{pmatrix} 51 & 34 & 22 & 55 & 10 & 23 & 157 \\ 59 & 21 & 23 & 65 & 17 & 27 & 182 \\ 44 & 28 & 37 & 53 & 33 & 22 & 197 \\ 39 & 32 & 27 & 56 & 29 & 16 & 207 \\ 52 & 31 & 28 & 53 & 28 & 21 & 201 \\ 50 & 45 & 23 & 43 & 18 & 35 & 217 \\ 43 & 27 & 25 & 54 & 31 & 40 & 213 \\ 47 & 31 & 33 & 67 & 27 & 31 & 216 \\ 28 & 32 & 36 & 64 & 53 & 55 & 223 \\ 21 & 12 & 40 & 65 & 40 & 65 & 252 \\ 37 & 21 & 26 & 56 & 51 & 71 & 256 \\ 46 & 33 & 27 & 74 & 47 & 63 & 241 \end{pmatrix}$$

**Matriz 5.9** Matriz de respuesta para las pacientes femeninas

$$Y = \begin{pmatrix} 44 & 16 & 98 & 19 & 15 & 15 & 416 & 127 \\ 61 & 22 & 107 & 40 & 14 & 23 & 434 & 136 \\ 53 & 20 & 102 & 21 & 20 & 25 & 379 & 153 \\ 58 & 17 & 89 & 26 & 24 & 25 & 430 & 156 \\ 49 & 19 & 106 & 30 & 15 & 19 & 447 & 146 \\ 64 & 19 & 124 & 28 & 21 & 26 & 435 & 179 \\ 47 & 15 & 112 & 30 & 23 & 25 & 430 & 192 \\ 39 & 18 & 129 & 35 & 26 & 21 & 457 & 202 \\ 46 & 20 & 151 & 43 & 28 & 43 & 422 & 217 \\ 43 & 24 & 152 & 54 & 81 & 44 & 457 & 196 \\ 47 & 19 & 177 & 35 & 53 & 34 & 463 & 200 \\ 40 & 25 & 196 & 35 & 47 & 33 & 529 & 212 \end{pmatrix}$$

Para encontrar el estimador de la matriz  $\beta$  se utiliza el mismo método con el que se resuelve el modelo univariado de regresión lineal, el método de los mínimos cuadrados del error, cuya fórmula es

$$\hat{\beta} = (Z^T Z)^{-1} Z^T Y$$

Fórmula 5.5

Si reemplazamos las matrices 5.7 y 5.8 en la fórmula anterior, obtenemos los valores estimados de las constantes  $\beta$ 's para predecir la cantidad de pacientes masculinos que se atenderán en Solca de Guayaquil en los próximos años. Esta matriz se la describe a continuación:

**Matriz 5.10** Matriz de estimadores de los parámetros desconocidos para encontrar el número de pacientes masculinos

$$\hat{\beta} = \begin{bmatrix} 3460.34 & 1010.25 & -1119.45 & -2063.99 & -6176.13 & -9474.94 & -14374.22 \\ -1.7167 & -0.493 & 0.5769 & 1.06643 & 3.11888 & 4.7797 & 7.3286 \end{bmatrix}$$

**Matriz 5.11** Matriz de los valores estimados de Y (número de pacientes masculinos con cáncer)

$$\hat{Y} = \begin{bmatrix} 52.53 & 31.63 & 25.64 & 52.88 & 14.85 & 12.79 & 173.19 \\ 50.81 & 31.14 & 26.32 & 53.95 & 17.97 & 17.57 & 180.52 \\ 49.09 & 30.64 & 26.90 & 55.02 & 21.08 & 22.35 & 187.85 \\ 47.38 & 30.15 & 27.47 & 56.08 & 24.20 & 27.13 & 195.18 \\ 45.66 & 29.66 & 28.05 & 57.15 & 27.32 & 31.91 & 202.51 \\ 43.94 & 29.16 & 28.63 & 58.22 & 30.44 & 36.69 & 209.84 \\ 42.22 & 28.67 & 29.21 & 59.28 & 33.56 & 41.47 & 217.16 \\ 40.51 & 28.18 & 29.78 & 60.35 & 36.68 & 46.25 & 224.49 \\ 38.69 & 27.68 & 30.36 & 61.42 & 39.80 & 51.03 & 231.82 \\ 37.07 & 27.19 & 30.94 & 62.48 & 42.92 & 55.81 & 239.15 \\ 35.36 & 26.70 & 31.51 & 63.55 & 46.03 & 60.59 & 246.48 \\ 33.64 & 26.21 & 32.1 & 64.62 & 49.15 & 65.37 & 253.81 \end{bmatrix}$$

**Matriz 5.12** Matriz de los errores estimados

$$\hat{\varepsilon} = \begin{bmatrix} -1.53 & 2.37 & -3.74 & 2.12 & -4.85 & 10.21 & -16.19 \\ 8.19 & -10.14 & -3.32 & 11.05 & -0.97 & 9.43 & 1.48 \\ -5.09 & -2.64 & 10.10 & -2.02 & 11.92 & -0.35 & 9.15 \\ -8.38 & 1.85 & -0.47 & -0.08 & 4.8 & -11.13 & 11.82 \\ 6.34 & 1.34 & -0.05 & -4.15 & 0.68 & -10.91 & -1.51 \\ 6.06 & 15.84 & -5.63 & -15.22 & -12.44 & -1.69 & 7.16 \\ 0.78 & -1.67 & -4.21 & -5.28 & -2.56 & -1.47 & -4.16 \\ 6.5 & 2.82 & 3.22 & 6.65 & -9.68 & -15.25 & -8.50 \\ -10.79 & 4.32 & 5.64 & 2.58 & 13.20 & 3.97 & -8.82 \\ -16.07 & -15.19 & 9.06 & 2.52 & -2.92 & 9.19 & 12.85 \\ 1.64 & -5.70 & -5.51 & -7.55 & 4.97 & 10.41 & 9.52 \\ 12.36 & 6.79 & -5.10 & 9.38 & -2.15 & -2.37 & -12.81 \end{bmatrix}$$

Si reemplazamos las matrices 5.8 y 5.10 en la fórmula anterior, obtenemos los valores estimados de las constantes  $\beta$ 's para predecir la cantidad de pacientes femeninos que se atenderán en Solca de Guayaquil en los próximos años. Esta matriz se la describe a continuación:

**Matriz 5.13** Matriz de estimadores de los parámetros desconocidos para encontrar el número de pacientes masculinos

$$\hat{\beta} = \begin{pmatrix} 24643 & -73216 & -167489 & -319634 & -821677 & -362614 & -137216 & -158173 \\ -1.21 & 0.378 & 8.48 & 1.62 & 4.144 & 1.836 & 7.12 & 8.035 \end{pmatrix}$$



**Matriz 5.14** Matriz de los valores estimados de Y (número de pacientes femeninos con cáncer)

$$\hat{Y} = \begin{bmatrix} 55.92 & 17.42 & 81.95 & 24.08 & 7.79 & 17.65 & 40245 & 13214 \\ 54.71 & 17.80 & 90.43 & 25.70 & 11.94 & 19.49 & 40956 & 14018 \\ 53.50 & 18.18 & 98.91 & 27.32 & 16.08 & 21.33 & 41668 & 14821 \\ 52.28 & 18.56 & 107.39 & 28.94 & 20.22 & 23.16 & 42379 & 15625 \\ 51.07 & 18.93 & 115.86 & 30.57 & 24.37 & 24.99 & 43091 & 16428 \\ 49.86 & 19.31 & 124.34 & 32.18 & 28.51 & 26.83 & 43803 & 17232 \\ 48.64 & 19.69 & 132.82 & 33.81 & 32.66 & 28.67 & 44514 & 18035 \\ 47.43 & 20.07 & 141.30 & 35.43 & 36.80 & 30.50 & 45226 & 18839 \\ 46.22 & 20.44 & 149.78 & 37.06 & 40.94 & 32.34 & 45937 & 19642 \\ 45.00 & 20.82 & 158.26 & 38.68 & 45.09 & 34.17 & 46649 & 20446 \\ 43.79 & 21.20 & 166.74 & 40.30 & 49.23 & 36.01 & 47360 & 21249 \\ 42.58 & 21.58 & 175.22 & 41.92 & 53.37 & 37.85 & 48072 & 22053 \end{bmatrix}$$

**Matriz 5.15** Matriz de los errores estimados

$$\hat{e} = \begin{bmatrix} -11.92 & -1.42 & 16.05 & -5.08 & 7.21 & -2.65 & 13.55 & -5.14 \\ 6.29 & 4.20 & 16.57 & 14.30 & 2.06 & 3.51 & 24.44 & -4.18 \\ -0.50 & 1.82 & 3.09 & -6.32 & 3.92 & 3.67 & -37.68 & 4.79 \\ 5.72 & -1.56 & -18.39 & -2.94 & 3.78 & 1.84 & 6.21 & -0.25 \\ -2.07 & 0.07 & -9.86 & -0.57 & -9.37 & -5.99 & 16.09 & -18.28 \\ 14.14 & -0.31 & -0.34 & -4.19 & -7.51 & -0.83 & -3.03 & 6.68 \\ -1.64 & -4.69 & -20.82 & -3.81 & -9.66 & -3.67 & -15.14 & 11.65 \\ -8.43 & -2.07 & -12.30 & -0.43 & -10.80 & -9.50 & 4.74 & 13.61 \\ -0.22 & -0.44 & 1.21 & 5.94 & -12.94 & 10.66 & -37.37 & 20.58 \\ -2.00 & 3.18 & -6.26 & 15.32 & 35.91 & 9.83 & -9.49 & -8.46 \\ 3.21 & -2.20 & 10.26 & -5.30 & 3.77 & -2.01 & -10.60 & -12.49 \\ -2.58 & 3.42 & 20.78 & -6.92 & -6.37 & -4.85 & 48.28 & -8.53 \end{bmatrix}$$

Al hallar los estimadores de las constantes  $\beta$ 's del modelo, podemos estimar la matriz de respuesta  $Y$  para un valor de  $Z_0$ . Si queremos estimar el número de pacientes masculinos que habrán en 1999, clasificados según el tipo de carcinomas que sufran podemos definir a  $Z_0$  como  $[1 \quad 1999]$  y lo multiplicamos por la matriz estimada de los  $\beta$ 's correspondiente. Así encontramos la respuesta  $Y$  estimada

$$\hat{Y}_0 = Z_0 \hat{\beta}$$

Fórmula 5.6

**Matriz 5.16** Matriz del número estimado de pacientes masculinos para 1999

$$\hat{Y}_0 = \begin{bmatrix} 28.49 & 24.72 & 33.82 & 67.81 & 58.51 & 79.71 & 275.79 \end{bmatrix}$$

Ya que los números representados en la matriz de respuesta estimada son número de pacientes masculinos, se los redondea; por lo tanto para el año 1999 se estima que habrá:

29 pacientes masculinos con carcinoma de piel,

25 pacientes masculinos con linfomas,

34 pacientes masculinos con carcinoma de pulmón,  
 68 pacientes masculinos con carcinoma de estómago,  
 59 pacientes masculinos con carcinoma leucemia,  
 80 pacientes masculinos con carcinoma de próstata, y  
 276 pacientes masculinos con otros carcinomas.

Si queremos estimar el número de pacientes femeninos que habrán en 1999, clasificados según el tipo de carcinomas que sufran podemos definir a  $Z_o$  de la misma forma que para el caso de pacientes masculinos y lo multiplicamos por la matriz de estimada de los  $\beta$ 's correspondiente. Así encontramos la respuesta  $Y$  estimada

**Matriz 5.17** Matriz del número estimado de pacientes femeninos para 1999

$$\hat{Y}_o = \begin{bmatrix} 38.94 & 22.71 & 200.65 & 46.79 & 65.80 & 43.35 & 502.06 & 244.63 \end{bmatrix}$$

Ya que los números representados en la matriz de respuesta estimada son número de pacientes femeninos, se los redondea; por lo tanto para el año 1999 se estima que habrá:

39 pacientes femeninos con carcinoma de piel,  
 23 pacientes femeninos con linfomas,

201 pacientes femeninos con carcinoma de mamas,  
 47 pacientes femeninos con carcinoma de estómago,  
 66      pacientes femeninos con carcinoma leucemia,  
 44      pacientes femeninos con carcinoma de ovario, y  
 503     pacientes femeninos con cervix  
 245     pacientes femeninos con otros carcinomas.

Al seguir el mismo procedimiento para los años 1997 y 1998, obtenemos las siguientes matrices como resultado:

**Matriz 5.18** Matriz de respuestas de número de pacientes masculinos para el año 1997

$$\hat{Y}_0 = [31.92 \quad 25.71 \quad 32.67 \quad 65.68 \quad 52.27 \quad 70.15 \quad 261.17]$$

**Matriz 5.19** Matriz de respuestas de número de pacientes femeninos para el año 1997

$$\hat{Y}_0 = [41.36 \quad 21.95 \quad 183.69 \quad 43.54 \quad 57.51 \quad 39.68 \quad 487.83 \quad 228.56]$$

**Matriz 5.20** Matriz de respuestas de número de pacientes femeninos para 1998

$$\hat{Y}_0 = [40.15 \quad 22.33 \quad 192.18 \quad 45.17 \quad 61.66 \quad 41.52 \quad 494.95 \quad 236.59]$$



**Matriz 5.21** Matriz de respuestas de número de pacientes masculinos para 1998

$$\hat{Y}_0 = [30.21 \quad 25.22 \quad 33.24 \quad 66.75 \quad 55.39 \quad 74.93 \quad 268.47]$$

### 5.3.2 Verificación del modelo de regresión.

Para verificar que el modelo se ajusta a la realidad se debe de hallar el vector de los índices de ajustes, el cual se lo halla por medio de la siguiente fórmula:

Donde

$$R^2 = \frac{SCR}{SCT} \quad \text{Fórmula 5.7}$$

$$SCR(\text{Suma Cuadrática de la Regresión}) = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 \quad \text{Fórmula 5.8}$$

$$SCT(\text{Suma Cuadrática Total}) = \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad \text{Fórmula 5.9}$$

Así obtenemos que el vector de índices de ajustes está dado por:

**Matriz 5.22** Vector de índices de ajustes para pacientes masculinos

$$R^2 = \begin{bmatrix} 0.3416 & 0.0467 & 0.1190 & 0.2108 & 0.6791 & 0.7854 & 0.8719 \end{bmatrix}$$

**Matriz 5.23** Vector de índices de ajustes para pacientes femeninos

$$R^2 = \begin{bmatrix} 0.2906 & 0.2061 & 0.7879 & 0.364 & 0.557 & 0.538 & 0.5198 & 0.8610 \end{bmatrix}$$

El índice de ajuste  $R^2$  nos indica en qué porcentaje las variables independientes predicen a la variable dependiente. Podemos observar que en el caso de los pacientes masculinos, el modelo se ajusta sólo en un 4.67% para los linfomas y en un 11.9% para el cáncer del pulmón. Al tener un índice de ajuste tan bajo, significa que estos linfomas no van a ser predecidos eficazmente.

El estimador de la varianza del error está dado por:

$$E(\text{MCE}) = \sigma^2$$

Fórmula 5.10

Donde

$$\text{MCE} = \frac{\text{SCE}}{n - p} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - p}$$

Fórmula 5.11

De donde obtenemos



**Matriz 5.24** Vector de medias cuadráticas del error para pacientes masculinos

$$MCE = \begin{bmatrix} 81.318 & 71.026 & 35.721 & 60.9649 & 65.7152 & 89.179 & 112.831 \end{bmatrix}$$

**Matriz 5.25** Vector de medias cuadráticas del error para pacientes femeninos

$$MCE = \begin{bmatrix} 51.36 & 7.8627 & 211.5 & 65.73 & 195.2 & 41.431 & 668.89 & 149.05 \end{bmatrix}$$

Cada uno de estos valores es la varianza estimada de los errores de predicción el modelo.

A continuación, presentamos los valores reales de las personas con carcinomas tratadas en Solca de Guayaquil, por sexo, comparado con los valores predichos por el modelo y con el porcentaje de error en base al valor real.

TABLA LXXVIII

COMPARACIÓN DE NÚMERO DE PACIENTES PRONOSTICADOS Y REAL DE 1997 Y 1998

1997								
	Mujer							
	Piel	Linf.	Mama	Estom	Leuc.	Ovario	Cervix	Otros
Real	65	23	238	36	36	34	466	368
Pronosticado	41.36	21.95	183.69	43.54	57.51	39.68	487.83	228.56
Error	23.64	1.05	54.31	-7.54	-16.51	-5.68	-21.83	139.44

% error	36%	5%	23%	-21%	-40%	-17%	-5%	38%
---------	-----	----	-----	------	------	------	-----	-----

	Hombre						
	Piel	Linf.	Pulmón	Estom	Leuc.	Próstata	Otros
Real	77	27	40	74	43	112	285
Pronosticado	31.92	25.71	32.67	65.58	52.27	70.15	261.17
Error	45.08	1.29	7.33	8.42	-9.27	41.85	23.83
% error	59%	5%	18%	11%	-22%	37%	8%

1998

	Mujer							
	Piel	Linf.	Mama	Estom	Leuc.	Ovario	Cervix	Otros
Real	71	28	209	65	51	34	385	1294
Pronosticado	40.15	22.33	192.18	45.17	61.66	41.52	494.95	236.59
Error	30.85	5.67	16.82	19.83	-10.66	-7.52	-109.95	1057.41
% error	43%	20%	8%	31%	-21%	-22%	-29%	82%

	Hombre						
	Piel	Linf.	Pulmon	Estom	Leuc.	Próstata	Otros
Real	76	27	30	93	50	100	291
Pronosticado	30.21	25.22	33.24	66.75	55.39	74.93	268.47
Error	45.79	1.78	-3.24	26.25	-5.39	25.07	22.53
% error	60%	7%	-11%	28%	-11%	25%	8%

Donde Linf son linfomas, esto significa estómago y leuc significa leucemia.

Los errores en leucemia de hombres son menores a los errores en leucemia de mujeres.

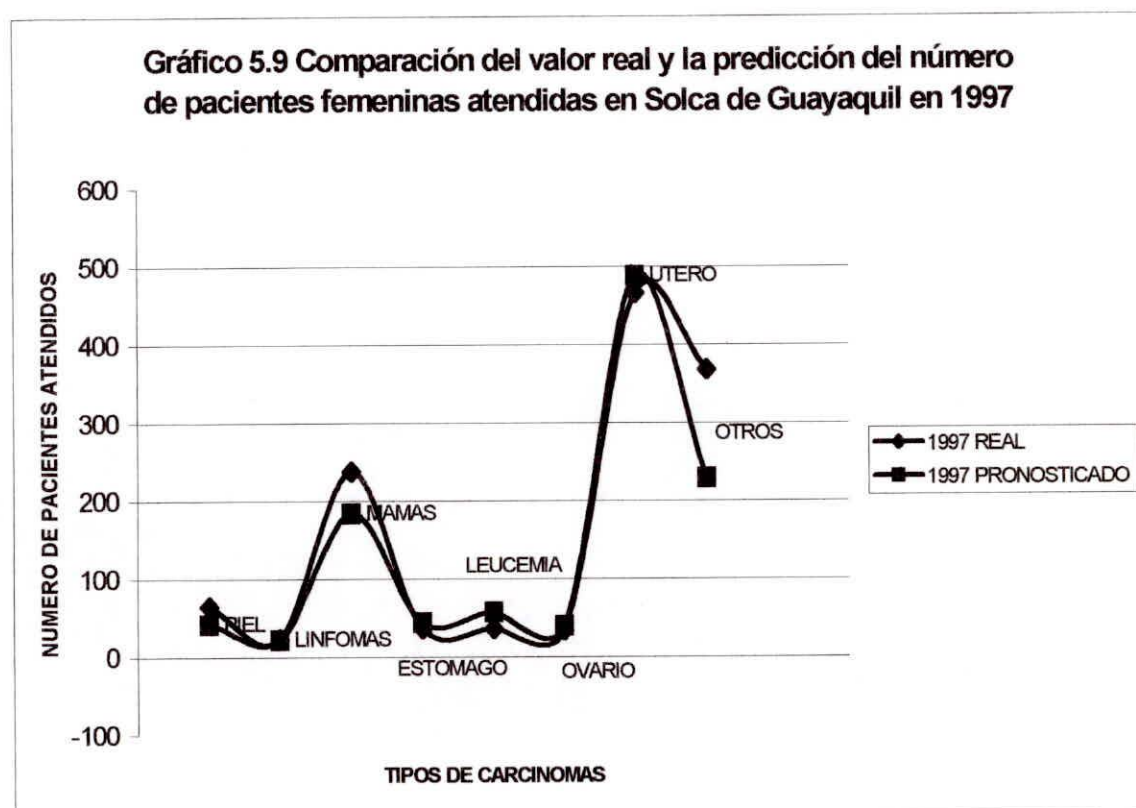
Los errores en la predicción de personas con linfomas es bastante bajo ( entre el 5 y el 20 por ciento).

El error en la cantidad de mujeres con carcinoma de cervix es bastante bajo en el año 1997 (-5%)

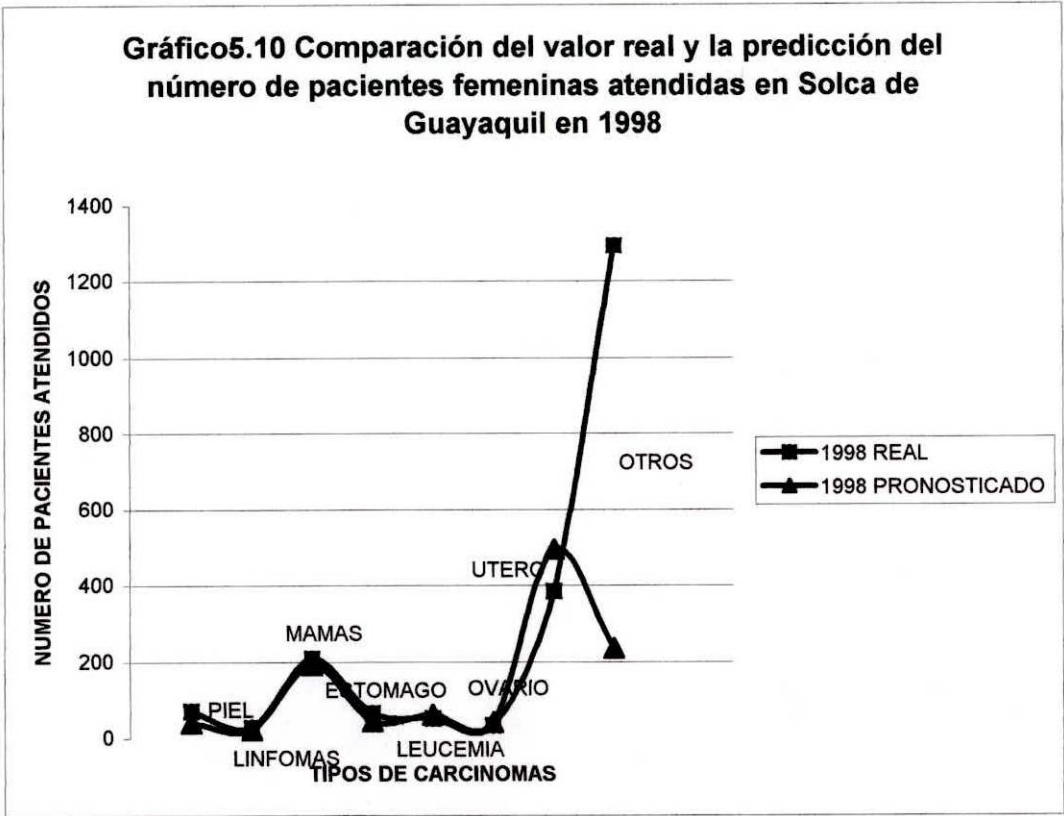
El error en la cantidad de mujeres con carcinoma de mamas es bastante bajo en el año 1998 (-8%)

El error para el año 1998 de personas con leucemia es menor al error en el año 1997.

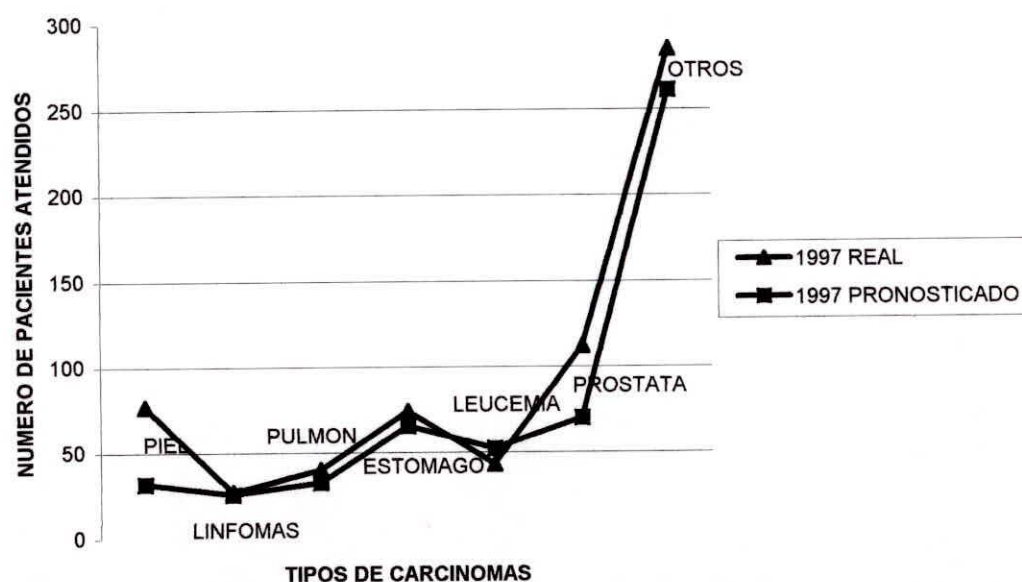
A continuación se muestra el gráfico de las predicciones y de los valores reales para cada caso



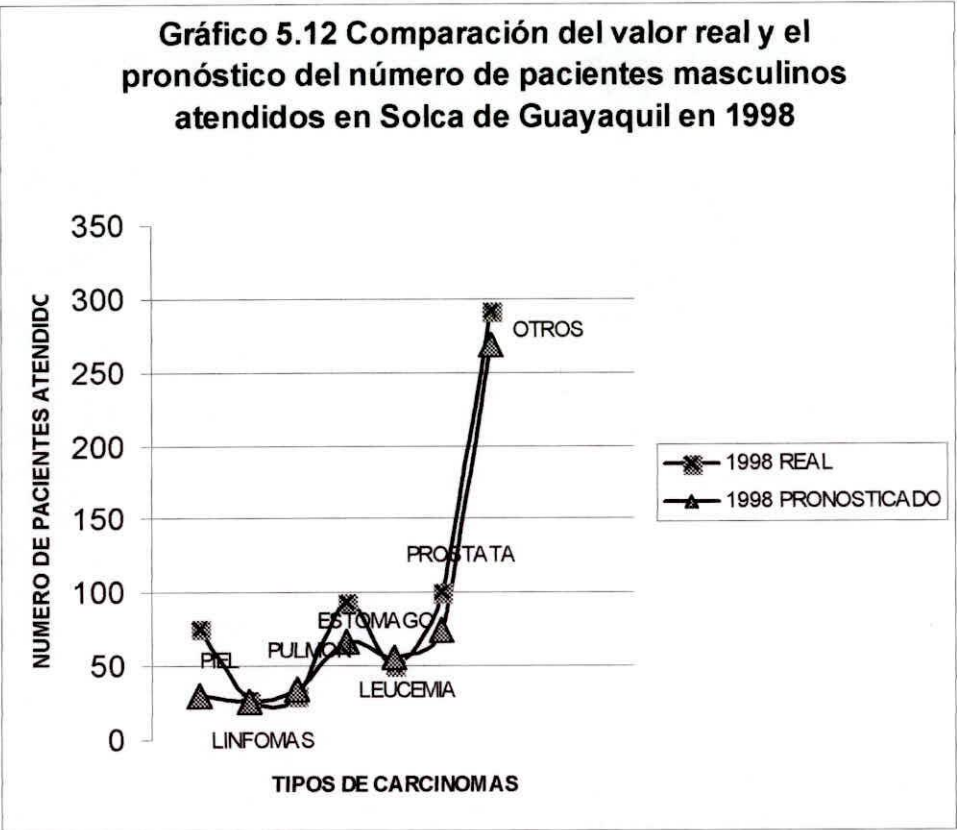
**Gráfico5.10 Comparación del valor real y la predicción del número de pacientes femeninas atendidas en Solca de Guayaquil en 1998**



**Gráfico 5.11 Comparación del valor real y la predicción del número de pacientes masculinos atendidos en Solca de Guayaquil en 1997**



**Gráfico 5.12 Comparación del valor real y el pronóstico del número de pacientes masculinos atendidos en Solca de Guayaquil en 1998**





# CONCLUSIONES

1. Al comparar los tres principales tipos de carcinomas representativos que afectan a la población Europea, Estadounidense y Ecuatoriana en el año de 1990, podemos notar que hay grandes diferencias. A continuación mostramos un cuadro comparativo en porcentajes de los principales carcinomas de estas poblaciones:

Ecuador		Estados unidos		Europa	
		Hombres			
Piel	11.60	Próstata	28.9	Pulmón	20.7
Linfomas	10.44	Pulmón	16.3	Colon y recto	13.2
Estómago	9.98	Colon y recto	11.81	Próstata	12.3
Próstata	8.12	Vejiga	6.24	Vejiga	7.39
Pulmón	5.34	Linfoma no hodkin	3.95	Estómago	6.59
Otros	54.00	Otros	32.8	Otros	39.9

### Mujer

Útero	48.55	Mamas	31.12	Mamas	27.8
Mamas	13.84	Colon y recto	13.55	Colon y recto	14.8
Piel	7.14	Pulmón	12.6	Pulmón	5.6
Estómago	3.13	Utero	8.87	Útero	5.13
Ovario	3.00	Ovario	4.36	Estómago	5.02
Otros	24.00	Otros	29.5	Otros	41.7

Con respecto a los varones se nota una gran diferencia entre la población ecuatoriana y la de los países desarrollados. Los carcinomas prostáticos, pulmonares y rectales son los que más afectan a los varones en países desarrollados, esto se debe al incremento en el consumo de tabaco y dietas altas en grasas y con poca fibra. En Ecuador el carcinoma más común en el hombre es el dermatológico, seguido por los linfomas y el cáncer estomacal. Ya que un factor de riesgo para el carcinoma de piel es la exposición severa al sol, podremos atribuir a la ubicación geográfica de nuestro país como un factor que influye en el tipo de cáncer que más afecta a la población masculina. Otra causa entre la diferencia podría ser la idiosincrasia de los hombres en nuestro país, que no se realizan el examen prostático y por lo tanto no se les detecta el cáncer en esta glándula, y al fin y al cabo mueren sin saber que están desarrollando esta enfermedad; muy al contrario de los hombres en países más avanzados, que se preocupan por realizarse este

examen pasada cierta edad para poder dar tratamiento adecuado para sus problemas prostáticos.

En el caso de las mujeres, en los países desarrollados encontramos características muy parecidas con respecto a los tres principales carcinomas que las afectan, los cuales son mamas, colon y pulmón. Las costumbres alimenticias en estos países no son de mucho beneficio para erradicar estos tipos de carcinomas; peor aún el consumo de sustancias dañinas al organismo como cigarrillo, alcohol y drogas. Las mujeres en nuestro país en cambio son más propensas al carcinoma uterino, seguido del mamario y el dermatológico. Una vez más se resalta el factor solar en la población ecuatoriana y su influencia en el porcentaje de personas con carcinoma de piel. Es válido analizar el caso del cáncer uterino más a fondo, ya que del total de la población femenina con cáncer, casi el 50% sufre de este tipo de carcinoma. De todos los factores de riesgo, ¿cuál es el que diferencia a la mujer ecuatoriana de las europeas o las estadounidenses? Realmente, las relaciones sexuales tempranas y la presencia infecciones de transmisión sexual han incrementado alrededor de todo el mundo; no sólo en los países desarrollados, aunque nos duela aceptarlo. El tener varios compañeros sexuales es una costumbre más de mujeres anglosajonas que latinas por el machismo de nuestra sociedad; aunque nunca falta una excepción a la regla. El único factor que nos diferencia es la abundancia de hijos.

2. Al realizar la comparación entre los niños de Ecuador y de Estados Unidos que sufren de carcinoma en el período desde 1991 a 1995, observamos los siguientes datos en porcentajes:

Ecuador				Estados Unidos			
	0-4	5-9	10-14		0-4	5-9	10-14
Leucemia	51.39	49.42	45.33	Leucemia	36.02	34.00	22.59
Ojo	15.01	9.23	15.11	SN	31.01	30.69	22.25
Ganglios	12.40	20.77	20.89	PBL	10.56	12.42	24.50
SN	10.11	14.81	6.89	Riñón	9.22	4.85	0.90
Otros	10.60	6.15	12.67	Otros	13.20	18.05	29.75

- Los niños de todas las edades, tanto de Ecuador como de Estados Unidos , sufren principalmente de leucemia.
  - Los niños de países desarrollados tienen una alta incidencia de carcinoma en el sistema nervioso y de piel (partes blandas), a diferencia de los niños ecuatorianos, que tienen un mayor índice en carcinomas oculares y en el sistema linfático (ganglios).
3. Pasando al análisis local de los niños con carcinomas, la ciudad donde vive el niño no influye en el tipo de carcinoma que este tenga, mientras

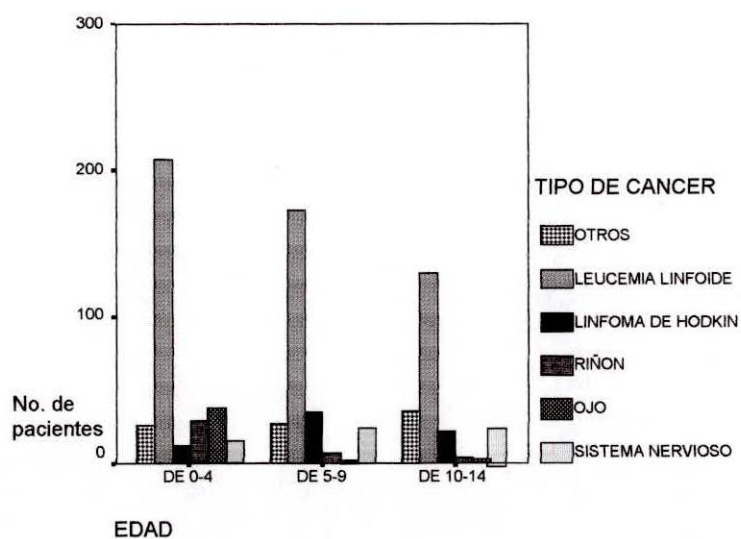
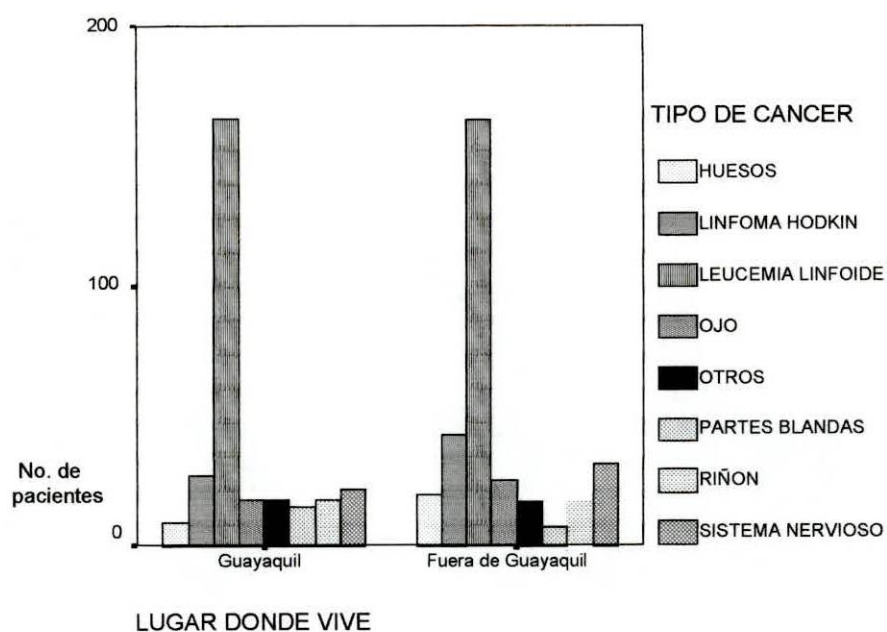


que su edad y su sexo sí influyen. A continuación presentamos una tabla con el valor p por los cuales se rechazó o aceptó estas hipótesis a un 95% de confianza:

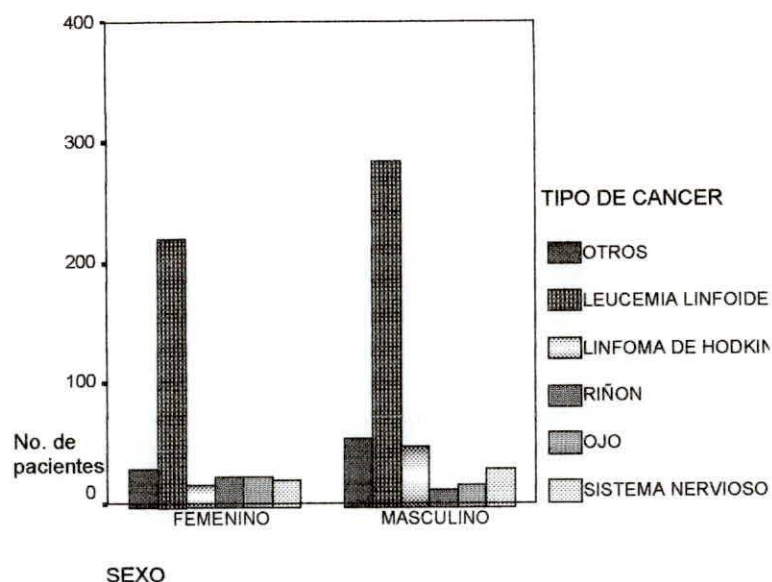
Tipo de cáncer versus...	Valor P
Sexo	0.002
Edad	0.000
Ciudad	0.107

Ya que trabajamos a un nivel de confianza del 95% para aceptar o rechazar la hipótesis de independencia entre variables, rechazamos la hipótesis de independencia entre las variables tipo de cáncer versus sexo y tipo de cáncer versus edad, pero aceptamos que existe evidencia de independencia entre el tipo de cáncer y la ciudad donde vive el niño.

Por ejemplo, es más probable que un niño entre 0 y 4 años tenga un carcinoma en el ojo, que un niño mayor; así mismo la leucemia afecta más a los niños menores de 9 años.





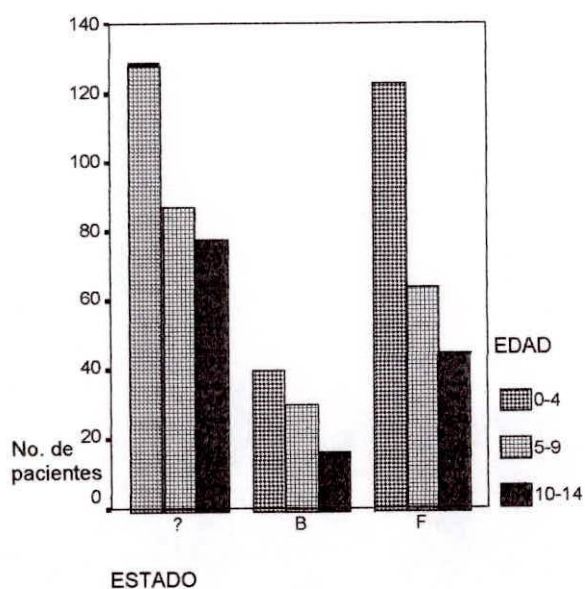


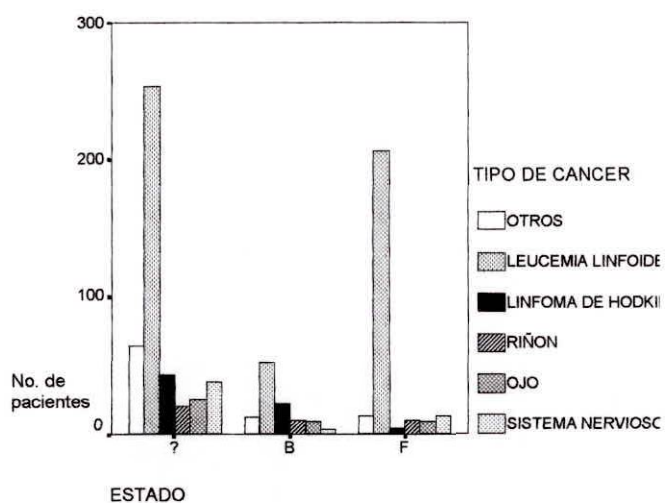
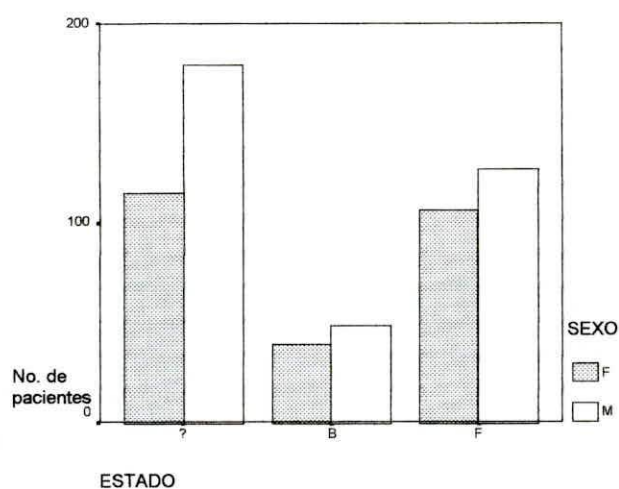
4. La edad en que al niño se le presenta la enfermedad y su sexo no influyen en su estado actual, mientras que el tipo de cáncer que padezcan y el lugar donde viva sí, estadísticamente hablando. De 443 niños que no se sabe su paradero, 186 viven en una ciudad diferente al tratamiento, y de los 108 que se encuentran en buen estado, 52 son de Guayaquil.

A continuación presentamos una tabla con el valor p por los cuales se rechazó o aceptó estas hipótesis a un 95% de confianza:

Estado versus...	Valor P
Sexo	0.297
Edad	0.134
Lugar donde vive	0.000
Tipo de cáncer	0.000

Al analizar los valores p de las hipótesis de independencia de variables a un 95% de confianza, tenemos evidencia estadística para aceptar que las variables estado versus lugar donde vive y estado versus tipo de cáncer son dependientes, al contrario de la evidencia de independencia entre estado y sexo y estado y edad.





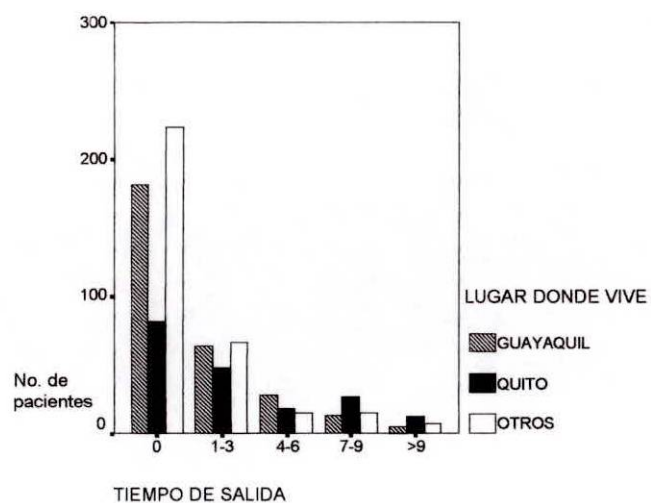
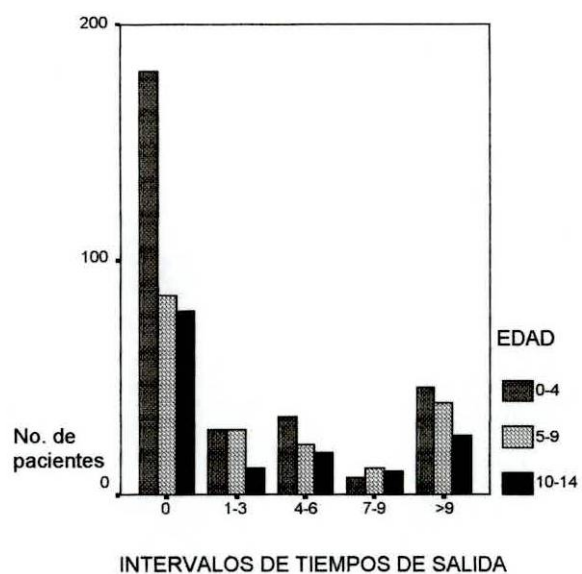
5. EL lugar donde se viva y la edad de los pacientes influyen mucho en el tiempo de salida de este, puesto que mientras menor es el niño o viva más alejado de la ciudad, hay mayor probabilidad de que los padres se lo lleven de la institución; mientras que el sexo no influye en nada en el tiempo que permanecen en tratamiento. Por ejemplo, de 343 niños que

egresaron en el año 0 al 1, 180 son de 0 a 4 años, mientras que 85 son de 5 a 9, y 78 son de 10 a 14. Los niños que viven fuera de Guayaquil son más propensos a dejar el tratamiento en el primer año, ya que de los 487 que lo hacen, 224 no viven en la ciudad que se le da el tratamiento.

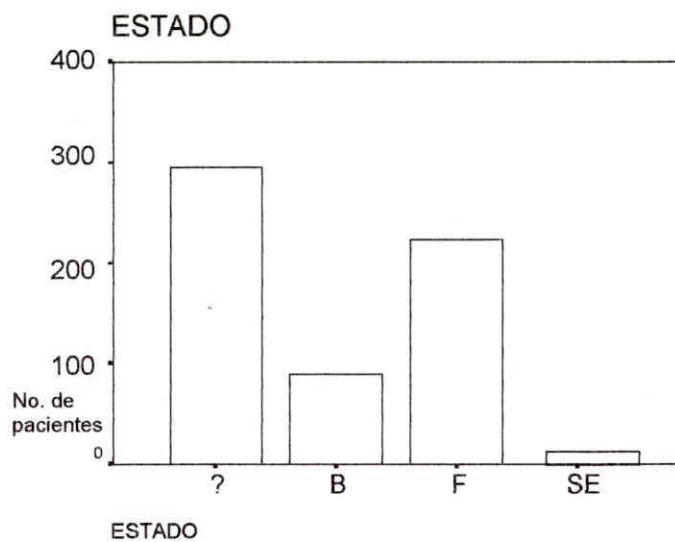
A continuación presentamos una tabla con el valor p por los cuales se rechazó o aceptó estas hipótesis a un 95% de confianza:

Tiempo de salida versus...	Valor P
Edad	0.024
Lugar donde vive	0.000
Sexo	0.065

Según los valores p calculados, hay suficiente evidencia estadística para aceptar la hipótesis que las variables tiempo de salida versus edad y versus lugar donde se vive son dependientes, mientras que las variables tiempo de salida y sexo son estadísticamente independientes.



6. En Guayaquil y Quito hay una alta incidencia de abandonos en el tratamiento (49.7%). Del resto de la población que continúa en el tratamiento, todavía continúan con vida el 29% de esta, es decir el 14.58% del total de la población.



Al preguntar al Dr. Paulson, el cual trató a la mayoría de estos niños, cuál en su opinión es el índice de abandono en estos niños, él responde que es de un 55%, lo cual no está alejado de la realidad del 49.7%. El da como principales razones de este fenómeno a la ignorancia de los padres, la situación económica de la familia, por idiosincrasia o por factores sociales.

7. Las probabilidades de los principales carcinomas que afectan a los niños de Quito y Guayaquil son detallados en la siguiente tabla:



Quito	Probabilidad de supervivencia al quinto año de tratamiento.
Leucemia	0.5192
Guayaquil	
Leucemia	0.1707
Linfoma de hodkin	0.8462

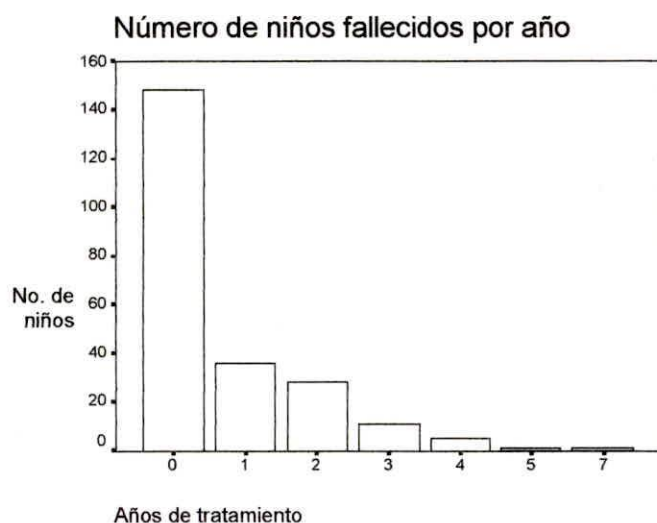
Al hablar con el Ab. Eduardo Vergara Saltos, secretario general de la Sociedad de Lucha contra el Cáncer, asegura que la diferencia de índices de mortalidad se debe a que se halla un fenómeno inmigratorio de los pacientes tratados en Quito, ya que estos se mudan a Guayaquil para obtener mejores tratamientos y mueren aquí, factor que aumenta el índice de mortalidad en nuestra ciudad.

En concordancia con la base de datos que se encuentra en el anexo 1 de este trabajo, no se encontró repetición de pacientes que se hayan atendido primero en Quito y después en Guayaquil, en el periodo desde 1985 hasta 1994. En mi opinión, más que este factor inmigratorio, el hecho de haber más de tres instituciones especializadas en tratar el cáncer de niños es una ventaja para los habitantes de la ciudad de Quito, puesto que se da atención más personalizada al paciente y a sus familiares, y por lo tanto hay más información disponible. En Guayaquil, la única institución encargada de la difusión de la prevención y tratamiento

para el cáncer es Solca, por lo que se requiere con urgencia capacitar al personal de otras instituciones para que detecten con eficacia síntomas de posibles carcinomas y cómo tratarlos, según explica el Dr. Vergara.

8. La etapa más difícil en el tratamiento de la leucemia en niños es el primer año, puesto que tiene un mayor índice de muerte, el cual va disminuyendo a medida que pasa el tiempo. La razón principal de este factor es que el 90% de los pacientes que ingresan a las instituciones tratantes se encuentran en estado de desnutrición; y de este 90%, 72% se encuentran en etapas avanzadas del cáncer; lo que impide el tratamiento y la erradicación de este mal a tiempo, según explica el doctor tratante.

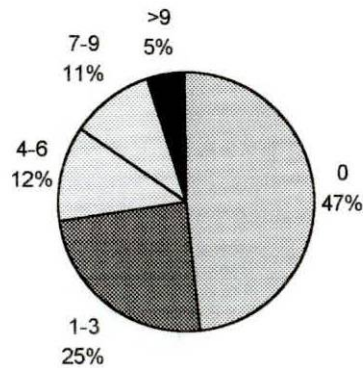
Tiempo de supervivencia	Número de niños	Porcentaje
0	148	64.3
1	36	15.7
2	28	12.2
3	11	4.8
4	5	2.2
5	1	0.4
6	1	0.4
Total	230	100.0



Al observar el gráfico del número de niños fallecidos por año, podemos notar que existe un comportamiento exponencial en los datos. Al realizar la prueba de Kolmogorov Smirnov, al 95% de confianza, encontramos que el p de la prueba es de 0.610, lo que implica que la hipótesis de que los datos tengan una distribución exponencial es correcta.

9. El 47% de los pacientes infantiles egresan de las instituciones tratantes en el primer año, sea en calidad de fallecidos o de retirados del tratamiento.

**Tiempos de salida de niños de Quito y Guayaquil,  
en porcentajes**



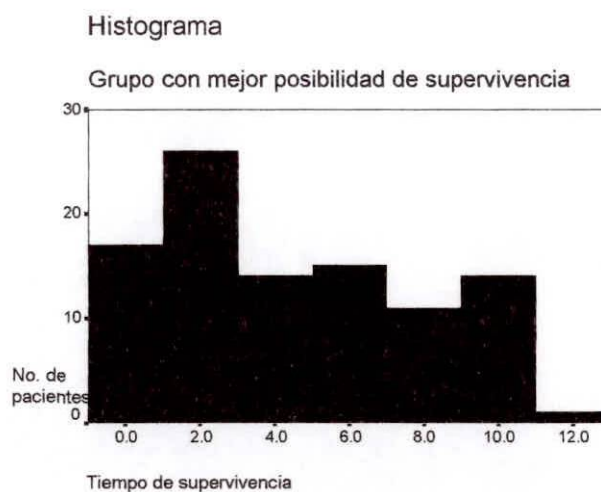
10. En general, el modelo de regresión se ajustó a los datos y las predicciones tuvieron un índice de error bajo, a excepción de la categoría otros carcinomas de pacientes femeninos en el año 1998.

Se espera que los valores reales de los pacientes con cáncer atendidos en Solca superen la cantidad pronosticada, a excepción de los carcinomas de cervix, ovario y leucemia en las mujeres, y leucemia en los varones; cuyos errores han resultado negativos, tanto en 1997 como en 1998. Esto se debe al error en el modelo de predicción.

11. Por el estudio realizado en el capítulo 3 del presente trabajo, se agrupó a

los niños tratados en Quito y Guayaquil en tres grupos según el tiempo de supervivencia que tuvieron con las diferentes combinaciones de los tratamientos recibidos:

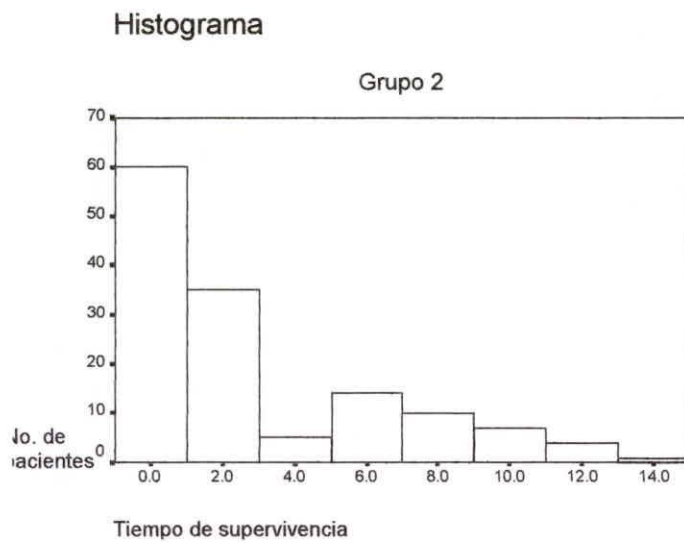
- El grupo conformado por las combinaciones de tratamiento: quimioterapia y cirugía o quimioterapia y radioterapia, los cuales estadísticamente tienen el mismo efecto en el tiempo de supervivencia proporcionado al individuo, es el que tiene un tiempo de supervivencia esperado mayor al resto de combinaciones de tratamientos. El tiempo de supervivencia esperado para este grupo es de 4.01 años y una desviación estándar de 3.33 años.



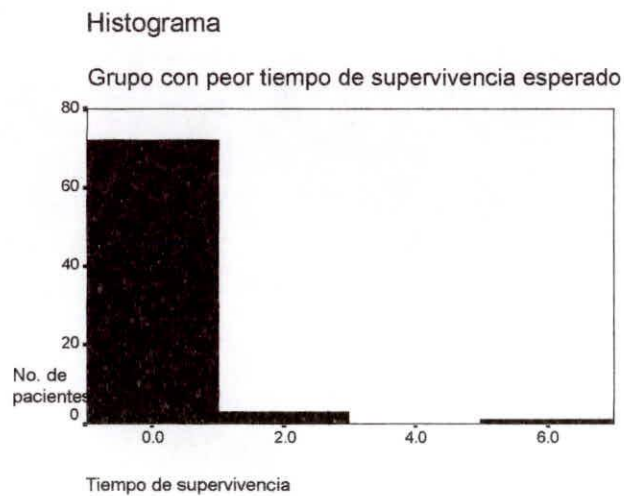
- El grupo conformado por los que recibieron tratamiento con sólo radioterapia, radioterapia y cirugía, sólo cirugía, sólo quimioterapia o los tres tratamientos juntos es el segundo grupo con mayor



tiempo esperado de supervivencia, el cual es de 2.56 años con una desviación estándar de 3.49 años



- El grupo que no recibe tratamiento, el cual tiene el tiempo de supervivencia esperado más bajo de todos, el cual es de 0.11 años con una desviación estándar de 0.6 años.





# RECOMENDACIONES

1. Se debería informar de una manera más eficaz a la población ecuatoriana las posibles causas que producen el cáncer y los síntomas principales que evidencia la presencia de los principales tipos que afectan a la población, como son el carcinoma en el útero y en las mamas, en las mujeres; y el carcinoma prostático, de piel y de estómago en los varones.

La Sociedad de Lucha contra el Cáncer tiene programas de información a la comunidad dirigidos por la Sociedad de Damas Voluntarias de Solca, los cuales son básicamente dos:

- Dar conferencias a nivel de empresas con médicos residentes explicando cómo detectar a tiempo los carcinomas de mama, útero, próstata, pulmón y piel, los exámenes correspondientes, los factores de riesgo y los síntomas de cada uno de estos;
- El segundo programa consiste en que los residentes y/o médicos realizando sus práctica rurales realicen exámenes de detección del cáncer como el papanicolaou y el tacto rectal (prostático) gratis en la

parte rural de la provincia del Guayas. Estas muestras son analizadas en la institución para detectar si los pacientes sufren o no de estas dolencias y darles la información respectiva sobre el tratamiento a seguir.

Sugiero que se realice un programa de información a la comunidad através de los medios de comunicación más utilizados sobres los síntomas y factores de riesgos de los principales carcinomas que afectan a nuestra comunidad; y tratar de crear una cultura médica preventiva, la cual es menos costosa que la curativa que predomina en nuestro medio.

2. Se recomienda a las instituciones que son encargadas del tratamiento para esta enfermedad que noten lo imperante que es implementar procesos adecuados de recolección de datos de los pacientes atendidos, y así poder realizar un seguimiento de los que desertan del tratamiento y de esta forma poder comparar la supervivencia de estos con los que continúan el tratamiento. Además, esto ayudaría a realizar estudios más profundos para poder asociar algunas de nuestras costumbres con los carcinomas más comunes en la población, y el porqué de la diferencia con otras poblaciones a nivel mundial.

3. Se recomienda dar mayor cuidado y atención, tanto psicológica como física, a los niños menores de 4 años y a sus padres; puesto que estos niños son los que tienen un mayor índice de mortalidad y de abandono. Además, esto ayudaría a realizar estudios más profundos para poder asociar algunas de nuestras costumbres con los carcinomas más comunes en la población, y el porqué de la diferencia con otras poblaciones a nivel mundial.
4. Por la diferencia en el tiempo de supervivencia entre los pacientes con las diferentes combinaciones de tratamientos, se recomienda realizar un diseño experimental con tres factores de dos niveles cada uno, donde se pueda medir el efecto de estos tratamientos en el tiempo de supervivencia de los pacientes y confirmar el análisis de varianza realizado en el capítulo 3.
5. Sería interesante realizar un experimento entre personas que utilizaron tratamientos convencionales para erradicar esta mortífera enfermedad y las que utilizaron los tratamientos de medicina alternativa; como tratamientos onmiopáticos, naturistas, entre otros; y analizar como variable de respuesta al tiempo de supervivencia de cada grupo para establecer si estadísticamente los resultados difieren entre los dos grupos.

## **GLOSARIO**

<b>Adenoma</b>	Tumor epitelial benigno en el cual las células forman estructuras glandulares identificables.
<b>Anemia perniciosa</b>	Anemia megaloblástica que ocurre en niños con mayor frecuencia de edades ulteriores, causada por la incapacidad de absorción de la vitamina B12 por parte de la mucosa gástrica.
<b>Antígeno carcinoembrionario.</b>	Antígeno glucoproteínico específico del cáncer del carcinoma cólico, que también se presenta en muchos adenocarcinomas de origen endodérmico y en el tejido gastrointestinal normal de embriones humanos.
<b>Biopsias</b>	Extirpación y examen, usualmente microscópico, de tejido de un cuerpo vivo, para establecer un diagnóstico preciso.
<b>Braquiterapia</b>	En radioterapia, tratamiento con radiación ionizante aplicada a la superficie del cuerpo o localizada a corta distancia del área corporal sometida a tratamiento.
<b>Carcinomas in situ</b>	Entidad neoplásica en la cual las células tumorales están circunscritas al epitelio de origen, sin invasión de la membrana basal; generalmente se les da este nombre a las células en el cuello uterino
<b>Cistosomiasis.</b>	Obstrucción de la vejiga



<b>Displasias</b>	Anomalía del desarrollo, alteración de la forma, dimensión y organización de las células adultas.
<b>Endodérmico</b>	Relativo al endodermo
<b>Endodermo</b>	La más interna de las tres capas germinativas del embrión; de donde se deriva el epitelio de la faringe, aparato digestivo, aparato respiratorio, vejiga y uretra.
<b>Epitelio</b>	Revestimiento de las superficies internas y externas del cuerpo, que incluye el revestimiento de vasos y otras pequeñas cavidades.
<b>Escroto</b>	Bolsa que contiene los testículos y sus órganos accesorios.
<b>Isótopos</b>	Elemento químico que tiene el mismo número atómico que otro, pero que posee una masa atómica diferente.
<b>Laparoscopías.</b>	Examen del interior del abdomen mediante un laparoscopio.
<b>Laparotomías exploradoras</b>	Incisión quirúrgica para explorar zona afectada por el cáncer
<b>Leucemia</b>	Enfermedad maligna progresiva de los órganos formadores de sangre, que se caracteriza por proliferación y desarrollo deformado de leucocitos y sus precursores en la sangre y en la médula ósea.
<b>Linfoide</b>	Que es semejante a la linfa o al tejido del sistema linfático.
<b>Linfomas</b>	Término general que se aplica a cualquier trastorno neoplásico del tejido linfoide; generalmente es maligno.

<b>Metástasis, metastásico</b>	Transferencia de enfermedad de un órgano o una parte hacia otro sitio no directamente relacionado. Puede ser causada por transferencia de microorganismos o de células.
<b>Nefroblastoma</b>	Tumor de Willms, tumor que afecta al riñón.
<b>Neoplasia</b>	Formación de un neoplasma, es decir multiplicación progresiva de células bajo condiciones que no producirían la multiplicación de las células normales
<b>Neoplásicos</b>	Relativo a la neoplasia
<b>Pólipos</b>	Excrecencia mórbida o neoplasia que hace protrusión desde una mucosa.
<b>Recidivas</b>	Recaída o recurrencia en una enfermedad
<b>Retinoblastoma</b>	Tumor que surge de las células terminales de la retina.
<b>Testosterona</b>	Importante sustancia andrógena, producida por las células intersticiales del los testículos como reacción al estímulo de la hormona luteinizante de la hipófisis anterior.



## **APÉNDICES**

## **APÉNDICE 1**

### **BASE DE DATOS DE NIÑOS CON CÁNCER ATENDIDOS EN SOLCA, HOSPITAL VACA ORTIZ Y CRUZ ROJA DE QUITO**

# BIBLIOGRAFÍA

1. SCHEAFFER RICHARD L. *Estadística Matemática con Aplicaciones*, Grupo Editorial Iberoamérica
2. NETER JOHN, WASSERMAN WILLIAM, *Applied Linear Statistical Models*. Richard D. Irwin, Inc. Illinois, 1974.
3. SEBER G.A.F., *Linear Regression Analysis*. John Wiley and sons, inc. New York, 1977.
4. HOLLANDER MYLES, WOLF DOUGLAS A., *Nonparametrical Statistical Methods*, John Willy and sons, inc. Estados Unidos, 1973.
5. REGISTRO NACIONAL DE TUMORES. *Cáncer en Quito 1995 y 1996*. AH/editorial. Quito, 1998.

6. DEPARTAMENTO DE ESTADÍSTICA DE SOLCA, *Historias clínicas y reportes generados*. Guayaquil, 1985-1998.
7. *Archivos de historias clínicas de la Cruz Roja de Quito*, 1985-1996.
8. KRAMAROVA E., DRAPER G.J., MASUYER E., MICHAELIS J., *International Incidence of Childhood Cancer*. D.M. Parkin ed. Lyon, Francia, 1998.
9. WHELAN S.L., FERLAY J., RAYMOND L., YOUNG J., *Cancer Incidence in five continents*, Volumen VII. D.M. Parkin ed. Lyon, Francia, 1997.
10. NORTH AMERICAN ASSOCIATION OF CENTRAL CANCER REGISTRIES, *Cancer in North America, 1991-1995*. Vivien W. Chen, Cheng Wu Xiao, Andrews Patricia. Louisiana, 1999.
11. HOSSFELD D.K., SHERMAN C.D., LOVE R.R., BOSCH F.X., *Manual de Oncología Clínica*, quinta edición. Ediciones Doyma S.A. Barcelona, 1992.
12. *Diccionario terminológico de ciencias médicas*. Salvat editores S.A. Barcelona, 1982.

13. *Diccionario enciclopédico ilustrado de medicina Dorland*. Nueva editorial Interamericana. México D.F., 1986

14. Lista tabular de las morfologías de las neoplasias.

15. Página web [www-dep.iarc.fr/iacr.htm](http://www-dep.iarc.fr/iacr.htm)