Bioestadística descriptiva

Rigurosamente sencilla

SISTEMA DE INFORMACIÓN EN LÍNEA
SIL
COMPLEMENTO VIRTUAL

ECOE EDICIONES

Gerardo Ardila Duarte

BIOESTADÍSTICA DESCRIPTIVA

RIGUROSAMENTE SENCILLA

Gerardo Ardila Duarte

Catalogación en la publicación - Biblioteca Nacional de Colombia

Ardila Duarte, Gerardo, autor
Bioestadística descriptiva : rigurosamente sencilla / Gerardo Ardila Duarte -- Primera edición -- Bogotá : Ecoe Ediciones, 2022
190 páginas. -- (Ciencias básicas. Estadística)

Incluye datos curriculares del autor -- Incluye referencias bibliográficas

ISBN 978-958-503-398-6 -- 978-958-503-399-3 (e-book)

1. Estadística vital - Procesamiento de datos - Problemas, ejercicios, etc. 2. Probabilidades - Procesamiento de datos - Problemas, ejercicios, etc.

CDD: 570.151950285 ed. 23 CO-BoBN - a1097188

Área: *Ciencias básicas*
Subárea: *Estadística*

© Gerardo Ardila Duarte

► Ecoe Ediciones S.A.S.
info@ecoeediciones.com
www.ecoeediciones.com
Carrera 19 # 63C 32, Tel.: 919 80 02
Bogotá, Colombia

Primera edición: Bogotá, enero del 2023

ISBN: 978-958-503-398-6
e-ISBN: 978-958-503-399-3

Directora editorial: Claudia Garay Castro
Coordinadora editorial: Paula Bermúdez B.
Corrección de estilo: Daniela Pérez
Diagramación: Nicol Pulido Casallas
Carátula: Wilson Marulanda Muñoz
Impresión: Carvajal Soluciones de Comunicación S.A.S.
Carrera 69 #15-24

Prohibida la reproducción total o parcial por cualquier medio sin la autorización escrita del titular de los derechos patrimoniales.

Impreso y hecho en Colombia - Todos los derechos reservados

AGRADECIMIENTOS

"A Dios, mi familia y lectores.
A la Editorial ECOE por creer en mi proyecto y
por su apoyo profesional en correcciones y sugerencias".

CONTENIDO

Prefacio xxi

Introducción xxiii

Capítulo 1. El papel de la bioestadística 1
1.1 Introducción 1
1.2 Excel y hojas de cálculo 2
1.2.1 Iniciando Excel 2
1.2.2 Ejecución de comandos en Excel 3
1.2.3 Selección de celdas 4
1.2.4 Creación de una base de datos 4
1.2.5 Fórmulas y funciones 5
1.3 Importancia de la bioestadística 6
1.4 Oportunidades que ofrece la bioestadística 7
1.5 Aplicación de la bioestadística 7
1.6 Algunas definiciones básicas 7
1.6.1 Población y parámetros 7
1.6.2 Muestras y estadísticos 8
1.6.3 Variables 8
1.6.4 Muestras aleatorias y recolección de datos 11
1.7 Las funciones de la bioestadística 11
1.8 Escalas de medida 12

Capítulo 2. Descripción de los conjuntos de datos 21
2.1 Introducción 22
2.2 Métodos de agrupación de datos 22
2.2.1 Tabla de frecuencias 22
2.2.2 Obtención de frecuencias usando Real Statistics 31
2.2.3 Obtención de frecuencias usando R 32
2.2.4 Tabla de frecuencias con datos discretos 34
2.2.5 Tablas de contingencia 35
2.3 Gráficos 42
2.3.1 El histograma 42
2.3.2 Diagrama circular o torta 43
2.3.3 Diagrama de barras 44
2.3.4 Diagramas de Pareto o diagramas 80/20 46
2.3.5 Diagramas de caja y bigotes o *box-plot* 47
2.3.6 Diagrama de dispersión 50
2.3.7 Diagrama radial 52
2.3.8 Diagrama de embudo o pirámide 53
2.3.9 Diagrama de área, clúster o rectángulo 53
2.3.10 Gráficos mixtos 54
2.3.11 Gráficos de paso 55
2.3.12 Gráficos de mapa 56

Capítulo 3. Medidas de tendencia central y dispersión 69
3.1 Medidas de tendencia central 71
3.1.1 Media aritmética: 71
3.1.2 Mediana ($\tilde{x}$) 71
3.1.3 Moda (mo): 71
3.1.4 Media geométrica (MG): 74
3.1.5 Media armónica (H): 74
3.2 Medidas de dispersión 78
3.2.1 Rango R o recorrido 78
3.2.2 Varianza 78
3.2.3 Varianza poblacional 78
3.2.4 Desviación estándar poblacional 79
3.2.5 Error estándar de la muestra: 79
3.2.6 Distribución normal y regla empírica 82
3.2.7 Gráficas de intervalo 85
3.2.8 El error de medición Dahlberg 88
3.2.9 Kappa de Cohen 89
3.2.10 Kappa de Fleiss 91
3.2.11 Carta de control de calidad 98
3.2.12 Método de Bland Altman 99

3.3 Otras medidas de dispersión o posición 105
3.3.1 Análisis de puntos atípicos 110
3.3.2 Diagramas *box-plot* 110
3.4 Cubos OLAP (*Online Analytical Processing*) o procesamiento analítico en línea 115
3.4.1 El coeficiente de sesgo o asimetría de Pearson 117
3.4.2 El coeficiente de variación 118
3.4.3 Curtosis 118

Capítulo 4. Probabilidad 125
4.1 Principios de probabilidad 126
4.2 Experimentos, resultados y conjuntos 127
4.3 Introducción a la probabilidad 127
4.4 Reglas de probabilidad y tablas de contingencia y probabilidad 130
4.5 Tabla y curva de clasificación ROC 131
4.6 La curva ROC 134
4.7 Tablas de contingencia 136
4.8 Teorema de Bayes 142
4.9 Conteo 144
4.10 Distribuciones de probabilidad 146
4.11 Distribución de probabilidad binomial 148
4.12 Distribución de Poisson 151
4.13 Distribución multinomial 154

Capítulo 5. Algunas distribuciones continuas de probabilidad 157
5.1 Distribución normal 158
5.2 Teorema central del límite 159
5.3 Nivel de confianza $z^{\sigma}/\sqrt{n}$ 160
5.4 Aproximación normal a binomial – cálculo de probabilidades 167
5.5 Distribución Gamma y Weibull 169
5.6 La distribución exponencial 173
5.7 Distribución chi cuadrado 175
5.8 Diagramas *box-plot* 177
5.9 Diagramas QQ plot y P- plot 179
5.10 Prueba de Geary 180
5.11 La prueba de Shapiro Wilk 180

Lecturas recomendadas 187
Referencias 189
Perfil del autor 191

ÍNDICE DE FIGURAS

Figura 1. Reconocimiento de CIEO-UniCIEO nacional e internacionalmente comité de investigación 2012-2014 xxv
Figura 2. Premio obtenido en 2020 xxv
Figura 3. Hoja de cálculo de Excel 3
Figura 4. Creación de una base de datos 4
Figura 5. Aplicación de fórmulas en Excel 5
Figura 6. Aplicación de funciones en Excel 6
Figura 7. Tipos de variables 10
Figura 8. Escalas de medición 13
Figura 9. Base resistencia a la des-cementación de brackets 15
Figura 10. Diferentes variables de pacientes que han acudido a un hospital 18
Figura 11. Paso 1 de Construcción de una tabla de frecuencias con tablas dinámicas 26
Figura 12. Paso 2 de Construcción de una tabla de frecuencias con tablas dinámicas 27
Figura 13. Paso 3 de Construcción de una tabla de frecuencias con tablas dinámicas 27
Figura 14. Paso 4 de Construcción de una tabla de frecuencias con tablas dinámicas 28

Figura 15. Paso 5 de Construcción de una tabla de frecuencias con tablas dinámicas 28
Figura 16. Paso 6 de Construcción de una tabla de frecuencias con tablas dinámicas 29
Figura 17. Paso 7 de Construcción de una tabla de frecuencias con tablas dinámicas 29
Figura 18. Paso 8 de Construcción de una tabla de frecuencias con tablas dinámicas 30
Figura 19. Paso 9 de Construcción de una tabla de frecuencias con tablas dinámicas 30
Figura 20. Paso 10 de Construcción de una tabla de frecuencias con tablas dinámicas 31
Figura 21. Frecuencia simple y acumulada en Real statistics 32
Figura 22. Tabla de Frecuencias usando R 34
Figura 23. Construcción de tabla de contingencias con tres variables: espaciador, zona y longitud 39
Figura 24. Tabla de contingencia: suma de longitud en micrómetros adaptación marginal cofias 39
Figura 25. Cálculo del promedio en una tabla dinámica 40
Figura 26. Datos para elaborar la tabla de frecuencias para la variable dolor 41
Figura 27. Distribución de frecuencia margen del muñón pre láser 43
Figura 28. Histograma de frecuencias 43
Figura 29. Distribución de la distancia en micrómetros entre el margen del muñón y el margen del borde 44
Figura 30. Calidad del poste prefabricado 44
Figura 31. Histograma de la calidad del poste prefabricado 45
Figura 32. Distribución de la presencia de células en hueso de conejo con implante de dispositivo especial 45
Figura 33. Diagrama de Pareto para el análisis de formación de células de hueso de conejo implantando un dispositivo especial 46
Figura 34. Análisis de calificación atípica por *box-plot* en R 48
Figura 35. Análisis de puntos atípicos por *box-plot* en Real Statistics 49
Figura 36. Uso de los datos del ejemplo 4 para elaborar diagrama de caja y bigotes en Excel 49
Figura 37. Diagrama de caja y bigotes en Excel 50
Figura 38. Datos para la elaboración de diagramas de dispersión 50
Figura 39. Diagrama de dispersión X1, X2 y Y en R 51

Figura 40. Diagrama de dispersión con Excel 52
Figura 41. Diagrama de dispersión con Real Statistics..... 52
Figura 42. Diagrama radial en Excel..... 53
Figura 43. Diagrama de embudo..... 53
Figura 44. Diagrama de muestra, clúster o rectángulo..... 54
Figura 45. Gráfico mixto 54
Figura 46. Gráfico de pasos..... 56
Figura 47. Gráfico de mapas 56
Figura 48. Datos de ingresos de pacientes a un hospital..... 68
Figura 49. Media, mediana y moda de los espacios en micrómetros usando Excel..... 72
Figura 50. Mediana 72
Figura 51. Moda de los espacios en micrómetros usando Excel..... 73
Figura 52. Media geométrica y media armónica usando Excel..... 75
Figura 53. Varianza, desviación estándar y error típico usando Excel..... 82
Figura 54. Tabla de frecuencia de medidas crestales 83
Figura 55. Distribución de las medidas en crestal preoperatoria..... 83
Figura 56. Histograma con curva de Normalidad en Real Statistics..... 84
Figura 57. Histograma con curva de Normalidad en Real Statistics (Continuación)..... 84
Figura 58. Histograma con Real Statistics..... 85
Figura 59. Gráfica del máximo, mínimo y desviación usando Excel 87
Figura 60. Comparación medidas de calibración 87
Figura 61. Error de medición de Dahlberg, coeficiente de variación y mediciones de pruebas piloto 88
Figura 62. Coeficientes de variación..... 89
Figura 63. Kappa de Cohen..... 91
Figura 64. Kappa de Fleiss..... 93
Figura 65. Índice de kappa de Cohen 95
Figura 66. Kappa de Fleiss..... 97
Figura 67. Carta control de calidad mediciones Bolton x operador..... 99
Figura 68. Diagrama de BlandAltman en Excel..... 100
Figura 69. Diagrama de BlandAltman en R..... 101
Figura 70. Placas de 15 pacientes para evaluar lesión tuberculosa..... 104
Figura 71. Rango intercuartílico en Excel..... 110
Figura 72. Diagrama de *box-plot* 111

Figura 73. Crear un diagrama *box-plot* en R 112
Figura 74. Descargue comando rcmdr 112
Figura 75. Importar datos desde Excel en R 113
Figura 76. Menú gráficos en Rcmdr 113
Figura 77. Diagrama de caja o *box-plot* en R 114
Figura 78. Pasos para elaborar cubo de OLAP en Excel 116
Figura 79. Cubo OLAP en Excel 116
Figura 80. Coeficiente de sesgo o asimetría de Pearson 118
Figura 81. Curtosis 119
Figura 82. Diagrama de Venn con operaciones de unión e intersección 130
Figura 83. Cantidad de mosquitos que murieron y vivieron en intervalos de dosificación de 2μg 133
Figura 84. Elaboración de tabla de clasificación 133
Figura 85. Tasa Positiva Real, Tasa de Falsos Positivos, Valor Predictivo Positivo y Valor Predictivo Negativo 134
Figura 86. Creación de curva ROC 135
Figura 87. Área bajo la curva (AUC) 135
Figura 88. Árbol de decisión Bayesiano 143
Figura 89. Distribución de probabilidad de las células formadas 147
Figura 90. Distribución de probabilidad de pacientes con dolor por el uso de ARCOC 151
Figura 91. Gráfica de la distribución de Poisson 153
Figura 92. Ubicación esperada de Media, Mediana y Moda en la distribución Normal 158
Figura 93. Probabilidad de que la flexión sea superior a 50,5 MPa 162
Figura 94. Probabilidad de que la flexión sea inferior a 48 MPa 163
Figura 95. Probabilidad de que la flexión esté entre 45 y 50 MPa 163
Figura 96. Índice dpd/creatinuria utilizando R 171
Figura 97. Distribución Gamma del índice de dpd/creatinuria de los pacientes 172
Figura 98. Promedio y desviación estándar cara cervical-vestibular 176
Figura 99. Análisis de datos atípicos por diagrama de *box-plot* usando R 178
Figura 100. Diagrama QQ plot usando Real Statistics 180
Figura 101. Gráfica del estadístico de Shapiro Wilk y diagrama *box-plot* 182
Figura 102. Pasos para la elaboración de un trazo de normal P-plot en R 184

ÍNDICE DE TABLAS

Tabla 1. Clases de variables en la investigación *Comparación de la resistencia a la descementación de postes prefabricados utilizando un cemento autoadhesivo, con y sin modificación de su protocolo* 16

Tabla 2. Parte de la base de datos de la investigación *Comparación de la resistencia a la descementación de postes prefabricados utilizando un cemento autoadhesivo, con y sin modificación de su protocolo* 17

Tabla 3. Características y escalas de variable en la investigación *Evaluación de los cambios vasculares a nivel macular, peripapilar y del disco óptico por angiografía por tomografía de coherencia óptica (angio oct) en pacientes con glaucoma* 18

Tabla 4. Datos mediciones en micrómetros entre margen de muñón y margen de borde protésico 22

Tabla 5. Intervalo o ancho de clase 23

Tabla 6. Clase 23

Tabla 7. Frecuencia 24

Tabla 8. Diagrama de tallo y hoja 24

Tabla 9. Frecuencia relativa 25

Tabla 10. Frecuencia, frecuencia relativa, frecuencia absoluta acumulada y frecuencia relativa acumulada 25
Tabla 11. Base de pesos para construcción tabla de frecuencias en R 32
Tabla 12. Distribución de frecuencias de la formación de células en hueso de conejo al realizar un implante con un dispositivo especial 34
Tabla 13. Base de datos para realizar ejercicio del ejemplo 1 35
Tabla 14. Análisis de Cualidad de material 37
Tabla 15. Tabla de contingencia: calidad del poste prefabricado 37
Tabla 16. Análisis marginal de la calidad del poste prefabricado 38
Tabla 17. Análisis total fila 38
Tabla 18. Tabla de contingencia: promedio de longitud de adaptación marginal 40
Tabla 19. Tabla de frecuencias para la variable dolor 41
Tabla 20. Base calificaciones evaluación 47
Tabla 21. Datos para elaborar gráfico de pasos 55
Tabla 22. Datos para elaborar gráfico de pasos (continuación) 55
Tabla 23. Eficacia de la coliforitropina alfa versus la hormona foliculo estimulante 58
Tabla 24. Distancia en micrómetros entre el margen del muñón protésico y el margen del borde protésico en prótesis fija de tres unidades con balanceo sin ser seccionada después de ser tratada con láser 61
Tabla 25. Reacción tisular que presenta un dispositivo de anclaje temporal al ser implantado en tibia de conejo 63
Tabla 26. Análisis del movimiento causado por tres clases de cierto tipo de brackets 64
Tabla 27. Evaluación Funcional de Buxcheker 65
Tabla 28. Comparación de las características histológicas y fuerza de adhesión entre el esmalte humano y bovino 67
Tabla 29. Media aritmética 71
Tabla 30. Mediana 71
Tabla 31. Espacio en micrómetros entre margen del muñón protésico *y* y margen borde protésico en prótesis fija de tres unidades con balanceo 72
Tabla 32. Media ponderada 74

Tabla 33. Marcadores de reabsorción ósea, calcio en suero y piridinolinas en orina como predictores precoces de cambios en la densidad ósea 75
Tabla 34. Descripción de la evolución clínica y complicaciones a los 30 días de los pacientes adultos sometidos a colecistectomía 76
Tabla 35. Adaptación marginal en micrómetros de cofias en discilicato de litio utilizando dos técnicas: inyección (e-max press®) y asistida por computador 77
Tabla 36. Partes por millón de contaminación en el medio ambiente 77
Tabla 37. Estudio de la correlación entre las calibraciones oseas: clínica preoperatoria, tomográfica e intraoperatoria 80
Tabla 38. Máximo, mínimo y desviación estándar 86
Tabla 39. Análisis de 50 criminales para la valoración del kappa de Cohen 90
Tabla 40. Estado del criminal según el criterio de seis jueces 93
Tabla 41. Resultados arrojados por el observador para hallar el índice kappa de Cohen 94
Tabla 42. Kappa de Cohen usando Real Statistics 96
Tabla 43. Índice OPI de 19 pacientes para calcular la kappa de Fleiss 96
Tabla 44. Kappa de Fleiss usando Real Statistics 98
Tabla 45. Comparación cambios posición sagital y tamaño mandibular, pacientes clase II con aparato SN1 con y sin tubos telescópicos 102
Tabla 46. Evaluación de 50 pacientes y su clasificación 105
Tabla 47. Interpretación Coeficiente de kappa 105
Tabla 48. Rango intercuartílico 108
Tabla 49. Cubo OLAP de las variables Edad y Número de dientes 117
Tabla 50. Base de datos para analizar la variación de la dimensión vertical posterior 120
Tabla 51. Mediciones SMS, PMI, SPI en dos periodos diferentes 122
Tabla 52. Base de datos de la investigación para comparar la resistencia flexural de 3 técnicas de soldadura: soplete, arco y láser 123
Tabla 53. Niveles de glucosa en ayunas de una muestra de 12 pacientes 124
Tabla 54. Información sobre pacientes que bruxan o no 128
Tabla 55. Modelo de probabilidad para cada uno de los experimentos 129

Tabla 56. Prótesis fijas requeridas por pacientes que asisten a un consultorio 129
Tabla 57. Tabla de clasificación para curva ROC 132
Tabla 58. Tabla de contingencia y tabla de probabilidad 136
Tabla 59. Caracteristicas de las trabéculas formadas en conejos al dividirlos en grupos según la aplicación control con citogel y hueso más citogel 137
Tabla 60. Resumen pre valencia Características trabécula vs Grupos 139
Tabla 61. Clasificación sobrepeso y obesidad 141
Tabla 62. Distribución de probabilidad células formadas 146
Tabla 63. Cantidad implantes efectuados por residente en dos semestres 148
Tabla 64. Propiedades mecánicas de los materiales usados en la investigación sobre la distribución de esfuerzos y deformaciones 152
Tabla 65. Aplicación de la distribución de Poisson 153
Tabla 66. Datos de flexión y tracción en MPa 160
Tabla 67. Coeficiente de curtosis o achatamiento 161
Tabla 68. Construcción de cubo OLAP para la flexión 166
Tabla 69. Datos de la investigación Determinación de marcadores de reabsorción ósea, calcio en suero y piridinolinas en orina como predictores tempranos de cambios en la densidad ósea... 170
Tabla 70. Tablas de frecuencias para determinar la adaptación de la cara cervical vestibular 176
Tabla 71. Prueba Chi-sq para evaluar normalidad 177
Tabla 72. Datos adaptación marginal e interna en coronas de disilicato de litio 177
Tabla 73. Estadístico de Shapiro Wilk 181
Tabla 74. Estimación del estadístico de Geary 182

Sistema de Información en Línea

Al final del libro encontrará el código para ingresar al **Sistema de información en Línea** - SIL - donde podrá descargar material complementario como artículos y evaluaciones e ingresar canal de YouTube "Bioestadística Sin Lágrimas", en el cual se resuelven los ejercicios paso a paso, investigaciones e incluso temas que saldrán en ediciones posteriores y ejercicios solicitados por los seguidores de todo el mundo. Aprovecho para agradecer a la comunidad iberoamericana especialmente chilena, peruana, mexicana y española, pues hay bastante investigador y/o estudiante que me sigue y solicita ayuda, la que se ofrece con mucho gusto.

PREFACIO

El supuesto de este libro se basa en un enfoque esencialmente no matemático con el fin de evitar los rigurosos desarrollos algebraicos que podrían atemorizar a un estudiante de odontología, ciencias médicas o biológicas que empleen la rigurosidad matemática. Sin embargo, en varios apartados será necesario emplear algunos elementos básicos para explicar un procedimiento o una demostración. Por otra parte, en el libro se aplicarán datos reales junto con sus respectivos procedimientos, así como el análisis de diseño de experimentos, de modo que los estudiantes, profesionales o investigadores puedan aplicar los métodos estadísticos, aunque se encuentren en proceso de aprendizaje. Además, se incorporarán técnicas que satisfacen las necesidades de un investigador.

Los ejemplos y ejercicios son tomados en su mayoría de casos reales de investigación desarrollados por odontólogos en sus diferentes especializaciones del Centro de Estudios e Investigaciones Odontológicas, además de los médicos investigadores de la Fundación Santa Fe de Bogotá, y el Colegio Odontológico Colombiano UNICOC.